DISS. ETH NO. 27620

IDENTIFYING THE CHEMICAL NATURE OF SOIL ORGANIC PHOSPHORUS WITH INCREASING MOLECULAR WEIGHT

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES

(Dr. sc. ETH Zurich)

presented by

JOLANDA ELENA REUSSER

MSc ETH Zurich in Environmental Engineering

born on 28.03.1990

citizen of Heiligenschwendi BE, CH

accepted on the recommendation of

Prof. Dr. Emmanuel Frossard

Dr. Timothy McLaren

Prof. Dr. Ruben Kretzschmar

Dr. Ronald Smernik

2021

Table of contents

List of abbreviations

ASL	Above sea level
CPMAS	Solid-state cross polarisation magic angle spinning
CPMG	Carr-Purcell-Meiboom-Gill pulse sequence
EDTA	Ethylenediaminetetraacetic acid disodium salt
HPLC	High-pressure liquid chromatography
HSQC	Heteronuclear single quantum correlation
ICP-MS	Inductively coupled plasma-mass spectrometry
ICP-OES	Inductively coupled plasma-optical emission spectrometry
IP	Inositol phosphates
K_{av}	Partition coefficient
MDP	Methylenediphosphonic acid
MRP	Molybdate reactive phosphorus
MUP	Molybdate unreactive phosphorus
NMR	Nuclear magnetic resonance
P	Phosphorus
P_{inorg}	Inorganic phosphorus
P_{org}	Organic phosphorus
P_{tot}	Total phosphorus
SCF	Sequential chemical fractionation
SD	Standard deviation
SDF	Spectral deconvolution fitting
SEC	Size exclusion chromatography
SOM	Soil organic matter
T_1	Longitudinal relaxation time (spin-lattice)
T_2	Transverse relaxation time (spin-spin)
UT	Untreated soil sample resp. untreated soil extract
UV-VIS	Ultraviolet-visible
XRD	X-ray diffraction
XRF	X-ray fluorescence

Abstract

Phosphorus (P) is essential for all living organisms because it is part of cell components and involved in major biochemical processes. Phosphorus exists in soils either as inorganic (P_{inorg}) or organic (P_{org}) forms. In general, pools of P_{org} comprise between 20 % to 80 % of total soil P. Identifiable P_{org} molecules include: phosphomonoesters, such as inositol phosphates (IP); phosphodiesters, such as DNA; and phosphonates, such as 2-aminoethylphosphonic acid. However, these identifiable molecules usually comprise on average 30 % of the total P_{org} pool in soil, the chemical nature of the majority of P_{org} in soil remains unresolved. This pool of 'unresolved' P_{org} can be estimated as an underlying broad signal in the phosphomonoester region of solution ^{31}P nuclear magnetic resonance (NMR) spectra on soil extracts. Little is known about the chemical composition of the broad signal but it appears to be of apparent large molecular weight and complex structure.

The overall aim of this PhD project was to identify the chemical nature of soil P_{org} with increasing molecular weight by combining several chemical extraction approaches, size separation procedures, and solution ^{31}P NMR spectroscopy techniques. Six topsoil samples of diverse soil types (a Ferralsol, a Vertisol, three Cambisols, and a Gleysol) and land-uses (arable, grassland, and forest), which had concentrations of P_{org} ranging from 70 to 1135 mg P/kg_{soil}, were analysed for this project.

In the first chapter, the two most commonly applied spectral deconvolution fitting (SDF) approaches of NMR soil spectra were compared to accurately quantify *myo*-IP_6 in soil extracts. In the first approach, sharp peaks in the phosphomonoester region were fitted to the baseline of the spectrum, whereas in the second approach, an underlying broad signal was included in the SDF process. The hypothesis was that carrying out SDF on the phosphomonoester region without a broad signal would result in an overestimation of *myo*-IP_6 in soil extracts. The results show that the average recovery of an added *myo*-IP_6 standard to soil extracts was 122 % (SD 32) and 95 % (SD 5) using SDF without and with an underlying broad signal, respectively. Therefore, an underlying broad signal needs to be included when carrying out SDF of the phosphomonoester region for accurate quantification of *myo*-IP_6, and for other sharp signals in this region (e.g. *neo*-IP_6).

High-resolution solution ^{31}P NMR spectra on soil extracts revealed a plethora of sharp signals of unknown identity. The hypothesis of the second chapter was that these sharp peaks mainly arise from lower-order IP (IP_1-IP_5), because of their previous detection in soil samples using chromatographic approaches. Hypobromite oxidation on alkaline soil extracts was carried out in order to oxidise all organic matter except IP. Subsequent NMR analyses of the extracts revealed the presence of four stereoisomers of IP_5, two isomers

of *myo*-IP_5 and *scyllo*-IP_4. These lower-order IP comprised 9 % to 50 % of the total soil IP pool. In addition, on average half of the underlying broad signal resisted hypobromite oxidation. Transverse relaxation experiments confirmed the differing molecular composition of these unresolved P_{org} compounds compared to IP, and suggested the presence of P_{org} in large molecular weight material.

The molecular size distribution of soil P_{org} was further investigated in the third chapter by combining size exclusion chromatography (SEC) with solution ^{31}P NMR spectroscopy. The hypothesis was that phosphomonoesters in large molecular weight material are dominated by multiple broad signals in the phosphomonoester region. The results revealed that these broad signals (on average 3-4) were most prominent in the 10-20 and 20-50 kDa fractions, where they comprised on average 77 % resp. 74 % of total phosphomonoesters across all soils. Hence, the findings indicate that the unresolved pool of P_{org} is comprised of multiple components rather than a single macropolymeric structure. Furthermore, the proportion of IP detected in molecular size fractions above 10 kDa was on average 23 % of the total IP pool. The presence of these P_{org} compounds in large molecular size material suggests a close association with the soil organic matter (SOM).

In the last chapter, the SOM of the Gleysol was fractionated according to its bonding types using the 'Humeomics' sequential chemical fractionation procedure. This was done in order to investigate the association of P_{org} within the SOM structure. The findings of this study revealed that 47 % of the unresolved pool of P_{org} was associated with the SOM structure. In addition, two pools of IP could be identified: one closely associated with the SOM (45 %) and the second being closely associated with the soil mineral phase (40 %). Especially the latter IP pool was highly underestimated using a single-step NaOH-EDTA extraction, which only extracted 58 % of total IP.

In conclusion, the chemical nature of soil P_{org} is much more diverse and complex than previously thought. Furthermore, we revealed that the unresolved P_{org} pool represented by broad signals is comprised of a range of smaller components, which are associated with the SOM through various bonding types. The identification of the chemical nature of these P_{org} compounds as well as their association types will help to improve our understanding on their dynamics in terrestrial (and aquatic) ecosystems. Consequently, soil management and agricultural strategies can be developed for using these P_{org} pools as a source of P for plant nutrition via mineralisation processes.

Zusammenfassung

Das Element Phosphor (P) ist ein essentieller Bestandteil der Zellen von lebenden Organismen und ist in den wichtigsten biochemischen Prozessen involviert. In Böden existiert P entweder als Bestandteil inorganischer (P_{inorg}) oder organischer Verbindungen (P_{org}), wobei letztere zwischen 20 % und 80 % des Gesamtphosphors ausmachen können. Identifizierte organische P-Verbindungen sind: Phosphomonoester, z. B. Inositolphosphate (IP); Phosphodiester, z. B. DNS; und Phosphonate, z. B. Aminomethylphosphonsäuren. Diese identifizierbaren P_{org}-Verbindungen umfassen jedoch durchschnittlich nur 30 % des totalen P_{org}-Pools in Böden; die chemische Zusammensetzung des restlichen P_{org} bleibt ungeklärt. Das Ausmass dieses unidentifizierten P_{org}-Pools in Böden kann durch ein verbreitertes Signal in der Resonanzregion von Phosphomonoestern bei der ^{31}P-Kernspinresonanzspektroskopie (NMR für Nuclear Magnetic Resonance) an Bodenextrakten angenähert werden. Über die chemische Zusammensetzung der Moleküle, welche dieses verbreiterte Signal produzieren, ist wenig bekannt. Jedoch scheinen sie von ausgedehnter Molekülgrösse und komplexer Struktur zu sein.

Die übergeordnete Zielsetzung dieses PhD-Projektes war die Identifizierung der chemischen Natur von P_{org} im Boden mit zunehmender Molekülgrösse, wofür verschiedene chemische Extraktionsmethoden, Grössenauftrennungs-Verfahren und ^{31}P-NMR-Spektroskopie an flüssigen Proben kombiniert wurden. Im Rahmen des Projektes wurden Oberboden-Horizonte von sechs Böden untersucht, welche verschiedene Bodentypen (ein Ferralsol, ein Vertisol, drei Cambisols und ein Gleysol), sowie verschiedene Landnutzungen (Ackerbau, Weideland und Waldboden) umfassen. Die totalen Gehalte von P_{org} in diesen Böden reichten von 70 bis 1135 mg P/kg_{Boden}.

Das erste Kapitel dieser Thesis behandelt den Vergleich der zwei am häufigsten angewandten spektralen Dekonvolutions-Methoden (SDM) im Hinblick auf die akkurate Quantifizierung von *myo*-IP_6 in Bodenextrakten. Bei der ersten Methode werden in der Resonanzregion von Phosphomonoestern schmale Signale bis zu der Grundlinie der Spektren gefittet. In der zweiten Methode wird zusätzlich ein verbreitertes Signal im Fitting-Prozess miteinbezogen, welches unter den schmalen Signalen liegt. Die Hypothese war, dass die erstere SDM ohne Berücksichtigung des unterliegenden, verbreiterten Signals die Konzentration von *myo*-IP_6 in Bodenextrakten überschätzt. In der Tat, die Resultate dieser Studie zeigen auf, dass die Konzentration einer zu den Bodenextrakten zugefügten *myo*-IP_6 Standardlösung zu 122 % (SD 32) der eigentlichen Konzentration berechnet wurde, wenn dieses unterliegende, verbreiterte Signal nicht in die SDM integriert wurde.

Mit Einbezug dieses Signales in die SDM betrug die berechnete Konzentration 95 % (SD 5) der eigentlichen Konzentration der Standardlösung. Basierend auf diesen Resultaten wird aufgezeigt, dass ein unterliegendes, verbreitertes Signal in die SDM der Phosphomonoester-Region integriert werden sollte, um *myo*-IP_6 und andere Signale in dieser Region, welche z. B. von *neo*-IP_6 produziert werden, akkurat quantifizieren zu können.

Höher aufgelöste ^{31}P NMR Spektren von Bodenextrakten zeigen eine Vielzahl von schmalen Signalen in der Resonanzregion von Phosphomonoestern, welche bis anhin noch nicht identifiziert wurden. Die Hypothese des zweiten Kapitels dieser Thesis war, dass diese Signale von tiefer wertigen IP (IP_1-IP_5) produziert werden, da diese bereits in Bodenextrakten mittels Chromatographie nachgewiesen wurden. Dazu wurde die gesamte organische Substanz ausser IP in alkalischen Bodenextrakten mittels Hypobromit-Oxidation oxidiert. Anschliessende NMR-Analysen der aufbereiteten Extrakte resultierte in der Identifikation von vier Stereoisomeren von IP_5, zwei Isomeren von *myo*-IP_5, sowie *scyllo*-IP_4. Diese tiefer wertigen IP machten insgesamt 9 % bis 50 % des Gesamtgehaltes an IP in den untersuchten Böden aus. Zudem konnte aufgezeigt werden, dass durchschnittlich die Hälfte des unterliegenden, verbreiterten Signals nicht oxidiert wurde. Experimente mit der transversalen Relaxationszeit bestätigten, dass die molekulare Zusammensetzung dieses unidentifizierten P_{org}-Pools sich von jener der IP unterschied. Zudem deuteten die Resultate an, dass dieser Pool von ausgedehnter Molekülgrösse ist.

Die molekulare Grössenverteilung von P_{org} in Böden wurde im dritten Kapitel weitergehend untersucht, indem Grössenausschluss-Chromatographie (SEC) mit ^{31}P NMR Spektroskopie an Bodenextrakten kombiniert wurden. Die Hypothese wurde aufgestellt, dass Phosphomonoester mit grossen Molekulargewichten resp. -grössen von mehreren (durchschnittlich 3-4) verbreiterten Signalen in NMR-Spektren dominieren. Die Resultate zeigen auf, dass im Bereich dieser verbreiterten Signale die Molekülgrössen-Fraktionen von 10-20 und 20-50 kDa am meisten vorkommen. In diesen Fraktionen waren im Durchschnitt 77 % resp. 74 % aller gemessenen Phosphomonoester Teil dieser verbreiterten Signale. Die Ergebnisse dieser Studie deuten stark darauf hin, dass der unidentifizierte P_{org}-Pool aus mehreren Komponenten unterschiedlicher Grösse besteht, anstatt einer einzigen Komponente von makropolymerischer Struktur. Zusätzlich wurde festgestellt, dass der Anteil von IP in Fraktionen mit Molekulargrössen über 10 kDa durchschnittlich 23 % des totalen IP-Pools betrug. Diese Präsenz von IP in Material von

ausgedehnter Molekülgrösse weist auf eine enge Verbindung mit der organischen Bodensubstanz (SOM) hin.

Im letzten Kapitel dieser Thesis wurde die SOM des Gleysols mithilfe der sequentiellen chemischen Fraktionierungsmethode "Humeomics" in ihre Bestandteile aufgeteilt. Das Ziel dieser Fraktionierung war, die unterschiedlichen Verbindungen des P_{org} mit der SOM-Suprastruktur zu untersuchen. Dies resultierte im Ergebnis, dass 47 % des unidentifizierten P_{org}-Pools mit der SOM-Suprastruktur verbunden war. Ausserdem konnten zwei unterschiedliche IP-Pools identifiziert werden: Einer verbunden mit der SOM durch schwache Esterverbindungen (45 %) und der Andere hauptsächlich verbunden mit der mineralischen Phase des Bodens (40 %). Mit einer direkten, einfachen NaOH-EDTA Extraktion wurde letzterer Pool stark unterschätzt, da nämlich nur 58 % des totalen IP-Gehaltes extrahiert werden konnte.

Zusammenfassend kann festgehalten werden, dass die chemische Natur von P_{org} im Boden viel differenzierter und komplexer ist als ursprünglich angenommen. Der bis anhin unidentifizierte P_{org}-Pool, welcher durch ein verbreitertes Signal repräsentiert wird, besteht aus mehreren Komponenten, die durch unterschiedliche Verbindungen mit der SOM assoziiert sind. Die Identifizierung dieser Moleküle sowie deren Verbindungsart mit der SOM wird unser Verständnis über die Dynamik dieser P_{org}-Komponenten in terrestrischen und aquatischen Ökosystemen vertiefen. Dadurch können Bodenbewirtschaftungs-Praktiken und landwirtschaftliche Strategien im Hinblick auf die Mineralisierung dieses P_{org}-Pools zwecks Pflanzenernährung optimiert werden.

General introduction

Phosphorus – essential for living organisms

Phosphorus (P) is involved in important biochemical processes in living organisms, e.g. the assimilation of carbon (C) into organic molecules by ribulose bisphosphate carboxylase in the Calvin cycle of photosynthesis, the energy transfer by nicotinamide adenine dinucleotide (NAD^+), the energy conservation by adenosine triphosphate (ATP) and is part of the cell genome, e.g. DNA (Madigan et al., 2012; Schubert, 2018). Furthermore, P is a key element of the cell wall as well as the genome, comprising 2.5 % of the cell dry weight (Madigan et al., 2012).

Phosphorus is primarily sourced by plants from the soil, where it can be the most limiting nutrient for plant growth and needs to be replenished by fertilisation in agricultural systems (Bünemann et al., 2010). However, despite its importance as an essential element for plants, an oversupply of P can threaten the balance of aquatic ecosystems due to P losses from soil into water bodies, causing eutrophication.

Phosphorus in soil

Plants take up P via the root-system as $H_2PO_4^-$ and HPO_4^{2-} ions from the soil solution (Blume et al., 2009). These ions can originate from a variety of organic (P_{org}) and inorganic (P_{inorg}) compounds, which are either already present in the soil system or are added via fertilisation, plant and animal residues/manure, atmospheric deposition, deposition of eroded soil material and/or weathering of P containing minerals in the bedrock (Tiessen, 2008). Total concentrations of P in soils range from less than 100 to over 3000 mg P/kg_{soil} (Condron et al., 2005).

Organic phosphorus forms

Between 20 % and 80 % of total P in soil exists in an organic form, mostly as phospho-monoesters (e.g. inositol phosphates (IP)), phosphodiesters (e.g. nucleic acids and phospholipids) and phosphonates (e.g. 2-aminoethyl phosphonic acids) (Anderson, 1980; Harrison, 1987; McLaren et al., 2020). The phosphate group in phosphomonoesters and phosphodiesters is linked to the organic moiety via one (C–O–P) resp. two (C–O–P–O–C) oxygen atoms. In contrast, phosphonates consist of P directly bound to the organic moiety (C–P). In a wider definition, condensed forms of P, such as polyphosphates and

pyrophosphates (George et al., 2018; McLaren et al., 2020), and phosphate associated with organic matter via a metal linkage (Gerke, 2010), are sometimes referred to as P_{org} because of their presence within living organisms or organic substances.

However, the chemical nature of about half of P_{org} compounds in soil remains 'unresolved' (McLaren et al., 2020). Without knowledge on the chemical structure, the chemical and physical properties of these compounds remain unknown, leaving a gap in the understanding of the cycling, transformation and bioavailability of a major proportion of P in terrestrial and aquatic ecosystems.

Inositol phosphates

Inositol phosphates are small identifiable biomolecules containing P. These compounds contribute to cellular functions, such as ion-regulation or signalling processes (Irvine and Schell, 2001). Furthermore, IP are constituents of a certain group of phospholipids (i.e. phosphoinositides), which are present in many animal and plant tissues (Cosgrove and Irving, 1980). In plant seeds, IP can represent up to 90 % of the organically-bound P, which act as an important P storage for germination (Cosgrove and Irving, 1980).

Inositol phosphates consist of a cyclohexanehexol ring (inositol) with one (IP_1) up to all six carbons (IP_6) being phosphorylated (Cosgrove and Irving, 1980). The spatial arrangement of these phosphate groups on the inositol ring can vary, forming in total nine stereo-isomeric forms: *myo*-, *scyllo*-, *neo*-, L-*chiro*-, D-*chiro*, *epi*-, *muco*-, *allo*- and *cis*-IP_6 (Turner et al., 2002). In addition, the phosphate groups can change their spatial arrangement from axial to equatorial and vice versa depending on pH, resulting in a ring flip from one chair conformation of the inositol ring to another (Turner et al., 2007b). These two different spatial arrangements of one stereoisomer are called conformers. In contrast to IP_6, lower-order IP (IP_1 to IP_5) are present as isomers, where the phosphate groups are attached to different carbons of the inositol ring e.g. *myo*-(1,2,4,5,6)-IP_5 and *myo*-(1,3,4,5,6)-IP_5.

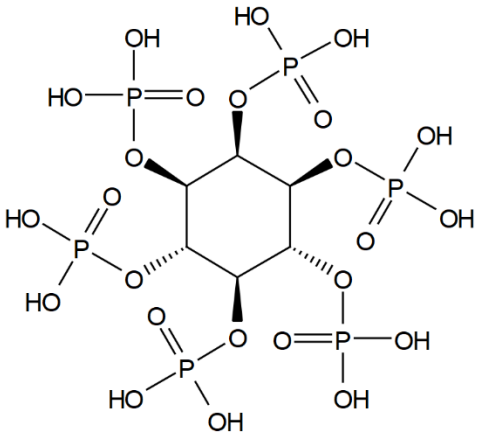


Figure 1. Chemical structure of *myo*-IP_6, drawn using ChemDraw Professional 17.0 PerkinElmer Inc.

So far, only the *myo*-, *scyllo*-, *neo*- and D-*chiro*-isomers have been detected in soils, with the *myo*-isomer of IP_6 (Figure 1) being the most abundant form (Cosgrove and Irving, 1980). A common name for *myo*-IP_6 is phytic acid (Turner et al., 2007b) and its salts are also known as phytate (Turner et al., 2002). Most input of IP into the soil system originates from manures and plant material containing almost exclusively the *myo*-isomer of IP_6, explaining the high abundance of this isomer in soil (Turner, 2007). Furthermore, plant litter can contain the *chiro*-isomer but to a much lesser extent (L'Annunziata, 2007). Indeed, other isomers than *myo* are extremely rare in biological tissues but occur widely in soils, especially *scyllo*-IP_6 (Turner, 2007). Microbial epimerisation of *myo*-IP was proposed as the origin of these isomers in soil (L'Annunziata et al., 1977; L'Annunziata, 2007). However, this process still needs to be proven to happen in soils and result in the presence of isomers other than *myo*. Other possible processes could include the phosphorylation of lower-order IP or direct biological synthesis of the isomers in soil by microorganisms (Cosgrove, 1969; Cosgrove and Irving, 1980; Turner et al., 2002).

Historically, phytate was considered to be one of the most abundant and persistent forms of P_{org} in soils (Potter and Snyder, 1918; Auten, 1923; McLaren et al., 2020). Dyer et al. (1940) documented the first direct isolation of phytin from soil (the common name for the calcium (Ca) and magnesium (Mg) salt of *myo*-IP_6) (Turner et al., 2002). Later, Wrenshall and Dyer (1941) confirmed the presence of phytin in soils and investigated the action of phytase, the enzyme which liberates the phosphate group from the inositol ring (Turner et al., 2002), on a range of metal-phytate compounds. The authors found that in contrast to Na-phytate, Fe-phytate and Al-phytate were mostly ‘immune’ to the enzymatic attack, possibly increasing their stability in soils. However, the authors suggested that the primary requirement for stability in soil is that the IP compounds are retained insoluble, either by sorption, complexation or precipitation (Celi and Barberis, 2007).

In the 1950s to 1970s, the combination of chromatographic techniques with chemical extraction and isolation methods (i.e. hypobromite oxidation) advanced the knowledge on the diversity and abundance of IP in soils (McLaren et al., 2020). Smith and Clark (1951) reported the presence of lower-order IP (IP_5) in soil extracts, which was already suggested by Yoshida (1940) on the case of IP_1 indirectly measured in Hawaiian soils. Further studies revealed the presence of all four stereoisomers found in soil as well as lower-order IP ranging from IP_1 to IP_5 (Halstead and Anderson, 1970; Anderson and Malcolm, 1974; Cosgrove and Irving, 1980; Irving and Cosgrove, 1982). Measured concentrations of the IP pool comprised on average 30 % of total P_{org}, with a range from less than 1 % to 60 %,

and up to 90 % in some organic soils (Cosgrove and Irving, 1980; Harrison, 1987; McLaren et al., 2020). Hence, these compounds comprise an important pool of P_{org} in soils, which is considered to accumulate and are relatively stable (Turner et al., 2002).

Furthermore, studies on the molecular size distribution of soil P_{org} reported that IP were present in large molecular size material (Moyer and Thomas, 1970; Hong and Yamane, 1981; Borie et al., 1989), despite their actual molecular size of less than 1 kDa. The authors suggested that IP bound to larger molecules, e.g. phosphoinositides, and/or the association with complex structures of organic matter caused their presence in the large molecular size material. However, the abundance of IP associated with SOM and the mechanisms involved in the formation of these associations are not yet fully understood.

With the advent of nuclear magnetic resonance spectroscopy (NMR), it was possible to directly and simultaneously detect a diversity of P_{org} compounds in soil extracts without the need of prior purification and isolation (Newman and Tate, 1980; Doolette and Smernik, 2011). The NMR technique is described later. Newman and Tate (1980) were the first to report solution ^{31}P NMR spectra on alkaline extracts of five soils originating from New Zealand. The authors were able to distinguish the broad P_{org} classes phosphonates, inorganic orthophosphate, phospho-monoesters and -diesters as well as polyphosphates (Figure 5). Furthermore, comparison of the soil spectra with the ^{31}P NMR spectrum of a *myo*-IP_6 standard solution revealed that this compound was present in the soil extracts. Advances in the spectroscopic technique as well as in the extraction procedure led to enhanced resolution and sensitivity. Cade-Menun and Preston (1996) reported that a mixture of sodium hydroxide (NaOH) with ethylenediamine-tetraacetic acid (EDTA), which chelates interfering paramagnetic ions in the extract, improved the detection of P_{org} compounds. Under these conditions, IP are deprotonated but otherwise stable, making them highly soluble and enhancing their detection in the soil extracts by NMR spectroscopy (Turner et al., 2003a; Doolette and Smernik, 2018). Using this technique combined with prior removal of other organic compounds by hypobromite oxidation, it was possible to detect all four stereoisomers in soil extracts (Turner et al., 2003c; Turner and Richardson, 2004; Turner et al., 2012). However, most research using solution ^{31}P NMR soil spectroscopy focused on the detection of *myo*- and *scyllo*-IP_6. The lack of studies reporting lower-order IP in solution ^{31}P NMR soil spectra is mainly due to limitations in the resolution and sensitivity of NMR instruments as well as the low abundance of these compounds and their occurrence in the phosphomonoester region, resulting in their concealment by

overlapping peaks. Thus, research on lower-order IP has been neglected in the past two decades, despite their confirmed presence in soil using chromatography.

The phosphomonoester region of solution ^{31}P NMR soil spectra usually exhibit a plethora of overlapping peaks (Doolette and Smernik, 2011). These peaks need to be separated by spectral deconvolution fitting (SDF) in order to calculate the net peak area and subsequently quantify IP and other phosphomonoesters. The SDF has commonly be carried out by fitting the peaks to the baseline of the spectra as first proposed by Turner et al. (2003c). This procedure, which is still widely in use (Hill and Cade-Menun, 2009; Deiss et al., 2016), focuses on the identification and quantification of sharp peaks arising from small molecules containing P. Larger, P-containing molecules that resonate in the phosphomonoester region and produce broad peaks are likely to be 'overlooked' using this SDF approach (Doolette and Smernik, 2011).

Smernik and Dougherty (2007) introduced another method for SDF for quantification of phytate in soil based on subtracting the ^{31}P NMR spectrum of an unspiked soil extract (phytate spike) from the spiked soil extract. The resulting concentrations of phytate in the four investigated soils originating from south-eastern Australia were less than 5 % of total P_{org}. The study underlined the importance of spiking with standard compounds for correct peak assignment as well as the adaption of the baseline for accurate quantification of phytate. Dougherty et al. (2007) reported the presence of an underlying broad feature in the phosphomonoester region of ^{31}P NMR soil spectra, which was then integrated in the SDF procedure by Bünemann et al. (2008b) (Figure 2). By comparing the two SDF approaches of Bünemann et al. (2008b) and Turner et al. (2003c), Doolette et al. (2010) revealed that concentrations of *myo*-IP_6 were overestimated by 54 % when an underlying broad signal was not integrated in the SDF procedure of the phosphomonoester region of ^{31}P NMR soil spectra. These studies proposed that the abundance of especially *myo*-IP_6 in soil had been overestimated. However, these studies were carried out on just a few soil samples not representing the whole range of P_{org} concentrations and compositions in soil. Furthermore, the chemical nature of the P_{org} compounds causing the underlying broad signal, which contributed to the majority of the NMR signal, was unknown.

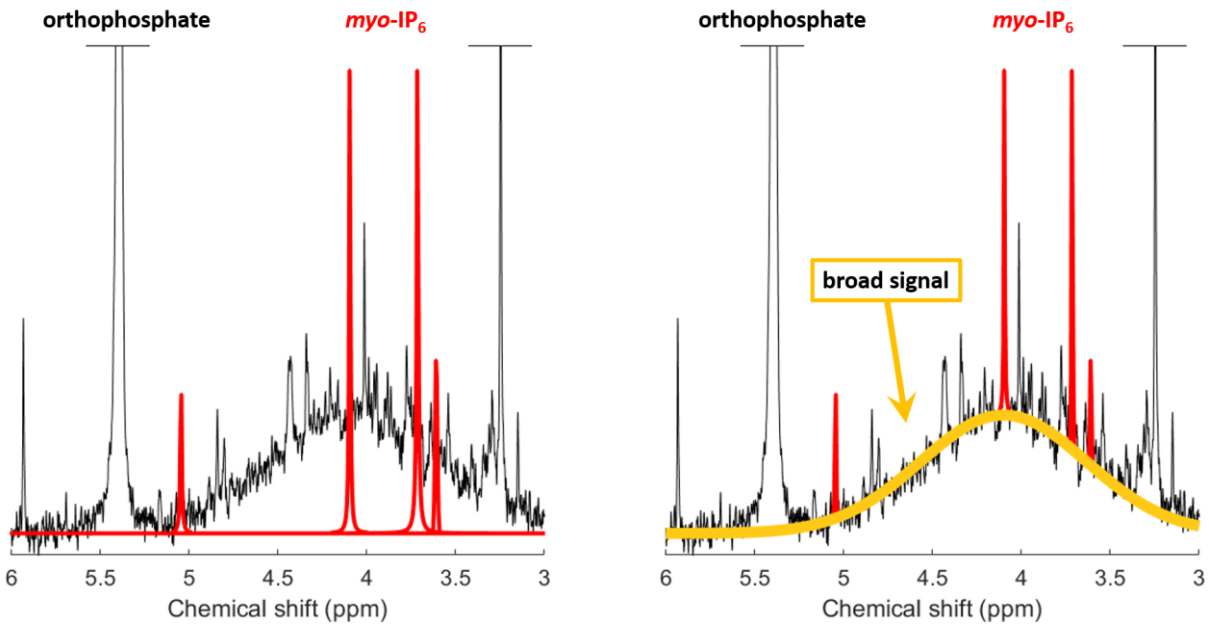


Figure 2. Phosphomonoester region of a solution ^{31}P NMR NaOH-EDTA soil spectrum. The SDF approach introduced by Turner et al. (2003c), who fitted the peaks of *myo*-IP$_6$ (in red) to the baseline of the spectrum, is illustrated on the left. The SDF approach including an underlying broad signal (yellow) introduced by Bünemann et al. (2008b), is illustrated on the right. The figure is modified from Reusser et al. (2020a).

The unresolved organic phosphorus pool

Already Auten (1923) proposed the existence of a soil P_{org} pool containing compounds other than biomolecules directly originating from organisms, such as e.g. phospholipids, phytate and nucleic acids, but formed by secondary synthesis processes in the soil. This hypothesis was based on the divergence between calculated total P contents in living organisms of the soil-plant systems and observed total P_{org} contents in soil. Furthermore, the author suggested that the majority of P_{org} in soil was associated with a 'humus-like' substance that became more complex over time and that phospholipids, phytate and nucleic acids were only minor components of the total soil P_{org}.

These findings on the existence of a complex pool of P_{org} in soil were further supported in the era of chromatography as defined by McLaren et al. (2020), where several studies reported that the majority of P_{org} in soil was found in large molecular size fractions of soil extracts. The association of P with fractions of high molecular weight material could not solely be explained by the existence of the to this date identified simple P compounds in soil (Thomas and Bowman, 1966). Therefore, Thomas and Bowman (1966) hypothesised that either unknown high molecular weight P_{org} compounds exist in soil or that the simple compounds are associated with other compounds of high molecular weight. Later, Moyer

and Thomas (1970) found that 36 % of extractable P_{org} was present in molecular weight fractions above 50 kDa, which did not contain simple compounds such as IP. The authors stated that the chemical nature of the P compounds in this large molecular size fraction has not been resolved. Steward and Tate (1971) reported that the majority of alkaline extractable P_{org} from soil was present in molecular weight fractions above 30 kDa. This P_{org} fraction did not adsorb onto an anion-exchange resin and partially degraded with the use of acid, suggesting that it consisted of phosphorylated macromolecules. The authors proposed that this large molecular size P_{org} could originate from phosphorylated uronic acid which is part of a polysaccharide fraction.

Anderson (1980) concluded that only a small percentage of soil P_{org} had been characterised. In the best case, 'only' 40 % remained to be identified, most of which occurring in very high molecular weight material. Similarly, Dalal (1977) summarised that the chemical nature of only half of soil P_{org} was resolved, in which the IP pool dominated. Furthermore, the author stated that not only little was known about the chemical nature of P_{org} but also about the processes and mechanisms involved in the incorporation of P in soil organic matter (SOM). The authors emphasised the importance of understanding the processes governing the P turnover in SOM, because the P supply by SOM is significant especially in pastures and arable crops in warmer climates.

With the advent of ^{31}P NMR spectroscopy for P_{org} speciation (Newman and Tate, 1980), studies on the P_{org} characterisation in soil humic acid extracts using NMR spectroscopy (Ogner, 1983; Bedrock et al., 1994) were consistent with findings on the unresolved P_{org} pool in high molecular weight material obtained during the chromatographic era as defined by McLaren et al. (2020). Ogner (1983) reported that a number of phosphomonoesters were present in the ^{31}P NMR spectra of high molecular weight material (>12000 m. w. u.) of humic acids prepared from four Norwegian forest soils, with the phosphomonoester region apparently being dominated by a broad feature (McLaren et al., 2020). Bedrock et al. (1994) compared the ^{31}P NMR spectra of an agricultural mineral soil with a blanket peat. The authors reported that the peaks in the phosphomonoester region were comparable between the two samples but less pronounced in the peat sample. The NMR spectra appear to exhibit a broad feature in the phosphomonoester region similarly to the study of Ogner (1983), but this could also be attributed to the low resolution of the NMR spectra for the phosphomonoester region, as stated by Bedrock et al. (1994). With time, the resolution and sensitivity of NMR spectroscopy greatly advanced but the broad feature still appeared

to be present in the phosphomonoester region of ^{31}P NMR spectra of humic substances extracts (He et al., 2006), providing more evidence for its existence (McLaren et al., 2020).

With ^{31}P NMR spectroscopy being the most commonly used technique for P_{org} characterisation, the focus of soil P_{org} research shifted away from the identification of the unresolved P_{org} pool in high molecular weight material towards the speciation of known biomolecules containing P, such as IP, DNA; phospholipids and phosphonates (Vestergren et al., 2012; Cade-Menun, 2015). The presence of an unresolved P_{org} pool was not introduced in the interpretation of solution ^{31}P NMR soil spectra until the study of Dougherty et al. (2007), who reported the existence of a broad feature in the phosphomonoester region ranging from chemical shifts δ 6 to 4 ppm. While phytate was removed by an HF treatment, the broad feature partially resisted the HF treatment. The authors proposed that this broad feature could represent P_{org} complexed in the organic matter matrix. As reviewed previously in the IP section, an underlying broad signal representing the unresolved P_{org} pool was integrated in the SDF approach by Bünemann et al. (2008b) (Figure 2). The authors found that whilst this broad signal accounted for 27 % to 47 % resp. 40 % to 71 % of total phosphomonoesters in a Chromic Luvisol resp. Calcisol (WRB, 2014), it was absent in model soils created by a mixture of 60 % sand and 40 % clay. The absence of the broad signal in the model soils was attributed to the absence of complex organic molecules formed during extended humification, which did not take place in the model soils.

Further research on P_{org} characterisation in soil extracts using solution ^{31}P NMR spectroscopy, which included an underlying broad signal in the SDF procedure, revealed that this signal dominated the high molecular weight fraction above 5 kDa (Jarosch et al., 2015) resp. above 10 kDa (McLaren et al., 2015b). In the study of Jarosch et al. (2015), the broad signal in the above 5 kDa fraction accounted for the majority of P (28 % of total NaOH-EDTA extractable P) and was coincident with P compounds stable to enzymatic hydrolysis, suggesting a high stability of the compounds causing this broad signal towards enzymatic attack. At the same time, McLaren et al. (2015b) reported that the unresolved P_{org} pool represented by an underlying broad signal in the phosphomonoester region dominated the above 10 kDa fraction of five soil extracts and accounted for 61 % to 73 % of total soil P_{org}. The authors suggested that the broad peak originated from P bound via phosphomonoester linkages to the supra-/macromolecular structures of SOM. However, the processes and mechanisms involved in the incorporation of these P compounds into the SOM as well as the chemical structure remained unknown. Without this knowledge,

the understanding of the cycling and transformation of this highly abundant, unresolved P_{org} pool in the soil-plant systems remains limited.

Soil organic matter

The term soil organic matter (SOM) describes the total of organic compounds in soil. Stevenson (1994) excluded undecayed plant and animal tissues, their partial decomposition products as well as the living microbial tissue from the definition of SOM. In contrast, Kögel-Knabner and Rumpel (2018) included plant litter constituents in the definition of SOM, with its compounds ranging from partially undecayed constituents to completely transformed organic substances, microbial cells and microbial products (Kögel-Knabner, 2002, 2017).

SOM either acts as atmospheric C sink or source, and therefore reduces or enhances the greenhouse effect. The effect of SOM on the global C cycle mainly depends on soil properties, soil biota, cover vegetation, climate and other environmental factors as well as soil management and agricultural practices (Stevenson, 1994; Batjes, 1996; Spaccini et al., 2000; Lal, 2004; Kögel-Knabner et al., 2008; Schmidt et al., 2011; Simpson and Simpson, 2012). However, SOM does not only play a crucial role in carbon sequestration but also in nutrient and pollutant cycling (Stevenson, 1982; Stevenson, 1986; McKercher and Anderson, 1989; Pant et al., 1994; Simpson and Johnson, 2006). Vice versa, functional groups containing essential elements, for example the phosphate groups in IP or proteins, are thought to be of great importance in the association processes of SOM with mineral surfaces and hence its stabilisation in the long-term (Kleber et al., 2007; Tipping et al., 2016).

Despite its importance, SOM is described by various models, especially with disagreement about the existence of humic substances and the role of humification processes (Lehmann and Kleber, 2015; Kleber and Lehmann, 2019; Olk et al., 2019; Hayes and Swift, 2020). The more traditional view of SOM is based on solubility characteristics in alkaline solutions and involves the fractions humin, humic acid and fulvic acids (Stevenson, 1994; Olk et al., 2019). According to Stevenson (1994), SOM consisted of known biochemical compounds, such as amino acids, carbohydrates, fats, waxes, resins, organic acids etc. as well as compounds of unknown chemical structure and dissimilar to molecules originating from microorganisms and plants. These compounds were summarised with the term humic

substances which consisted of high molecular weight molecules of complex structure formed by secondary synthesis processes in soil (humification) (Stevenson, 1994).

Later, Piccolo (2001) demonstrated that the apparent macromolecular structure of the humic substances is caused by the association of smaller molecules forming a 'suprastructure'. Based on these results, Nebbioso and Piccolo (2011) introduced the sequential chemical fractionation method 'Humeomics' which fractionates the SOM suprastructure into weak ester-, strong ester- and ether-bound compounds, as well as organomineral compounds. The differentiation between weakly and strongly ester-bound is based on the physical accessibility of these compounds for the extraction solution, i.e. the BF_3 extractant used as a catalyst to break weak ester linkages is considered to be bulkier compared to the smaller alkaline nucleophile KOH used as a catalyst to break strong ester bonds. These findings outdated the theory that humic substances, and hence a large proportion of SOM, are comprised of large macropolymers formed by secondary synthesis processes. Furthermore, the model by Nebbioso and Piccolo (2011) on the different bonding types of organic molecules forming a SOM suprastructure can also be used to investigate the association of SOM with P compounds i.e. the unresolved P_{org} pool, which is hypothesised to be linked to the SOM through ester bonds (McLaren et al., 2015b; McLaren et al., 2020). This shift from individual macropolymers building up the SOM to an association point of view was summarised in the soil continuum model of Lehmann and Kleber (2015). In this model, plant and animal residues degrade from large biopolymers to monomers and finally to CO_2. During this process, desorption/adsorption of organic molecules on mineral surfaces as well as aggregate formation/destruction are potentially happening on each degradation level (Figure 3). The formation of aggregates resp. associations of the organic molecules with soil constituents and/or other organic molecules will influence their solubility and apparent molecular size (Lehmann and Kleber, 2015). These factors are assumed to govern the stabilisation of SOM along with other environmental characteristics (Kleber, 2010; Schmidt et al., 2011).

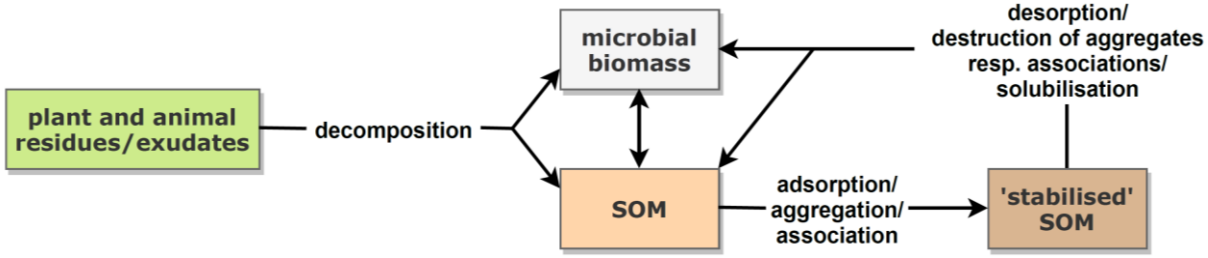


Figure 3. Simplified cycle resp. fate of organic matter in soil (SOM). Arrows depict biotic and abiotic processes. Decomposition and mobilisation processes involve the potential emission of CO_2 as a final product (modified from Lehmann and Kleber (2015) and Kögel-Knabner (2017)).

Structure elucidation of organic phosphorus compounds in soil extracts using ^{31}P NMR spectroscopy

The solution ^{31}P NMR technique is based upon the so called nuclear spin property, which differs for each element and its isotopes (Claridge, 2016a). This property is characterised by the nuclear spin quantum number *I*. Nuclei with *I*=0 such as ^{12}C do not possess a nuclear spin and are therefore not NMR active (Claridge, 2016b). In contrast, ^{31}P has a nuclear spin quantum number of ½ and can hence be observed by NMR. The atoms are brought into a magnetic field (B_0) and subsequently the magnetic moments of the nuclei (M) adjust along the magnetic field vector. Afterwards, a resonant secondary magnetic field is applied by a strong radiofrequency pulse, which deflects the magnetic moments of each nucleus (Figure 4). The nuclei start to precess around B_0 with the same Larmor frequency (ω_0) as the intrinsic magnetic dipoles of each particle. This results in an oscillating magnetic field, which can be detected by a radio frequency coil.

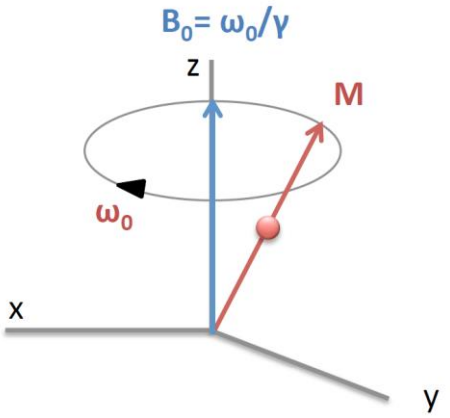


Figure 4. Simplified scheme of NMR experiment. Deflected magnetic moment (M) of a nucleus (red) in laboratory frame x-y-z, Larmor frequency ω_0, initial magnetic field B_0 and gyromagnetic ratio γ. Taken from: Structure determination by NMR, Dr M.-O. Ebert, Prof Dr B. Jaun, ETHZ.

The detected signal as well as the time needed for the magnetic moments to go back into their initial states after the excitation pulse depend on the spin of the nucleus, on neighbouring nuclei/atoms, on the matrix of the measured solution and on molecular motion. The time until the initial state is reached is called relaxation time and it is divided into the recovery in the z-direction (longitudinal) and in the x-y-plane (transversal).

Neighbouring atoms can shield the nucleus of interest from the applied oscillating magnetic field and therefore a difference in the emitted frequency is generated. This difference in frequency results in different chemical shifts (in ppm). In an NMR spectrum, the measured intensity of signals on the y-axis is plotted to the chemical shift on the x-axis. Usually, the chemical shifts of measured molecules are plotted relatively to the chemical shift of a standard compound, which is set to δ 0 ppm.

Since the study of Newman and Tate (1980), solution ^{31}P NMR spectroscopy of alkaline soil extracts became the predominantly applied analytical technique for speciation of P

compounds that can be brought into solution (Cade-Menun and Liu, 2014; McLaren et al., 2020). The NMR soil spectra are usually divided in broad classes of P compounds (Figure 5): phosphonates (δ 19.9 to 13.3 ppm), the combined orthophosphate and phosphomonoester region (δ 6.2 to 2.9 ppm), phosphodiester (δ 2.8 to -2.0 ppm), pyrophosphates (δ -4.8 to -5.5 ppm) and polyphosphates (δ -17.1 to -20.6 ppm). In addition to the qualitative information about the presence of specific P compounds in the extracts, it is possible to acquire quantitative information about the content of these compounds in the initial soil by the addition of a P standard of known concentration, calculation of the net peak area and SDF, if necessary. Furthermore, the recycle delay (time between pulses) of the NMR experiment has to be set long enough for complete relaxation of each nucleus.

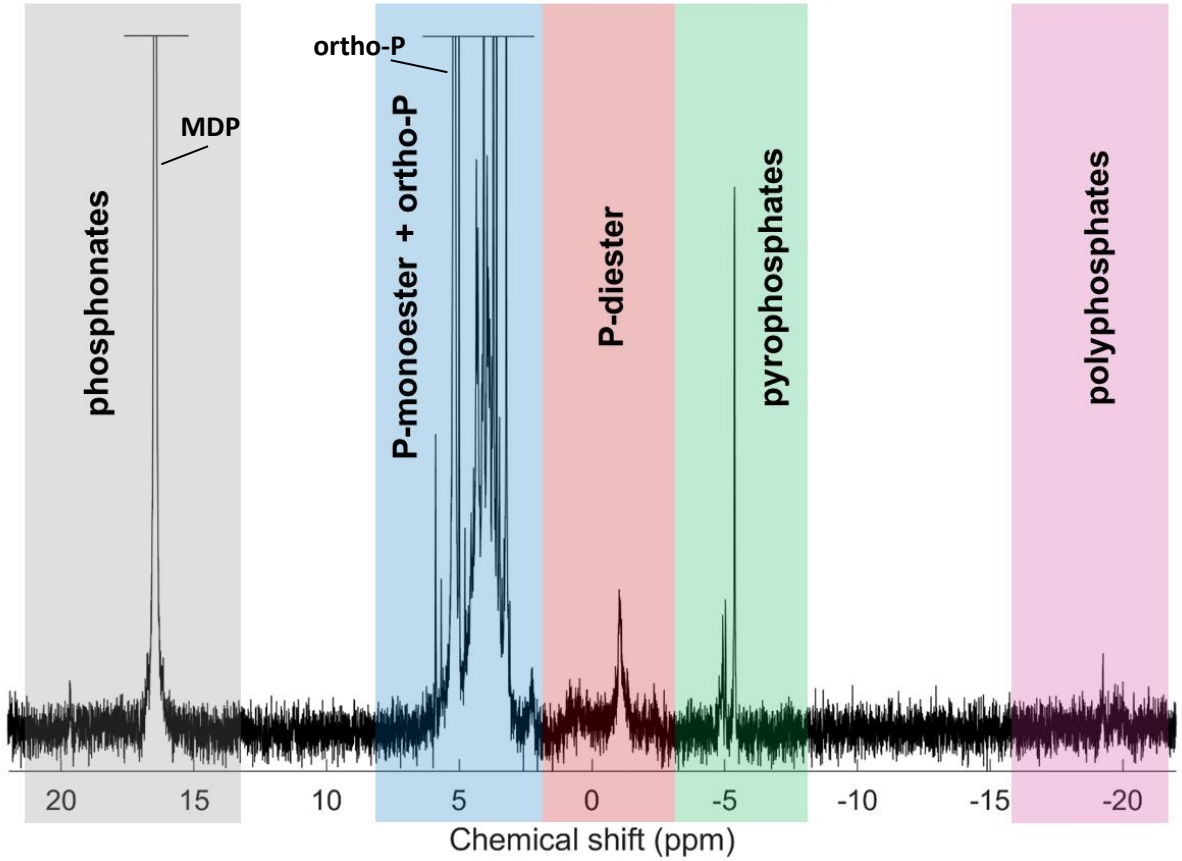


Figure 5. Solution ^{31}P NMR spectrum (500 MHz) of a 0.25 M NaOH + 0.05 M EDTA soil (Gleysol) extract. The broad classes of P compounds are marked, including the phosphonate, combined orthophosphate + phosphomonoester, phosphodiester, pyrophosphate and polyphosphate region. The peaks of orthophosphate as well as the added P-standard methylene diphosphonate (MDP) are indicated.

Specific NMR experiments such as Heteronuclear Single Quantum Correlation (HSQC) can also render information about the molecular structure of specific P compounds by measuring $^{1}H–^{31}P$ correlations (Vestergren et al., 2012). Furthermore, transverse relaxation (T_2) experiments can be used in order to determine the structural composition of detected peaks, i.e. whether a NMR signal is composed of a series of sharp peaks (inhomogeneous line broadening) or contains a single or a few peaks of relatively wide line width (homogeneous line broadening) (Figure 6). Additionally, the T_2 time of a molecule is assumed to be inversely related to its tumbling in the extract and hence its molecular size (Bloembergen et al., 1948; Claridge, 2016b; McLaren et al., 2019).

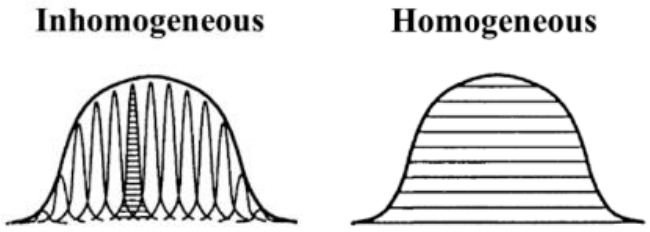


Figure 6. Visual representation of the two types of the composition of a NMR signal. Inhomogeneous line broadening describes a series of sharp peaks with relatively narrow line widths that comprise a broad signal. Homogenous broadening on the right describes a broad NMR signal with a relatively large line width, containing a single (or a few) component(s). The figure was taken and adapted from Schmidt-Rohr and Spiess (1994).

Objective and hypotheses of the PhD project

The objective of this PhD project is to identify the chemical nature of soil P_{org} with increasing molecular weight by combining several chemical extraction approaches, size separation procedures, and solution ^{31}P NMR spectroscopy techniques. Especially the pools of IP and unresolved P_{org} are investigated in depth due to their high abundance in soil systems. This objective will be addressed in four chapters involving the combination of wet chemistry methods with size exclusion chromatography and solution ^{31}P NMR spectroscopy.

In chapter 1, the aim is to assess the SDF approach that renders the most accurate quantification of P_{org} species in alkaline soil extracts using solution ^{31}P NMR spectroscopy. Therefore, the two SDF approaches without and with an underlying broad signal in the phosphomonoester region will be compared based on their quantification accuracy of an added *myo*-IP_6 standard. The main hypothesis is:

- The SDF procedure including an underlying broad signal will provide quantitative analysis of sharp signals (e.g. *myo*-IP_6) in the phosphomonoester region compared to the procedure without a broad signal.

In chapter 2, solution ^{31}P NMR spectroscopy will be carried out on hypobromite oxidised alkaline soil extracts in order to identify IP and investigate their diversity and abundance in soils. This purification and isolation method is assumed to destroy all organic matter except IP (Wrenshall and Dyer, 1941; Turner and Richardson, 2004). Furthermore, the structural composition of compounds in the phosphomonoester region resisting hypobromite oxidation will be probed by carrying out transverse relaxation NMR experiments. The main hypotheses of this chapter are:

- The majority of unidentified sharp peaks in the phosphomonoester region will be resistant to hypobromite oxidation, suggesting the presence of a range of IP, including lower-order IP (IP_1-IP_5).

- The underlying broad signal will be caused by homogeneous broadening, suggesting that it is not comprised of a range of smaller, sharp peaks arising from i.e. IP.

In chapter 3, the aim is to determine the chemical nature of P_{org} in large molecular size fractions. Thus, molecular size fractions of alkaline soil extracts obtained by size-exclusion

chromatography will be analysed by solution ^{31}P NMR spectroscopy. In addition, different metals in these large molecular size fractions will be measured which could be involved in the complexation of P with large molecular size material. We hypothesise that:

- Low molecular size fractions will be dominated by sharp signals mostly arising from IP.

- High molecular weight fractions will be dominated by several broad signals but also contain some sharper signals. The sharp signals could arise from smaller P_{org} molecules, i.e. IP, associated with larger molecules of the SOM.

- The broad signals are caused by a range of P_{org} compounds present in a continuum of molecular sizes rather than by a macropolymeric structure at distinct molecular size.

In chapter 4, the chemical nature of P_{org} associated with the SOM will be investigated by fractionating the complex structure of SOM into smaller components based on their bonding types. For this purpose, the Humeomics sequential chemical fractionation method introduced by Nebbioso and Piccolo (2011) will be carried out on a selected soil sample. Subsequently, the C and P composition of the different extracts and soil residues will be analysed using total element measurements, solid state ^{13}C NMR and solution ^{31}P NMR spectroscopy. The main hypotheses of this chapter are:

- Two IP pools exist in soil: one associated with the SOM and one associated with the mineral phase.

- The broad signals are caused by P_{org} associated with the SOM through ester linkages.

Each chapter will involve detailed analysis on six topsoil samples, which include two Cambisols and a Gleysol from Switzerland, a Cambisol from Germany, a Ferralsol from Colombia and a Vertisol from Australia (WRB, 2014). These soils were chosen from a set of 29 soil samples due to their diverse origin and composition of P_{org}, with concentrations ranging from 70 to 1135 mg P_{org}/kg_{soil}. A detailed list of the origin, properties and chemical composition of the 29 soil samples as well as the information on the mineral composition based on X-ray diffraction (XRD) analyses can be found in Appendices 1 and 2.

Chapter 1

Quantitative measures of *myo*-IP_6 in soil using solution ^{31}P NMR spectroscopy and spectral deconvolution fitting including a broad signal

Published as:

Reusser, J.E., Verel, R., Frossard, E., McLaren, T.I., 2020. Quantitative measures of *myo*-IP_6 in soil using solution ^{31}P NMR spectroscopy and spectral deconvolution fitting including a broad signal. Environmental Science: Processes & Impacts 22, 1084-1094.

https://pubs.rsc.org/en/Content/ArticleLanding/2020/EM/C9EM00485H#!divAbstract

Abstract

Inositol phosphates, particularly *myo*-inositol hexakisphosphates (*myo*-IP_6), are an important pool of soil organic phosphorus (P) in terrestrial ecosystems. To measure concentrations of *myo*-IP_6 in alkaline soil extracts, solution ^{31}P nuclear magnetic resonance (NMR) spectroscopy is commonly used. However, overlap of the NMR peaks of *myo*-IP_6 with several other peaks in the phosphomonoester region requires spectral deconvolution fitting (SDF) to partition the signals and quantify *myo*-IP_6. At present, two main SDF approaches are in use; the first fits a Lorentzian/Gaussian line shape to the *myo*-IP_6 peaks directly to the baseline without an underlying broad signal, and the second fits a Lorentzian/Gaussian line shape to the *myo*-IP_6 peaks simultaneously with an underlying broad peak. The aim of this study was to compare the recovery of added *myo*-IP_6 to soil extracts using both SDF procedures for six soil samples of diverse origin and differing concentrations of organic P (112 to 1505 mg P/kg_{soil}). The average recovery of total added *myo*-IP_6 was 95 % (SD 5) and 122 % (SD 32) using SDF with and without an underlying broad signal, respectively. The recovery of individual peaks of *myo*-IP_6 differed, most notably, the C5 phosphate peak of *myo*-IP_6 was overestimated by up to 213 % when a broad peak was not included in SDF. Based on the SDF procedure that includes a broad peak, concentrations of *myo*-IP_6 ranged from 0.6 to 90.4 mg P/kg_{soil}, which comprised 1 to 23 % of total phosphomonoesters. Our results demonstrate that the SDF procedure with an underlying broad signal is essential for the accurate quantification of *myo*-IP_6 in soil extracts.

Introduction

Phosphorus (P) is an essential macronutrient for all living organisms, which is primarily sourced from the soil environment. It is estimated that between 20 % and 80 % of the total P (P_{tot}) in soil exists in an organic form (Anderson, 1980; Harrison, 1987). A major pool of identifiable organic P (P_{org}) in soil is that of inositol phosphates (IP), of which the *myo* stereoisomer of inositol hexakisphosphate (*myo*-IP_6) is the most abundant (Irving and Cosgrove, 1982; Turner et al., 2012). Studies have reported that pools of *myo*-IP_6 comprise on average one third of the total P_{org} in soil (McLaren et al., 2020). The mechanism for its accumulation in soil is thought to be due to its high binding affinity to aluminium and iron (hydro-)oxides (Ognalaga et al., 1994).

Solution ^{31}P nuclear magnetic resonance (NMR) spectroscopy has been used since 1980 to identify the chemical nature of P_{org} in soil extracts (Newman and Tate, 1980; Cade-Menun and Liu, 2014). The majority of P (~80 %) in NaOH soil extracts is detected in the phosphomonoester region (organic moiety–O–PO_3) of the NMR spectrum (McLaren et al., 2020). However, due to many overlapping signals in this region, spectral deconvolution fitting (SDF) procedures are required to partition the NMR signal (Doolette and Smernik, 2015). The two main SDF approaches applied to soil extracts are that of Turner et al. (2003c) or modifications thereof (Hill and Cade-Menun, 2009), and Bünemann et al. (2008b) or modifications thereof (McLaren et al., 2019).

Turner et al. (2003c) were the first to propose a SDF procedure that could partition the NMR signal within the phosphomonoester region and quantify *myo*-IP_6 in soil extracts. The SDF procedure was carried out using the Bruker WinNMR program, and involved fitting a series of sharp signals from the peak maxima of *myo*-IP_6 to the baseline of the spectra. The procedure was applied to 29 soils under grassland in the United Kingdom and the authors tested the efficacy of the procedure by calculating the recovery of added *myo*-IP_6 in a 1 M NaOH solution containing a mixture of P_{org} compounds. The authors reported that the recovery of added *myo*-IP_6 was on average 102 %. However, a limitation of this study was that recoveries of added *myo*-IP_6 were determined in non-soil extracts.

In contrast, Bünemann et al. (2008b) later proposed a SDF for the quantification of P_{org} compounds (e.g. *myo*-IP_6) that involved fitting a broad feature in the phosphomonoester region, which was then subtracted from the original NMR spectrum and then the overlaying sharp signals were fitted. The authors hypothesised that the broad signal was caused by phosphomonoesters in large and complex molecules, which was later confirmed by

McLaren et al. (2015b) and McLaren et al. (2019). Bünemann et al. (2008b) did not describe the SDF procedure in detail or test its efficacy at the time. However, Doolette et al. (2010) compared the two SDF procedures (with and without a broad signal) to quantitatively recover *myo*-IP_6, which was added to a soil extract. The authors reported that the concentration of *myo*-IP_6 was overestimated by 54 % when a broad signal was not included. Doolette et al. (2011a) later explained the SDF procedure in more detail. A limitation of this study was that the recovery of added *myo*-IP_6 was only tested on one soil extract. In addition, both SDF methods were not able to identify the recovery of the four individual peaks of *myo*-IP_6 due to poor spectral resolution.

A plethora of studies have been carried out since 2003 to identify the chemical nature of soil P_{org} using solution ^{31}P NMR spectroscopy and SDF procedures (Cade-Menun and Liu, 2014). Despite this, there has been no detailed assessment on the efficacy of the two main SDF procedures, which has major consequences for how we understand the composition of soil P_{org}. The implicit assumption of the SDF procedure without a broad signal is that the phosphomonoester region is comprised of an array of sharp peaks (e.g. *myo*-IP_6) with very similar line widths, which are presumably from small organic molecules (Bünemann et al., 2008a; Noack et al., 2012). In contrast, the implication of the SDF procedure with a broad signal is that it interprets the phosphomonoester region as comprised of a broad signal, which is considered to be P_{org} in the form of large molecular structures and associated with the soil organic matter (SOM) (McLaren et al., 2015b; McLaren et al., 2019), and in addition an array of sharp peaks from small organic molecules containing P. It is also important to reconcile these views with previous studies using non-NMR techniques, which often reported large concentrations of P_{org} in large molecular weight fractions (Jarosch et al., 2015; McLaren et al., 2015b).

The aim of this study was to assess the efficacy of the two SDF procedures to quantitatively determine *myo*-IP_6 in soil extracts. *myo*-Inositol hexakisphosphate was chosen because it can be easily detected in ^{31}P NMR spectra on soil extracts as four distinct peaks in a 1:2:2:1 ratio. This ratio is caused by the chemical structure of *myo*-IP_6 (see Figure SI-1 in the supporting information), with the two phosphate groups bound to the C1 and C3 carbon nuclei of the inositol ring being chemically equivalent, and therefore exhibiting the same chemical shift. This is similarly the case for the phosphate groups bound to the C4 and C6 carbon nuclei. The phosphate groups bound to C5 and C2 have distinct chemical environments caused by the conformation of the molecule. The combined four peaks associated with the phosphate groups of *myo*-IP_6 can be used to probe a relatively wide

chemical shift range in the phosphomonoester region of the ^{31}P NMR spectrum using a single compound.

Materials and methods

Soil collection and preparation

Six soil samples were collected from the upper horizon of soil profiles including different soil types and land use systems across four countries. We included soils that covered a wide diversity of organic P contents and soil properties: two Cambisols (Cambisol (P) and Cambisol (A)) and a Gleysol from Switzerland, a Ferralsol from Colombia, a Cambisol (F) from Germany, and a Vertisol from Australia (WRB, 2014). The Ferralsol was collected in 1997 from the 0-20 cm soil layer of the improved grassland treatment of the long-term Culticore field experiment at the Carimagua Research Station in Colombia (Bühler et al., 2003). The Vertisol was collected in 2017 from the 0-15 cm soil layer of a field under cropping from southern Queensland, Australia. The site has been under cultivation for the past 25 years and prior to this was shrubland containing sparse *Eucalyptus camaldulensis* L. and native grasses. The Cambisol (P) was collected in 2017 from the 0-20 cm soil layer of a cultivated field but was under grassland for more than 6 years prior. The Cambisol (F) was collected in 2014 from the 0-7 cm topsoil layer of a beech forest in Bad Brückenau, Germany, as described in Bünemann et al. (2016). The Gleysol was collected in 2017 from the 0-10 cm soil layer of a drained marshland, which has been under grassland for more than 20 years, near Lucerne, Switzerland. The Cambisol (A) sample was collected in 2013 from the 0-20 cm soil layer from an unfertilised border strip of a cultivated field in Rümlang, Switzerland (Meyer et al., 2017).

Background information on the studied sites, and some chemical and physical properties of the soils, are reported in Table 1. The Cambisol (P), Cambisol (F) and Gleysol soil samples were passed through a 5 mm sieve and dried at 60°C for 5 days. The Ferralsol and Cambisol (A) samples were received dried and sieved at <2 mm, whereas the Vertisol sample was received dried (at 40°C for 2 days) and ground at <2 mm. In order to include soils with a diversity of organic P contents and soil properties, soils were sourced from the field in Switzerland and from previous studies (Bühler et al., 2003; Bünemann et al., 2016; Meyer et al., 2017). Consequently, there were small differences in soil preparation among soils used in this study, which may slightly affect soil P extraction. Nevertheless, these differences will not affect the application of SDF procedures to the resultant NMR spectra for the quantification of added *myo*-IP_6 in soil extracts. Concentrations of total carbon (C_{tot}) and nitrogen (N_{tot}) in soil were measured using combustion of 50 mg ground soil weighed into tin foil capsules (vario PYRO cube®, Elementar Analysesysteme GmbH).

Concentrations of 0.5 M H_2SO_4 extractable P_{tot} and P_{org} in soil were measured using the ignition-H_2SO_4 method of Saunders and Williams (1955) as described in Kuo (1996). Total concentrations of soil P were determined by X-ray fluorescence spectroscopy (SPECTRO XEPOS ED-XRF, AMETEK®) using 4.0 g of ground soil sample mixed with 0.9 g of wax (CEREOX Licowax, FLUXANA®). The XRF instrument was calibrated using commercially available reference soils. Soil pH was measured in H_2O at a soil-to-solution ratio of 1:2.5 (w/w) with a glass electrode.

Table 1. Some background information on the sites and some chemical and physical properties of the soils reported in this study.

Parameter	Unit	Ferralsol	Vertisol	Cambisol (P)	Cambisol (F)	Gleysol	Cambisol (A)
Soil type	-	Ferralsol	Vertisol	Cambisol	Dystric Skeletic Cambisol	Gleysol	Calcaric Cambisol
Country	-	Colombia	Australia	Switzerland	Germany	Switzerland	Switzerland
Coordinates of sampling	-	4°30' N / 71°19' W	27°52' S / 151°37' E	46°55' N / 7°36' E	50°21’ N / 9°55’ E	47°05' N / 8°06' E	47°26′ N / 8°31′ E
Elevation	m ASL	150	402	748	800	612	443
Sampling depth	cm	0-20	0-15	0-20	0-7	0-10	0-20
Land use	-	Pasture	Arable	Pasture	Forest	Pasture	Arable
C_{tot}	g C/kg_{soil}	26.7	23.9	21.0	90.3	148.3	27.6
N_{tot}	g N/kg_{soil}	1.7	1.9	2.3	6.6	10.9	2.5
pH in H_2O	-	3.6	6.1	5.1	3.6	5.0	7.7[a]

[a] Meyer et al. (2017)

Extraction of soil organic P

Concentrations of P_{org} were determined based on the method of Cade-Menun et al. (2002). Briefly, 3.0 g of soil was extracted with 30 mL of 0.25 M NaOH + 0.05 M EDTA. Soil extracts were shaken for 16 h on a horizontal shaker at 150 rpm at 24°C, centrifuged for 10 min at 4643 *g*, and then the supernatant passed through a Whatman no. 42 filter paper. A 20 mL aliquot of the filtrate was frozen at -80°C and then lyophilised prior to NMR analysis. This resulted in 420 to 782 mg of lyophilised material across all soils. Concentrations of P_{tot} in the remaining filtrates were measured using inductively coupled plasma-optical emission spectrometry (ICP-OES). Concentrations of molybdate reactive

P (MRP) were measured using the malachite green method of Ohno and Zibilske (1991). The difference between P_{tot} and MRP in NaOH-EDTA filtrates is molybdate unreactive P (MUP), which is largely considered to be P_{org} (Bowman and Moir, 1993) but may also include a small proportion of condensed phosphates (e.g. pyrophosphate) (Vaz et al., 1992).

Preparation of lyophilised material for solution ^{31}P NMR spectroscopy

Preparation of lyophilised material for solution ^{31}P NMR spectroscopy was based on a modification of the methods as reported in Vincent et al. (2013) and Spain et al. (2018). Briefly, 120 mg of lyophilised material was weighed into 1.5 mL microcentrifuge tubes and then a 600 µL aliquot of 0.25 M NaOH + 0.05 M EDTA solution was added. The solution was briefly vortexed and then allowed to rest overnight in order for complete hydrolysis of RNA and phospholipids (Makarov et al., 2002; Turner et al., 2003a; Doolette et al., 2009; Vestergren et al., 2012). This is because the hydrolysis products of RNA (RNA mononucleotides (Makarov et al., 2002)) and phospholipids (α- and β-glycerophosphate (Doolette et al., 2009)) generate peaks in the phosphomonoester region (Turner et al., 2003a; Doolette et al., 2009). The microcentrifuge tubes were then centrifuged at 10621 *g* for 15 min, and a 500 µL aliquot of the supernatant was transferred to another 1.5 mL microcentrifuge tube, which then received a 25 µL aliquot of a 0.03 M methylene-diphosphonic acid (MDP) standard in D_2O (Sigma-Aldrich, product no. M9508) and a 25 µL aliquot of sodium deuteroxide (NaOD) at 40 % (w/w) in D_2O (Sigma-Aldrich, product no. 372072). The solution was briefly vortexed and then transferred to 5 mm NMR tubes for analysis.

Subsequent NMR analyses of the Cambisol (F) and Gleysol samples revealed considerable line broadening of all peaks in the NMR spectra, which might have been caused by high sample viscosity. Therefore, the preparation of lyophilised material was repeated for these samples but at a wider ratio, as recommended by Cade-Menun and Liu (2014). For these samples, 80 mg of lyophilised material was dissolved in 600 µL of 0.25 M NaOH + 0.05 M EDTA solution, and then prepared as previously described. This overcame the issue and resulted in high resolution NMR spectra.

Solution ^{31}P NMR spectroscopy.

All NMR analyses were carried out with a Bruker Avance IIIHD 500 MHz NMR spectrometer equipped with a 5 mm liquid-state Prodigy™ CryoProbe (Bruker Corporation;

Billerica, MA) at the NMR facility of the Laboratory of Inorganic Chemistry (Hönggerberg, ETH Zürich). Solution ^{31}P NMR spectra were acquired using a ^{31}P frequency of 202.5 MHz, with gated broadband proton decoupling and 90° pulses (duration of 12 µs) for excitation. Careful shimming of the samples resulted in a spectral resolution of <0.1 Hz. The recycle delay of each sample was set based on an inversion recovery experiment (Vold et al., 1968). Briefly, the spin-lattice relaxation times (T_1) were calculated from 10 separate experiments with increasing τ values, the time period between the applied pulses in the inversion recovery sequence. Each spectrum was obtained with the collection of 24 scans and a recycle delay of 5 s. Total duration of the inversion recovery experiment for each sample was 56 min. The recycle delay for each sample was calculated by multiplying the longest T_1 value from the inversion recovery experiment by five. This resulted in recycle delays ranging from 6.7 to 31.0 s across all soils. The number of scans was set to 1024 or 4096, depending on the signal-to-noise ratio of the obtained spectrum.

Processing of NMR spectra

Spectral processing involved Fourier transformation, phase correction and baseline adjustment using the TopSpin® software of Bruker (Version 3.5 pl 7, Bruker Corporation; Billerica, MA). All NMR spectra were processed with an exponential line broadening of 0.6 Hz. Since the concentration of added MDP is known, its integral (net peak area) is directly proportional to that of all other NMR signals (Doolette et al., 2011a). Therefore, quantification of P species in NMR spectra was carried out based on spectral integration (Turner, 2008). In general, integral regions included: 1) phosphonates, in particular the added MDP (δ 16.9 to 16.3 ppm), which includes its two carbon satellite peaks (at δ 16.96 and 16.36 ppm), 2-aminoethylphosphonic acid (δ 19.8 to 19.6 ppm), unknown phosphonate 1 (δ 19.3 to 19.2 ppm), unknown phosphonate 2 (δ 18.3 to 18.1 ppm) and unknown phosphonate 3 (δ 16.5 to 16.4 ppm); 2) the combined orthophosphate and phosphomonoester region (δ 6.0 to 3.0 ppm); 3) phosphodiesters, in particular unknown phosphodiester 1 (δ 2.5 to 2.2 ppm), unknown phosphodiester 2 (δ 0.6 to 0.5 ppm), DNA (δ -0.7 to -1.4 ppm) and unknown phosphodiester 3 (δ -2.3 to -2.4 ppm); and 4) pyrophosphate (δ -4.8 to -5.4 ppm). These integral regions are highlighted in Figure SI-3. Due to overlapping signals in the orthophosphate and phosphomonoester region, SDF was needed to partition the NMR signals within this region.

Deconvolution fitting procedures

Two spectral deconvolution fitting approaches were applied to the orthophosphate and phosphomonoester region. The first approach involved fitting all identifiable sharp peaks (i.e. distinguishable from the noise of the spectrum) to the baseline of the spectra, which is based on the method of Turner et al. (2003b). The second approach involved fitting all identifiable sharp signals and simultaneously an underlying broad signal in the phosphomonoester region, which has been described in McLaren et al. (2019), and is based on the method of Bünemann et al. (2008a). We used MATLAB® R2017a (The MathWorks, Inc.) scripts containing a non-linear optimisation algorithm for the SDF of the NMR spectra. Figure 1 shows a graphical representation of the two deconvolution fitting procedures used in this study (peak assignment according to (Doolette et al., 2011a)). For both SDF procedures, a Lorentzian or Gaussian line shape model was selected for each peak based on a visual assessment of the most optimised fit and calculated residuals. In general, a Lorentzian line shape was used to fit the broad signal, signals of *myo*-IP_6 and orthophosphate, whereas a Gaussian line shape was used to fit the remaining sharp signals, which were generally of lower signal intensity. Each of the identifiable sharp peaks were fitted between an upper and lower bound on both their line widths at half-height and their peak positions at highest intensity. The same was true for the peak position of the underlying broad signal and its line width at half-height. The upper and lower bound for both parameters of the broad signal being set to δ 3.8 ppm and δ 4.5 ppm, and from 19 Hz to 293 Hz, respectively. Within these visually assessed boundaries, all signals (including the underlying broad signal) were fitted by the non-linear optimisation algorithm. The residues of all fitted spectra are plotted in Figures SI-4 and SI-5. The goodness of fit parameters of the SDF including the Root-mean-square fitting error and R^2 (coefficient of determination) are listed in Table SI-1 (supporting information). Table SI-1 also includes the results of a reduced Chi^2-test in MATLAB® with the fitted peaks and the according residuals (chemical shift range δ 5.2 to 3.0 ppm). We carried out the Chi^2-test in order to determine if an overfitting of the spectra occurred ($Chi^2<1$). Briefly, the results ($Chi^2>1$) show that there was no overfitting of the spectra using both SDF approaches, and that the Chi^2 values of the SDF approach with a broad signal were closer to 1 than without a broad signal, which suggests the former approach is a better fit of the spectra. Furthermore, we carried out the Bootstrap sampling function of MATLAB® (n=100) for all peaks within the given boundaries of the peak positions, the area and line widths at half-height. The resulting means of the peak position and line widths at half-height including the standard deviations for the broad signal are listed in Table SI-2. The calculated standard deviations

were very low for all three parameters, indicating a good fit of the broad peak within the given boundaries.

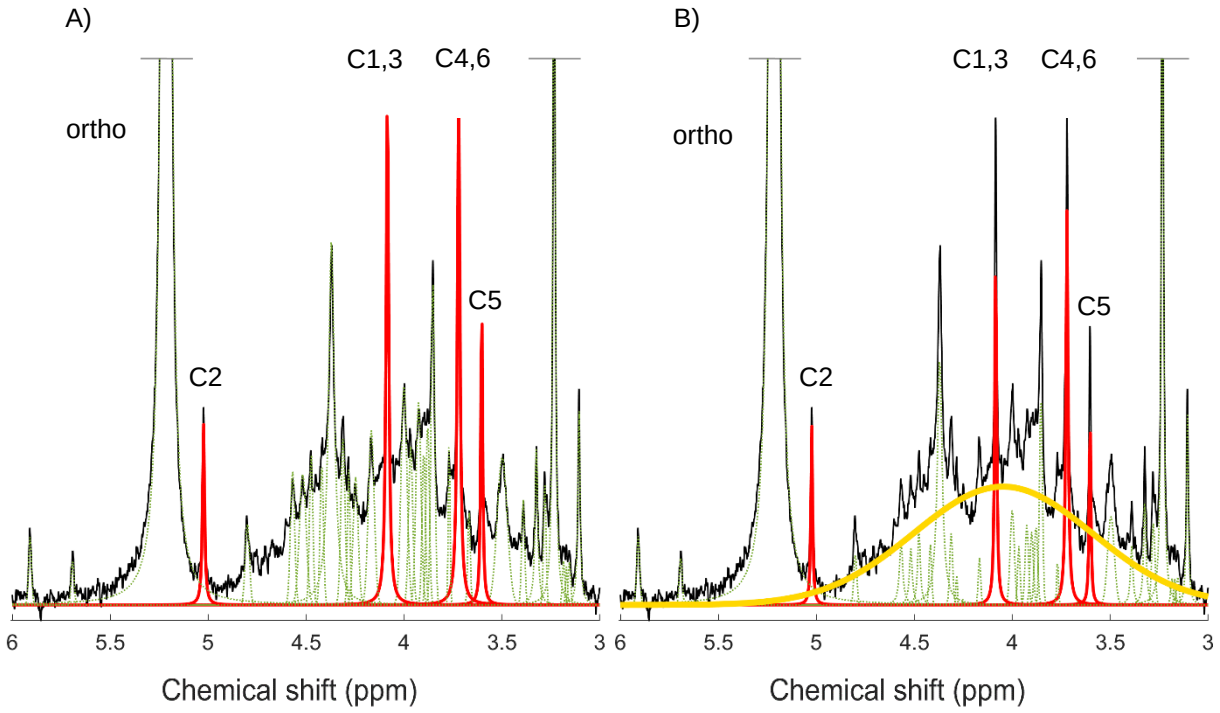


Figure 1. Solution ^{31}P NMR spectrum of 0.25 M NaOH + 0.05 M EDTA Cambisol (A) extract (black). Graphical representation of the two spectral deconvolution fitting approaches without (A) and with an underlying broad signal (B). All fitted peaks are illustrated; the broad peak (yellow) and the four *myo*-IP_6 peaks (red) have been highlighted. In addition, the phosphate groups bound to carbons C1-6 of *myo*-IP_6 and also that of the orthophosphate (ortho) peak, have been identified.

Spiking experiment

The two deconvolution fitting procedures were assessed by determining the recovery of a known amount of added *myo*-IP_6 to soil extracts. After NMR analysis of the unspiked soil extract, a 10 µL aliquot of a 5.5 mM *myo*-IP_6 standard in D_2O was added to all soil extracts except for the Gleysol sample (Sigma-Aldrich, product no. P5681). For the Gleysol, a 10 µL aliquot of a 11 mM *myo*-IP_6 standard in D_2O was added. The aim was to add *myo*-IP_6 at a concentration that would result in an increase of peak intensity of approximately 3-times the peak intensity in unspiked extracts (Doolette et al., 2009). The NMR tube was then sealed with parafilm, inverted several times, and then allowed to rest prior to NMR analysis. The NMR analysis parameters on the spiked soil extract were the same as that carried out on unspiked extracts. Similarly, spectral processing and quantification were carried out as previously described. The recovery of added *myo*-IP_6 was calculated using Equation 1.

$$Recovery\ of\ added\ myo-IP_6\ (\%) = \frac{A\ (mg\ P\ l^{-1}) - B\ (mg\ P\ l^{-1})}{C\ (mg\ P\ l^{-1})} * 100\ , \quad (1)$$

where *A* refers to the concentration of *myo*-IP_6 in the spiked soil extract, *B* to the concentration of *myo*-IP_6 in the unspiked extract and *C* to the concentration of the added *myo*-IP_6. Solution ^{31}P NMR recovery of the phytate standard revealed impurities, therefore *C* represents the actual concentration of *myo*-IP_6 in the standard (see Figure SI-2).

Statistical analyses and graphics

All graphics were created using MATLAB® R2017a (The MathWorks, Inc.). All statistical analyses were carried out using Microsoft® Excel 2016. This included calculating the mean values and standard deviations (SD) of the added *myo*-IP_6 recovery across the six soil samples. The NMR observability was calculated by comparing the concentration of P_{tot} as detected by NMR with that measured by ICP-OES (Dougherty et al., 2005; Doolette et al., 2011a), which ranged from 52 % to 89 % (on average of 68 %) across all soils.

Results

Pools of soil P

Concentrations of P_{tot} in soil using XRF ranged from 320 to 3841 mg P/kg_{soil} (Table 2). Concentrations of P_{tot} in NaOH-EDTA extracts comprised 28 % to 51 % of the P_{tot} in soil as measured by XRF. Concentrations of P_{org} using the ignition-H_2SO_4 extraction technique ranged from 143 to 1377 mg P/kg_{soil}. The concentration of P_{org} in NaOH-EDTA extracts ranged from 93 to 1326 mg P/kg_{soil}, which comprised 8 % to 35 % (an average value of 24 %) of the total soil P using XRF.

Table 2. Total soil phosphorus (P_{tot}) as measured by X-ray fluorescence (XRF) spectroscopy, and pools of extractable P using the ignition-H_2SO_4 extraction technique of Saunders and Williams (1955) and the NaOH-EDTA extraction technique of Cade-Menun et al. (2002). The percentage of extractable P to that of P_{tot} in soil as measured by XRF is shown in parentheses.

Soil	XRF	Ignition-H_2SO_4	NaOH-EDTA		
	P_{tot}	**P_{org}**	**P_{tot}**	**MRP[a]**	**MUP[b]**
	mg P/kg_{soil}	mg P/kg_{soil} (%)	mg P/kg_{soil} (%)	mg P/kg_{soil} (%)	mg P/kg_{soil} (%)
Ferralsol	320	109 (34)	160 (50)	67 (21)	93 (29)
Vertisol	1726	143 (8)	484 (28)	351 (20)	133 (8)
Cambisol (P)	2553	729 (29)	863 (34)	323 (13)	540 (21)
Cambisol (F)	3841	1377 (36)	1850 (48)	525 (14)	1326 (35)
Gleysol	2913	939 (32)	1490 (51)	610 (21)	880 (30)
Cambisol (A)	1724	430 (25)	510 (30)	128 (7)	382 (22)

[a] Molybdate reactive P (MRP) based on the malachite green method of Ohno and Zibilske (1991).

[b] The difference between P_{tot} and MRP is molybdate unreactive P (MUP), which is considered to be P_{org}.

Solution ^{31}P NMR spectra of soil extracts

The majority of NMR signals occurred in the orthophosphate and phosphomonoester region (δ 6.0 to 3.0 ppm), which comprised on average 96 % of total NMR signal (Table 3). In general, the peak of greatest intensity across all soils was that of orthophosphate (apart from the added MDP). The largest pool of P_{org} as determined by the integral over the various regions of peaks was that of phosphomonoesters, which accounts on average for 94 % of the total NMR signal arising from organic forms. Concentrations of

phosphomonoesters range from 36.3 to 501.1 mg P/kg_{soil}. The remaining NMR signal is distributed between phosphodiesters, pyrophosphates and phosphonates.

Table 3. Relative concentrations of P classes (phosphonates, orthophosphates, phosphomonoesters, phosphodiesters and pyrophosphates) determined by solution ^{31}P NMR spectroscopy as percentage (%) of total P in NaOH-EDTA soil extracts. Chemical shift regions were attributed to P species according to peak positions in spectra.

Soil	Phosphonate[a] (δ 19.8 to 16.5 ppm)	Ortho-P (δ 5.5 to 5.0 ppm)	P-monoester (δ 6.0 to 3.0 ppm)	P-diester (δ 2.5 to -2.4 ppm)	Pyrophosphate (δ -4.8 to -5.3 ppm)
Ferralsol	1.0	55.0	37.0	5.1	1.9
Vertisol	1.0	85.0	13.3	0.0	0.7
Cambisol (P)	0.2	50.4	47.8	0.4	1.2
Cambisol (F)	1.5	46.7	47.7	2.9	1.3
Gleysol	0.0	48.5	45.3	3.3	2.9
Cambisol (A)	0.2	51.1	46.9	0.0	1.8

[a] The added methylenediphosphonic acid (MDP) standard is not included.

The phosphomonoester region comprised of two main spectral features based on a visual assessment; the presence of (i) 18 to 47 sharp signals, and (ii) an underlying broad signal. The four peaks of *myo*-IP_6 were observed in the NMR spectra of all soils except in the Vertisol, in which only the C1,3 and C4,6 peaks could be clearly identified due to the low initial concentration of *myo*-IP_6. In all soils, the C2 peak of *myo*-IP_6 at δ 5.04 ppm in the NMR spectra exhibited little overlap with the broad signal compared to the other peaks of *myo*-IP_6. Some slight overlap between the C2 peak of *myo*-IP_6 and the base of the orthophosphate peak was observed in the Cambisol (F) and Gleysol.

Spike recoveries of total myo-IP_6

Spiking the soil extracts with *myo*-IP_6 resulted in a clear increase in its peak intensities at δ 5.04 ppm, 4.10 ppm, 3.72 ppm and 3.61 ppm across all soils (Figure 2). On average, the increase in peak intensity of *myo*-IP_6 was 3-fold relative to the unspiked samples in terms of absolute intensity from the peak maximum to the baseline of the spectrum. The exception was the Vertisol sample, where the increase of the C1,3 and the C4,6 peak was 7-fold relative to the unspiked sample. The increase in peak intensity due to spiking varied for the four individual peaks of *myo*-IP_6. The C2 peak showed the greatest increase of

intensity (3.2-fold) whereas the C5 peak showed the least (2.6-fold). The addition of the phytate standard to the soil extract also resulted in an increase in the intensity of other peaks to that of *myo*-IP_6, particularly peaks at δ 3.98, 4.14, 4.17 and 4.57 ppm.

The recovery of total *myo*-IP_6 in the six soil extracts using the SDF procedure with a broad signal was on average 95 %, whereas this was on average 122 % using the SDF procedure without a broad signal (Table 4). In addition, the variation in recovery of *myo*-IP_6 across the six samples was least using the former approach (SD of 5) compared to the latter approach (SD of 32).

Table 4. Calculated recoveries of the added *myo*-IP_6 (Sigma-Aldrich, product no. P5681) in 0.25 M NaOH + 0.05 M EDTA soil extracts across 6 soil samples (Ferralsol, Vertisol, Cambisol (P), Cambisol (F), Gleysol and Cambisol (A)), and their standard deviation (SD). Concentrations of the added *myo*-IP_6 were obtained from solution ^{31}P NMR spectra using two spectral deconvolution fitting (SDF) procedures; one (i) with a broad underlying signal based on the method of Bünemann et al. (2008b), and (ii) one without an underlying broad signal based on the method of Turner et al. (2003c). The carbon nuclei C1-C6 of the inositol ring on which the phosphate group is attached has been indicated.

SDF procedure	*myo*-IP_6	Fer. (%)	Ver. (%)	Ca. (P) (%)	Ca. (F) (%)	Gley. (%)	Ca. (A) (%)	Average (%)	SD
with broad signal	total *myo*-IP_6	96	90	99	92	91	105	95	5
	C2-*myo*-IP_6	103	101	73	106	95	104	97	12
	C1,C3-*myo*-IP_6	97	91	108	74	81	106	93	14
	C4,C6-*myo*-IP_6	103	91	103	96	97	116	101	9
	C5-*myo*-IP_6	73	77	96	106	95	81	88	13
without broad signal	total *myo*-IP_6	146	120	73	151	94	147	122	32
	C2-*myo*-IP_6	146	113	76	111	101	119	111	23
	C1,C3-*myo*-IP_6	148	109	77	130	83	175	120	38
	C4,C6-*myo*-IP_6	150	127	67	158	91	124	120	35
	C5-*myo*-IP_6	135	133	76	213	117	167	140	46

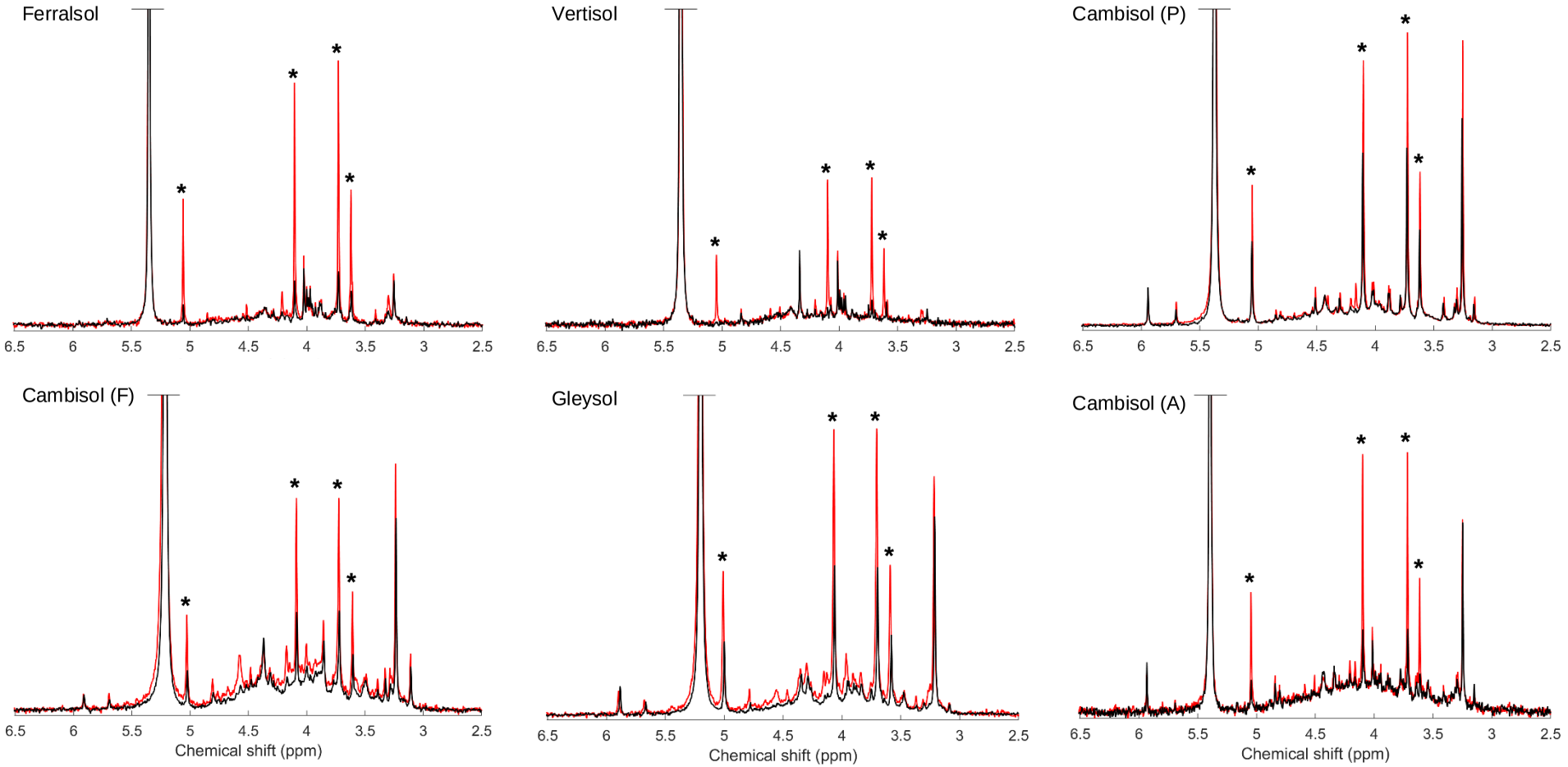


Figure 2. Solution ^{31}P nuclear magnetic resonance (NMR) spectra of the orthophosphate and phosphomonoester region on unspiked (black) and spiked (red) 0.25 M NaOH + 0.05 M EDTA soil extracts. Samples were spiked with a 10 µL aliquot of 5.5 mM *myo*-IP_6 (Sigma-Aldrich, product no. P5681) except for the Gleysol sample, which was spiked with a 11 mM *myo*-IP_6-standard solution. Signal intensities were normalised to the MDP peak intensity for direct comparison of unspiked and spiked extracts. However, in order to highlight spectral features, the vertical axes of the spectra have been increased by a factor of 1.7 for the Cambisol (P), the Cambsol (F) and the Gleysol, and by a factor of 10 for the Ferralsol, the Vertisol and the Cambisol (A). The four peaks generated by *myo*-IP_6 are marked with an asterisk.

Spike recoveries of individual peaks of myo-IP$_6$

Spike recoveries for each of the four peaks of *myo*-IP$_6$ differed between the SDF approaches across all soils (Table 4). Overestimation of spike recoveries using the SDF procedure without a broad peak was greatest (up to 213 %) for the C5 peak of *myo*-IP$_6$ compared to all other peaks. Furthermore, spike recoveries of the C1,3 and C4,6 peaks of *myo*-IP$_6$ were overestimated on average 8 % more than that of the C2 peak. The peak ratios of *myo*-IP$_6$ in the unspiked soils were on average 1.0:2.2:1.8:0.9 with a broad signal and 1.0:2.1:1.9:0.7 without a broad signal, when the C2 peak was set to 1 (the Vertisol sample was not included due to unreliable measures of the C2 and C5 peaks of *myo*-IP$_6$).

Quantification of myo-IP$_6$ (and the broad signal) in soil extracts

Concentrations of total *myo*-IP$_6$ in soil extracts obtained with SDF with a broad peak ranged from 0.6 to 90.4 mg P/kg$_{soil}$, which comprised between 1 % and 23 % of total phosphomonoesters (Table 5). On average, the broad signal accounted for 64 % of total phosphomonoesters across all soils (Table 5). When concentrations of *myo*-IP$_6$ as determined by SDF with a broad peak were subtracted from the *myo*-IP$_6$ values obtained by SDF without a broad peak, the amount of P$_{org}$ that would have been previously attributed to *myo*-IP$_6$ ranged from 1.3 to 56.1 mg P/kg$_{soil}$.

Table 5. Concentrations of organic P compounds obtained from solution ^{31}P NMR spectra of 0.25 M NaOH + 0.05 M EDTA soil extracts. The spectral deconvolution fitting of the phosphomonoester region has been carried out with the inclusion of an underlying broad signal.

Parameter	Unit	Ferralsol	Vertisol	Cambisol (P)	Cambisol (F)	Gleysol	Cambisol (A)
myo-IP$_6$	mg P /kg$_{soil}$	4.4	0.6	80.7	46.2	90.4	6.3
Proportion of *myo*-IP$_6$ to total phosphomonoesters	%	12	1	23	9	23	4
Broad peak	mg P /kg$_{soil}$	21.6	30.9	186.0	305.8	216.7	110.3
Proportion of broad peak to total phosphomonoesters	%	59	79	53	61	54	78

Discussion

Extractability of soil organic P and solution ^{31}P NMR spectra

Concentrations of P_{org} in NaOH-EDTA extracts were similar to reported values in previous studies (Turner, 2008; Jarosch et al., 2015). The difference between total soil P as measured by XRF and NaOH-EDTA extractable P is likely due to P_{inorg} held within mineral silicates and other insoluble mineral phases containing P (McLaren et al., 2015a). Organic P extracted with the NaOH-EDTA technique is strongly correlated to pools of soil P_{org} (McLaren et al., 2015a). The extraction method allows for detailed characterisation of P_{org} forms using ^{31}P NMR spectroscopy (Cade-Menun and Liu, 2014). Solution ^{31}P NMR spectra were highly resolved and exhibited a high signal-to-noise ratio across all samples. The broad classes of P detected by NMR across all soils were phosphonates, orthophosphate, phosphomonoesters, phosphodiesters and pyrophosphate, except for the absence of phosphonates in the Gleysol and phosphodiesters in the Vertisol. Their distribution within the total NMR signal is consistent with previous studies, which typically show the majority of NMR signal occurs in the orthophosphate and phosphomonoester region (Doolette et al., 2011b; Jarosch et al., 2015; McLaren et al., 2015b). The peaks of *myo*-IP_6 were also clearly observed within this region, and therefore allowed for their quantification using SDF (Doolette and Smernik, 2015).

Recovery of myo-IP_6 in soil extracts

Concentrations of added *myo*-IP_6 were overestimated when the SDF procedure did not include a broad signal. The reason for this is likely twofold: 1) the intensities of the *myo*-IP_6 peaks are higher when a broad signal is not included; and 2) the line widths of these peaks when fitted to the baseline are greater if a broad signal is not included in the SDF compared to that when a broad signal is included. Therefore, the peak area belonging to compounds other than *myo*-IP_6 is being attributed to *myo*-IP_6, which may result in an overestimation of its concentration in soil. This supports the finding of Doolette et al. (2010) who reported an overestimation of a phytate spike by 54 % when a broad signal was not fitted. Whilst the finding of Doolette et al. (2010) was higher than that found in the current study, it is likely due to a greater overlap among sharp peaks in the phosphomonoester region. The authors observed and fitted up to six sharp peaks in the orthophosphate and phosphomonoester region, but the C2 peak of *myo*-IP_6 was not visible because it overlapped with orthophosphate. This was not the case in the current study where all four

peaks of *myo*-IP_6 were visible and the C2 peak of *myo*-IP_6 was clearly separated from orthophosphate.

Whilst the recovery of added *myo*-IP_6 was generally overestimated using SDF without a broad signal across all soils, this was not the case for the Cambisol (P) soil sample. The reason for this is unclear but might be due to the high proportion of *myo*-IP_6 to total phosphomonoesters in this soil. It might also be due to an underestimation of the side regions at the base of the peak, as the fitting of neighbouring sharp peaks can influence the partitioning of signals within a particular region. In comparison to SDF with a broad signal of this soil, the fitting of side peak areas of *myo*-IP_6 appeared to improve based on a visual assessment, i.e. the peak height was reduced but the line width at half-height increased. This highlights the sensitivity of peak fitting to small changes in peak shape for the quantification of P species in some soils.

The variability between the two SDF approaches differed in their recovery of added *myo*-IP_6 across all soils, which was least for the SDF approach with a broad signal than that without. The reason for this appears to be the relative proportion of the underlying broad signal to the *myo*-IP_6 peaks among different soils. This is consistent with other compounds that are present within this region, which overlay the broad signal. Doolette et al. (2010) reported that concentrations of glycerophosphate were doubled when SDF was carried out without an underlying broad signal. Of course, since *myo*-IP_6 exhibits 4 peaks in a NMR spectrum, this would vary between the individual peaks of *myo*-IP_6, based on differing proportions of an underlying broad signal.

Individual recoveries of the four peaks of *myo*-IP_6 differed. The greatest difference in the recovery of peaks from *myo*-IP_6 occurred when SDF was carried out without a broad signal, which supports that the baseline of the spectra varies within these chemical shifts (Doolette and Smernik, 2015). The C1,3 and C4,6 peaks of *myo*-IP_6 have chemical shifts within the phosphomonoester region where the intensity of the broad peak is high relative to that for the C2 and C5 peaks. However, the recovery of the C5 peak of *myo*-IP_6 at δ 3.61 ppm was generally overestimated more than that of the other *myo*-IP_6 peaks. The C5 peak of *myo*-IP_6 is present on the shoulder of the broad signal, which has a maximum intensity at about δ 4.06 ppm. Doolette and Smernik (2015) hypothesised that fitting peaks from the peak maxima to the baseline would result in substantially more signal from an underlying broad peak to the C5 peak of *myo*-IP_6 compared to the C2 peak of *myo*-IP_6. Moreover, the C5 peak appears to be most sensitive to changes in the fitting procedures, as it shows the highest variation for both deconvolution approaches. Another source of

variation in the quantification of the C5 peak is that it overlaps with the base of the upfield C4,6 peak of *myo*-IP_6, and possibly peaks from uridine-2'-monophosphate, adenosine-2'-diphosphate, and some unidentified compounds (Vestergren et al., 2012). There is also evidence that the broad peak itself is comprised of more than one component (Doolette and Smernik, 2015; McLaren et al., 2019), which may result in an imperfect Lorentzian/Gaussian distribution within the phosphomonoester region. Nevertheless, our study shows that the quantification of *myo*-IP_6 using SDF with a broad signal generally results in measures of *myo*-IP_6 that are more accurate and consistent across a diversity of soils compared to that of SDF without a broad signal.

The comparison of the peak ratios to the theoretical value of 1:2:2:1 was used in previous studies to evaluate the accuracy of the applied SDF procedure (Vincent et al., 2012). In our study, the peak ratios of *myo*-IP_6 were similar between the two SDF approaches and close to the theoretical ratio of 1:2:2:1 (Costello et al., 1976). Therefore, the ratio of peaks from *myo*-IP_6 cannot be used to assess the efficacy of the SDF for accurate quantification of organic P compounds. Since the peak ratios of *myo*-IP_6 were not a useful assessment of the SDF approaches, it is likely that other organic P compounds exhibiting sharp signals may provide some insight into the validity of the SDF approach. In particular, an overestimation would be likely for the C2,5 peak of *neo*-IP_6 in the 4-eq-2-ax conformation (Turner et al., 2012), since its chemical shift is present in the region of the broad peak. Whereas the C1,3,4,6 peak of *neo*-IP_6 is located upfield of the orthophosphate peak and therefore does not overlap with the broad signal. We identified both of these peaks (δ 5.92 and 3.78 ppm) through spiking in the Cambisol (P), the Cambisol (F), the Gleysol and the Cambisol (A) soil samples (data not shown), calculated their ratios, and compared to the theoretical ratio of 4:2 (Turner et al., 2012). Calculated ratios were on average 4.0:6.2 when carrying out SDF without a broad signal and 4.0:1.5 with an underlying broad signal. These results provide supporting evidence that a broad signal should be included when carrying out SDF.

In the current study, there was a large number of sharp peaks detected in the phosphomonoester region, which was likely due to optimised extraction techniques and high resolution NMR (Cade-Menun and Liu, 2014). Interestingly, the intensity of several unidentified sharp peaks increased when the *myo*-IP_6 standard was added (Figure 2). The *myo*-IP_6 standard contained many phosphomonoester impurities (see Figure SI-2), which are likely those of lower-order *myo*-IP (Doolette and Smernik, 2018). Doolette and Smernik (2018) investigated the chemical composition of a variety of purchased or synthesised

phytate standards using solution ^{31}P NMR spectroscopy. The authors found that the majority of phytate standards were impure and contained a mixture of lower- (and higher-) order IP, and orthophosphate. The authors suggested that thermal degradation during storage was the primary mechanisms in the degradation of higher-order IP to lower-order IP. This suggests the presence of lower-order IP in soil extracts that could be detected using solution ^{31}P NMR spectroscopy.

Concentrations of myo-IP$_6$ (and the broad signal) in soil

Our results demonstrate that a quantitative determination of *myo*-IP$_6$ in soil extracts using solution ^{31}P NMR spectroscopy requires SDF procedures that include an underlying broad signal. In the current study, concentrations of *myo*-IP$_6$ and their proportion to the total pool of P_{org} in soil were generally in range or lower than that more broadly reported in the literature. This most likely reflects the majority of published studies that primarily carry out SDF without a broad signal. Clearly, pools of *myo*-IP$_6$ in soil are an important portion of the soil P_{org}, which are found in the majority of soils across the world (Turner et al., 2002). However, they do not account for the majority of P_{org} in soil and sometimes found at negligible concentrations in some soils (Williams and Anderson, 1968; Steward and Tate, 1971; Turner, 2006).

The largest pool of soil P_{org} was that of the broad signal across all soils. This is consistent with previous studies, where it generally comprises 40 % to 70 % of the total P_{org} in soil (Doolette et al., 2011a; Jarosch et al., 2015; McLaren et al., 2015b). The exact chemical nature of this pool remains unclear but can be described as phosphomonoesters in the form of large molecular structures (McLaren et al., 2015b), which contain pools of P_{org} resistant to enzymatic hydrolysis (Jarosch et al., 2015), associated with humic fractions (He et al., 2011), and are structurally complex (McLaren et al., 2019). The presence of a broad peak is consistent with previous studies using non-NMR techniques that report a large proportion of the P_{org} in soil is unresolved and can occur in large molecular weight fractions (Steward and Tate, 1971; Dalal, 1977; Harrison, 1987; Bowman and Moir, 1993).

Conclusion

myo-Inositol hexakisphosphate is an important pool of soil organic P. However, its accurate quantification using solution ^{31}P NMR spectroscopy followed by SDF is uncertain. Our aim was to compare the recovery of added *myo*-IP_6 using two SDF procedures in NMR spectra on soil extracts. The average recovery of total added *myo*-IP_6 by SDF with a broad signal was close to 100 % and exhibited less variation than that by SDF without a broad signal. The recovery of individual peaks of *myo*-IP_6 differed between its four peaks, which was overestimated for the C5 phosphate peak by up to 140 % when a broad peak was not fitted. We recommend that the accurate quantification of *myo*-IP_6 using solution ^{31}P NMR spectroscopy on soil extracts includes a broad signal when carrying out SDF. This is also relevant for other sharp signals in the phosphomonoester region, which overlay the broad signal. Furthermore, our results show that previous studies reporting concentrations of *myo*-IP_6 using SDF without an underlying broad signal may be unreliable. It is essential that pools of *myo*-IP_6 (and the broad signal) are accurately determined for an improved understanding of the abundance and cycling of organic P in soil.

Supporting information

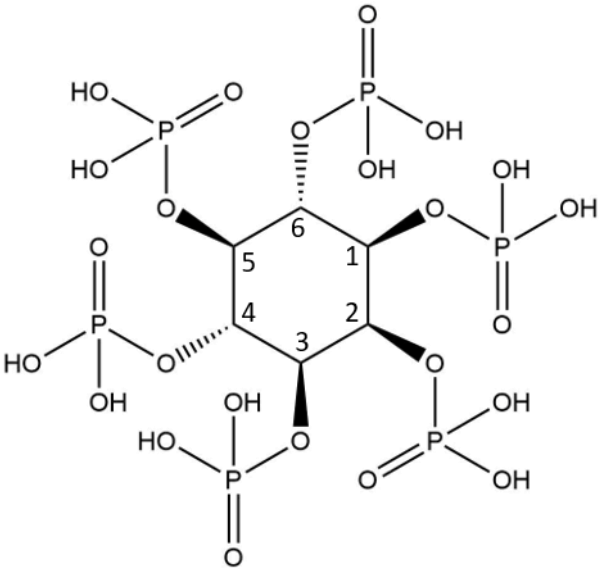


Figure SI-1. Structural formula of *myo*-inositol hexakisphosphate (*myo*-IP_6). The carbon positions of the inositol ring have been identified (1-6). ChemDraw Professional 17.0, ©PerkinElmer Informatics Inc.

Analysis of myo-IP_6 standard

Doolette and Smernik (2018) reported the P impurities of commercially available 'phytate' standards. The authors reported that concentrations of *myo*-IP_6 in the standards ranged from 8 % to 93 % of the P_{tot}. Consequently, the phytate standard (Sigma-Aldrich, product no. P5681) used in this study was analysed using solution ^{31}P NMR spectroscopy prior to the spiking experiments. The spectrum (Figure SI-2) revealed P impurities similar to that reported by Doolette and Smernik (2018), which comprised 28.97 % of P_{tot} in the purchased standard.

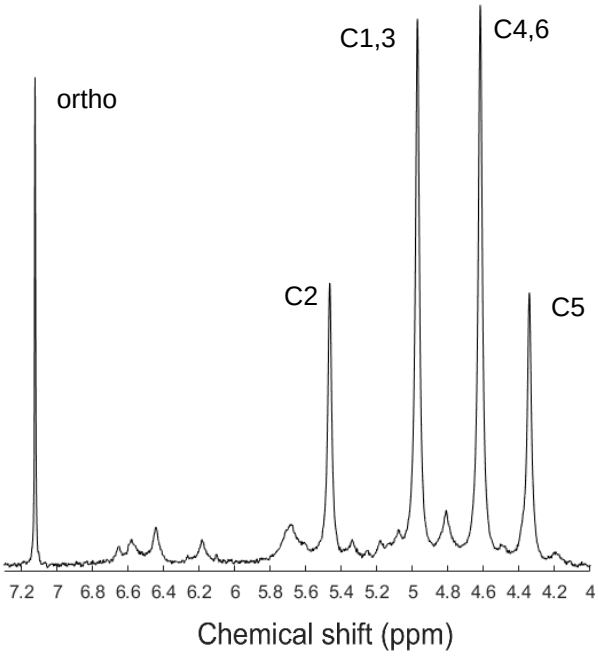


Figure SI-2. Solution ^{31}P NMR spectrum (500 MHz) of *myo*-IP_6 standard (Sigma Aldrich, product no. P5681), 0.2 mg P in 150 µL D_2O and 400 µL NaOD. Identified are P peaks of *myo*-IP_6 labeled with corresponding carbons C1-6 and orthophosphate (ortho). All other peaks are assumed to arise from impurities, possibly representing lower order IP.

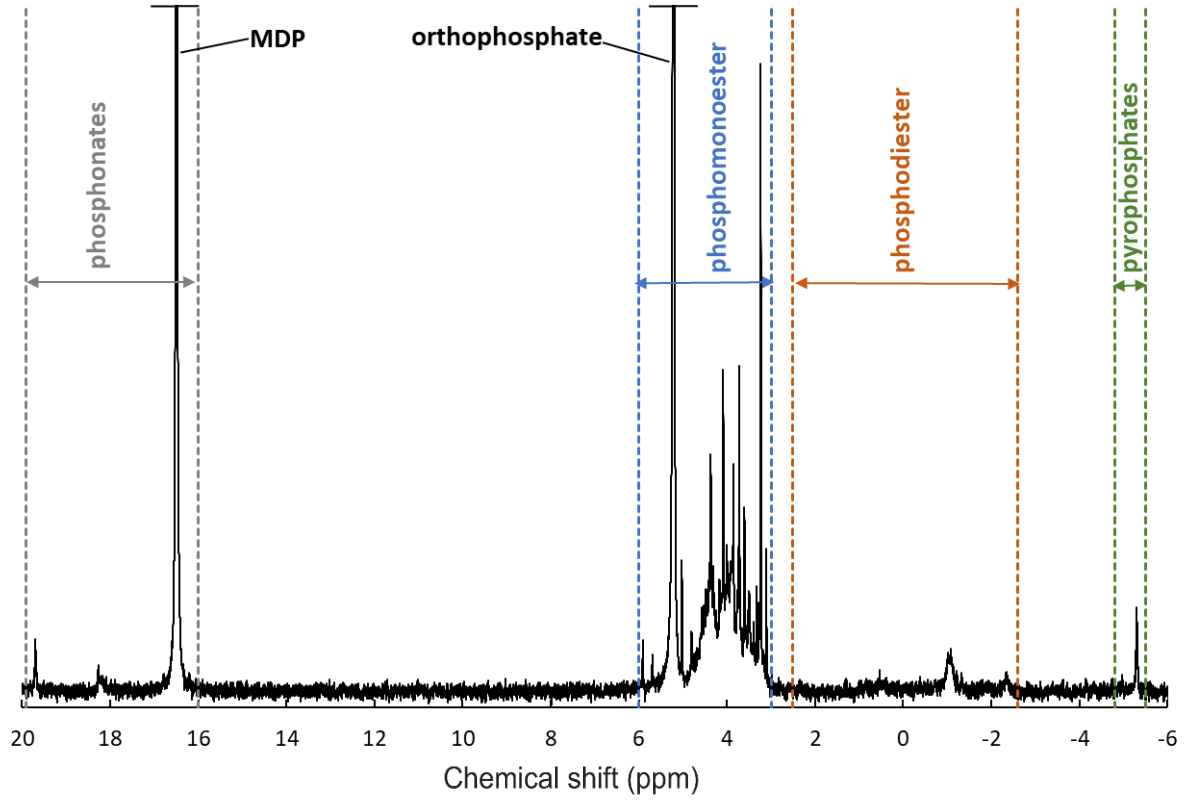


Figure SI-3. Solution ^{31}P NMR spectrum (500 MHz) of 0.25 M NaOH + 0.05 M EDTA Cambisol (F) extract (black) with marked integral regions of phosphonates, phosphomonoester, phosphodiester and pyrophosphates. Signal intensities were normalised to the MDP peak intensity and the vertical axis has been increased by factor 2.

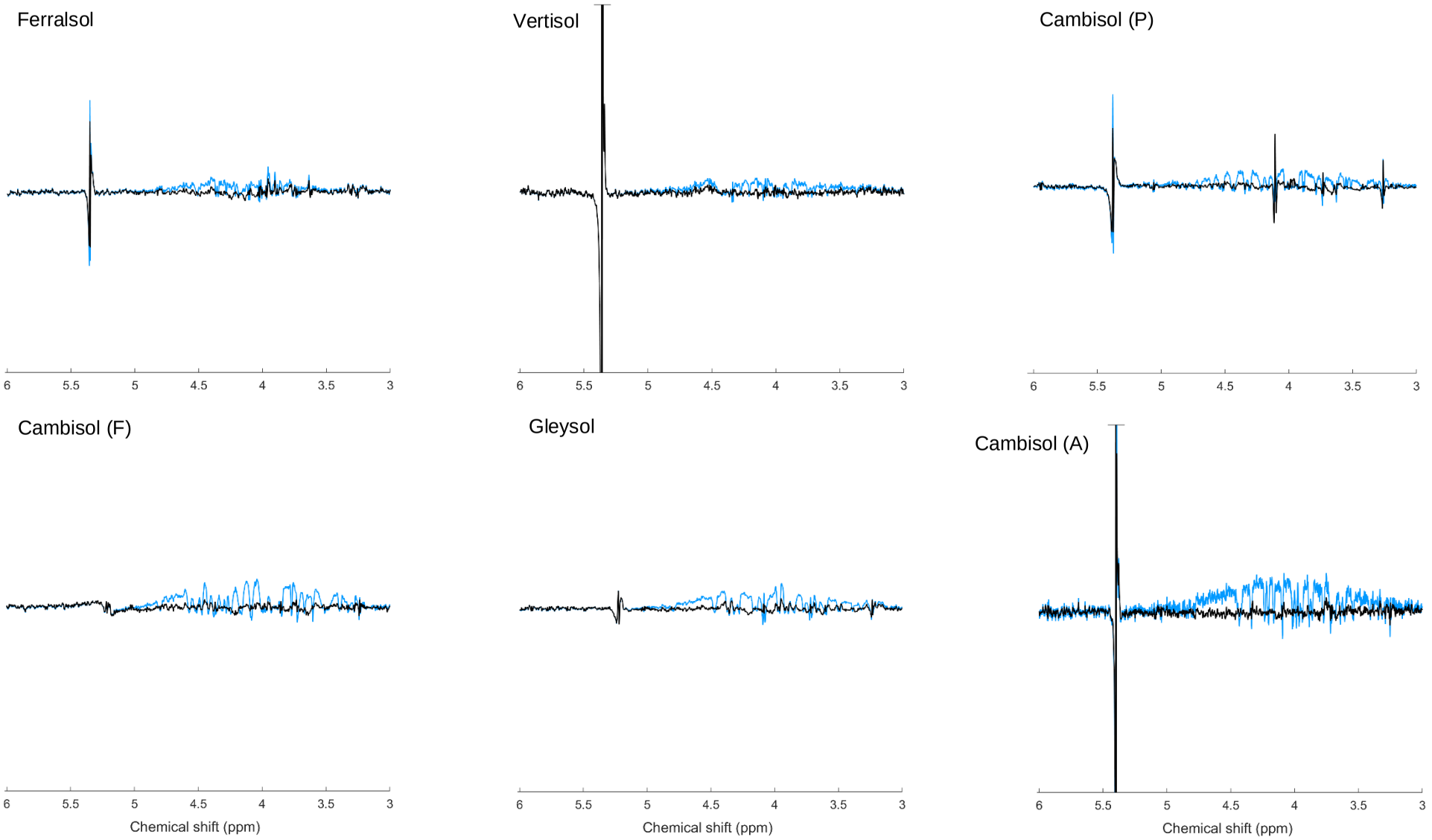


Figure SI-4. Residues of the spectral deconvolution fitting (SDF) procedure of solution ^{31}P NMR spectra of unspiked 0.25 M NaOH + 0.05 M EDTA soil extracts. The residues of the SDF approach including an underlying broad signal are shown in black and the approach without an underlying broad signal in blue. Residue intensities were normalised to the MDP peak intensity and the vertical axes of the spectra have been increased by a factor of 1.7 for the Cambisol (P), the Cambisol (F) and the Gleysol, and by a factor of 10 for the Ferralsol, Vertisol and Cambisol (A) for direct comparison with the initial soil spectra before SDF (Figure 2).

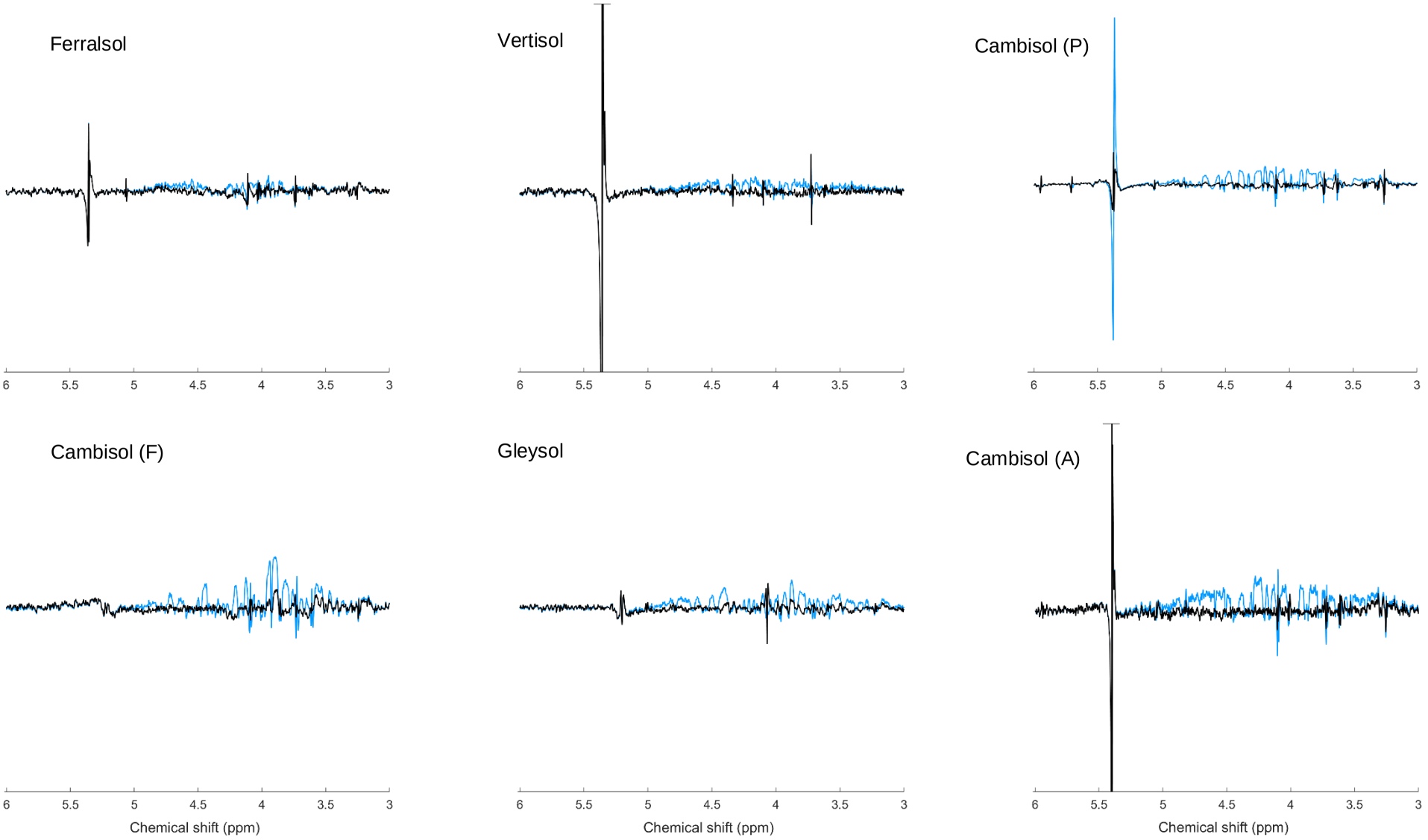


Figure SI-5. Residues of the spectral deconvolution fitting (SDF) procedure of solution ^{31}P NMR spectra of spiked 0.25 M NaOH + 0.05 M EDTA soil extracts. The residues of the SDF approach including an underlying broad signal are shown in black and the approach without an underlying broad signal in blue. Residue intensities were normalised to the MDP peak intensity and the vertical axes of the spectra have been increased by a factor of 1.7 for the Cambisol (P), the Cambisol (F) and the Gleysol, and by a factor of 10 for the Ferralsol, the Vertisol and the Cambisol (A) for direct comparison with the initial soil spectra before SDF (Figure 2).

Table SI-1. Goodness of fit parameters of the spectral deconvolution fitting carried out with the non-linear optimisation algorithm in MATLAB®. The Root-mean-square fitting error, R^2 and Chi^2 for the fitted spectra with and without an underlying broad signal are listed. Chi^2 was calculated using the chemical shift range from δ 5.2 to 3.0 ppm in order to avoid the domination of the orthophosphate peak and its residuals in the calculations. The variances of the 'noise' of the spectra were calculated over the chemical shift range δ 14.1 to 8.1 ppm.

	With underlying broad signal			**Without underlying broad signal**		
Soil	**Root-mean-square fitting error**	**R^2**	**Chi^2**	**Root-mean-square fitting error**	**R^2**	**Chi^2**
Ferralsol	0.33	0.997	4.6	0.40	0.995	19.2
Ferralsol spiked	0.17	0.999	4.4	0.70	0.987	9.0
Vertisol	0.29	0.997	1.7	0.39	0.993	10.5
Vertisol spiked	0.30	0.997	2.7	0.35	0.996	9.5
Cambisol (P)	0.43	0.988	27.5	0.44	0.988	62.9
Cambisol (P) spiked	0.43	0.988	16.1	0.44	0.988	151.1
Cambisol (F)	0.23	0.999	4.3	0.56	0.994	52.2
Cambisol (F) spiked	0.25	0.999	7.5	0.44	0.997	53.1
Gleysol	0.21	0.999	5.5	0.62	0.993	48.5
Gleysol spiked	0.28	0.999	6.7	0.67	0.993	23.0
Cambisol (A)	0.33	0.994	1.6	0.46	0.985	34.2
Cambisol (A) spiked	0.34	0.994	2.6	0.40	0.992	21.2

Table SI-2. Chemical shift, line width at half-height and area of the underlying broad signal as determined by the MATLAB® bootstrap function of the spectral deconvolution fitting procedure. Mean values with standard deviation in parentheses (100 trials).

Soil	Chemical shift (ppm)	Line width (Hz)	Area (-)
Ferralsol	4.20 (1.1E-06)	162 (3.1E-04)	1130800 (839)
Ferralsol spiked	4.22 (2.5E-03)	162 (8.4E-03)	1144189 (2754)
Vertisol	4.06 (1.2E-03)	212 (8.4E-01)	621167 (1450)
Vertisol spiked	4.04 (1.1E-03)	219 (7.2E-01)	661114 (1422)
Cambisol (P)	4.10 (2.2E-04)	229 (1.7E-01)	4370326 (2026)
Cambisol (P) spiked	4.11 (9.9E-04)	202 (1.6E-03)	3644372 (3649)
Cambisol (F)	4.05 (2.0E-04)	217 (5.8E-02)	8564373 (3212)
Cambisol (F) spiked	4.10 (9.4E-08)	202 (3.7E-05)	9035211 (2375)
Gleysol	4.03 (4.5E-04)	221 (2.3E-01)	7383970 (5627)
Gleysol spiked	4.04 (1.2E-04)	243 (5.7E-05)	6842632 (3397)
Cambisol (A)	4.10 (2.3E-06)	223 (5.5E-03)	2535481 (846)
Cambisol (A) spiked	4.15 (4.7E-06)	233 (2.4E-03)	2414768 (1338)

Chapter 2

Identification of lower-order inositol phosphates (IP_5 and IP_4) in soil extracts as determined by hypobromite oxidation and solution ^{31}P NMR spectroscopy

Published as:

Reusser, J.E., Verel, R., Zindel, D., Frossard, E., McLaren, T.I., 2020. Identification of lower-order inositol phosphates (IP_5 and IP_4) in soil extracts as determined by hypobromite oxidation and solution ^{31}P NMR spectroscopy. Biogeosciences 17, 5079-5095.

https://bg.copernicus.org/articles/17/5079/2020/bg-17-5079-2020-discussion.html

Abstract

Inositol phosphates (IP) are a major pool of identifiable organic phosphorus (P) in soil. However, insight into their distribution and cycling in soil remains limited, particularly of lower-order IP (IP_5 and IP_4). This is because the quantification of lower-order IP typically requires a series of chemical extractions, including hypobromite oxidation to isolate IP, followed by chromatographic separation. In this study, we identified the chemical nature of organic P in four soil extracts following hypobromite oxidation using solution ^{31}P NMR spectroscopy and transverse relaxation (T_2) experiments. Soil samples analysed include A horizons from a Ferralsol (Colombia), a Cambisol and a Gleysol from Switzerland, and a Cambisol from Germany. Solution ^{31}P nuclear magnetic resonance (NMR) spectra of the phosphomonoester region in soil extracts following hypobromite oxidation revealed an increase in the number of sharp signals (up to 70), and an on average 2-fold decrease in the concentration of the broad signal compared to the untreated soil extracts. We identified the presence of four stereoisomers of IP_6, four stereoisomers of IP_5, and *scyllo*-IP_4. We also identified for the first time two isomers of *myo*-IP_5 in soil extracts: *myo*-(1,2,4,5,6)-IP_5 and *myo*-(1,3,4,5,6)-IP_5. Concentrations of total IP ranged from 1.4 to 159.3 mg P/kg_{soil} across all soils, of which between 9 % and 50 % were comprised of lower-order IP. Furthermore, we found that the T_2 times, which are considered to be inversely related to the tumbling of a molecule in solution and hence its molecular size, were significantly shorter for the underlying broad signal compared to for the sharp signals (IP_6) in soil extracts following hypobromite oxidation. In summary, we demonstrate the presence of a plethora of organic P compounds in soil extracts, largely attributed to IP of various orders, and provide new insight into the chemical stability of complex forms of organic P associated with soil organic matter.

Introduction

Inositol phosphates (IP) are found widely in nature and are important for cellular functions in living organisms. They are found in eukaryotic cells where they operate in ion-regulation processes, as signalling or P storage compounds (Irvine and Schell, 2001). The basic structure of IP consists of a carbon ring (cyclohexanehexol) with one to six phosphorylated centres (IP_{1-6}) and up to nine stereoisomers (Angyal, 1963; Cosgrove and Irving, 1980). An important IP found in nature is *myo*-IP_6, which is used as a P storage compound in plant seeds. Another important species of IP is *myo*-(1,3,4,5,6)-IP_5, which is present in most eukaryotic cells at concentrations ranging from 15 to 50 µM (Riley et al., 2006). Species of IP_{1-3} are present in phospholipids, such as phosphatidylinositol diphosphates, and are an essential structural component of the cell membrane system (Strickland, 1973; Cosgrove and Irving, 1980).

Inositol phosphates have been reported to comprise more than 50 % of total organic phosphorus (P_{org}) in some soils (Cosgrove and Irving, 1980; McDowell and Stewart, 2006; Turner, 2007). Four stereoisomers of IP have been detected in soils, with the *myo* stereoisomer being the most abundant (56 %), followed by *scyllo* (33 %), *neo* and D-*chiro* (11 %) (Cosgrove and Irving, 1980; Turner et al., 2012). The largest input of *myo*-IP_6 to the soil occurs via the addition of plant seeds (Turner et al., 2002). However, the addition of *myo*-IP_6 to soil can also occur via manure input because monogastric animals are mostly incapable of digesting *myo*-IP_6 without the addition of phytases to their diets (Leytem et al., 2004; Leytem and Maguire, 2007; Turner et al., 2007b). An exception to this is pigs, which were found to at least partially digest phytate (Leytem et al., 2004), and transgenic pigs expressing salivary phytase (Golovan et al., 2001; Zhang et al., 2018). The accumulation of *myo*-IP_6 in soil occurs due to the negative charge of the deprotonated phosphate groups, which can coordinate on the charged surfaces of Fe and Al (hydro)-oxides (Anderson et al., 1974; Ognalaga et al., 1994), clay minerals (Goring and Bartholomew, 1951), and soil organic matter (SOM) (McKercher and Anderson, 1989) or form insoluble precipitates with cations (Celi and Barberis, 2007). These processes lead to the stabilisation of IP in soil resulting in its accumulation and reduced bioavailability (Turner et al., 2002). In contrast, the sources and mechanisms controlling the flux of *scyllo*-, *neo*- and D-*chiro*-IP_6 in soil remain unknown but are thought to involve epimerisation of the *myo* stereoisomer (L'Annunziata, 1975).

Chromatographic separation of alkaline soil extracts revealed the presence of four stereoisomers of IP_6 and lower-order IP_{1-5} (Halstead and Anderson, 1970; Anderson and

Malcolm, 1974; Cosgrove and Irving, 1980; Irving and Cosgrove, 1982). Irving and Cosgrove (1981) used hypobromite oxidation prior to chromatography to isolate the IP fraction in alkaline soils. The basis of this approach is that IP are considered to be highly resistant to hypobromite oxidation, whereas other organic compounds (e.g. phospholipids and nucleic acids) will undergo oxidation (Dyer and Wrenshall, 1941; Turner and Richardson, 2004). The resistance of IP to hypobromite oxidation is thought to be due to the high charge density and steric hindrance, which is caused by the chair conformation of the molecule and the bound phosphate groups, with the P in its highest oxidation state. Hypobromite oxidation of inositol (without phosphate groups) mainly results in the formation of inososes, which have an intact carbon ring (Fatiadi, 1968). Fatiadi (1968) considered that the oxidation of bromine with inositol is stereospecific and comparable to catalytic or bacterial oxidants.

A limitation of chromatographic separation of alkaline extracts is that there is a mixture of unknown organic compounds that can co-elute with IP and result in an overestimation of IP concentrations (Irving and Cosgrove, 1981). However, this can also occur for IP, and historically, studies often reported the combined concentration of IP_6 and IP_5 due to a lack of differentiation in their elution times (McKercher and Anderson, 1968b). More recently, Almeida et al. (2018) investigated how cover crops might mobilise soil IP using hypobromite oxidation on NaOH-EDTA extracts followed by chromatographic separation. The authors found that pools of *myo*-IP_6 and 'unidentified IP' accounted for 30 % of the total extractable pool of P and hypothesised that the unidentified IP pool consists solely of lower-order *myo*-IP. Pools of lower-order $IP_{1\text{-}5}$ comprise on average 17 % of the total pool of IP in soil and account for an important pool of soil organic P in terrestrial ecosystems (Anderson and Malcolm, 1974; Cosgrove and Irving, 1980; Turner et al., 2002; Turner, 2007).

Since the 1980s, solution ^{31}P nuclear magnetic resonance spectroscopy (NMR) has been the most commonly used technique to characterise the chemical nature of organic P in soil extracts (Newman and Tate, 1980; Cade-Menun and Liu, 2014). An advantage of this technique is the simultaneous detection of all forms of organic P that come into solution, which is brought about by a single-step extraction with alkali and a chelating agent (Cade-Menun and Preston, 1996). However, a limitation of the technique has been the loss of information on the diversity and amount of soil IP compared to that typically obtained prior to 1980 (Smith and Clark, 1951; Anderson, 1955; Cosgrove, 1963). To date, solution ^{31}P NMR spectroscopy on soil extracts has only reported concentrations of *myo*-, *scyllo*-, *chiro*-

and *neo*-IP_6. The fact that lower-order IP were not reported in studies using NMR spectroscopy might be due to overlap of peaks in the phosphomonoester region, which makes peak assignment of specific compounds difficult (Doolette et al., 2009).

Turner et al. (2012) carried out hypobromite oxidation prior to solution ^{31}P NMR analysis of alkaline soil extracts to isolate the IP fraction. This had the advantage of reducing the number of NMR signals in the phosphomonoester region and consequently the overlap of peaks. The authors demonstrated the presence of *neo*- and *chiro*-IP_6 in NMR spectra via spiking of hypobromite oxidised extracts. Interestingly, the authors also reported the presence of NMR signals in the phosphomonoester region that could not be assigned to IP_6 and were resistant to hypobromite oxidation. They were not able to attribute the NMR signals to any specific P compounds but hypothesised based on their resistance to hypobromite oxidation that they were due to lower-order IP.

The aim of this study reported here was to identify and quantify IP in soil extracts following hypobromite oxidation using solution ^{31}P NMR spectroscopy. In addition, the structural composition of phosphomonoesters in soil extracts following hypobromite oxidation was probed using solution ^{31}P NMR spectroscopy and transverse relaxation experiments. We hypothesise that a large portion of sharp peaks in the phosphomonoester region of untreated soil extracts are resistant to hypobromite oxidation, which would indicate the presence of a wide variety of IP. This would have major consequences on our understanding of P cycling in terrestrial (and aquatic) ecosystems, as many more organic P compounds and mechanisms would be involved than previously thought. Furthermore, a better understanding of these organic P compounds in soil would also help improve strategies to increase their biological utilisation, which may reduce the amount of fertiliser needed in agricultural systems and thus influence the transfer of P to aquatic and marine ecosystems.

Materials and methods

Soil collection and preparation

Soil samples were collected from the upper horizon of the profile at four diverse sites. These comprise a Ferralsol from Colombia, a Vertisol from Australia, a Cambisol from Germany, and a Gleysol from Switzerland (WRB, 2014). The four soil samples were chosen from a larger collection based on their diverse concentration of P_{org} and composition of the phosphomonoester region in NMR spectra (Reusser et al., 2020a). Background information and some chemical properties of the soils are reported in Table 1. Briefly, the Ferralsol was collected from an improved grassland in 1997 at the Carimagua Research Station's long-term Culticore field experiment in Columbia (Bühler et al., 2003). The Vertisol was collected from an arable field in 2018 located in southern Queensland. The site had been under native shrubland prior to 1992. The Cambisol (F) was collected from a beech forest in 2014 and is part of the "SPP 1685–Ecosystem Nutrition" project (Bünemann et al., 2016; Lang et al., 2017). The Gleysol was collected from the peaty top soil layer of a drained marshland in 2017, which has been under grassland for at least 20 years.

Soil samples were passed through a 5 mm sieve and dried at 60°C for 5 days, except for the Ferralsol (sieved <2 mm) and the Vertisol (ground <2 mm), which were received dried. Total concentrations of C and N in soils were obtained using combustion of 50 mg of ground soil (to powder) weighed into tinfoil capsules (vario PYRO cube®, Elementar Analysensysteme GmbH). Soil pH was measured in H_2O with a soil-to-solution ratio of 1:2.5 (w/w) using a glass electrode.

Table 1. General characteristics of soil samples used in this study.

Parameter	Unit	Ferralsol	Vertisol	Cambisol (F)	Gleysol
Soil type	-	Ferralsol	Vertisol	Dystric Skeletic Cambisol I	Gleysol
Country	-	Colombia	Australia	Germany	Switzerland
Coordinates sampling site	-	4°30' N / 71°19' W	27°52' S / 151°37' E	50°21’ N / 9°55’ E	47°05' N / 8°06' E
Elevation	m ASL	150	402	800	612
Sampling depth	cm	0-20	0-15	0-7	0-10
Year of sampling	year	1997	2017	2014	2017
Land use	-	Pasture	Arable field	Forest	Pasture
C_{tot}	g C/kg_{soil}	26.7	23.9	90.3	148.3
N_{tot}	g N/kg_{soil}	1.7	1.9	6.6	10.9
pH in H_2O	-	3.6	6.1	3.6	5.0

Soil phosphorus analyses

Total concentrations of soil P were obtained by X-ray fluorescence (XRF) spectroscopy (SPECTRO XEPOS ED-XRF, AMETEK®) using 4.0 g of soil sample ground to powder mixed with 0.9 g of wax (CEREOX Licowax, FLUXANA®). The XRF instrument was calibrated using commercially available reference soils. Concentrations of organic P for NMR analysis were obtained using the NaOH-EDTA extraction technique of Cade-Menun et al. (2002) at a soil-to-solution ratio of 1:10, i.e. extracting 4 g of soil with 40 mL of extractant.

Hypobromite oxidation

Hypobromite oxidation of NaOH-EDTA soil filtrates was carried out based on a modified version of the method described in Suzumura and Kamatani (1993) and Turner et al. (2012). The hypobromite oxidation procedure is similar to that reported in Turner (2020). Briefly, 10 mL of the NaOH-EDTA filtrate was placed in a three-necked round-bottom flask equipped with a septum, condenser, magnetic stir bar and thermometer (through a Claisen adapter with N_2 adapter). After the addition of 1 mL of 10 M aqueous NaOH and vigorous stirring, an aliquot of 0.6 mL Br_2 (which was cooled prior to use) was added, resulting in an exothermic reaction where some of the soil extracts nearly boiled. The optimal volume

of Br_2 for oxidation was assessed in a previous pilot study using 0.2, 0.4, 0.6 and 0.8 mL Br_2 volumes, and then observing differences in their NMR spectral features (Figure SI-8). The reaction was heated to 100 °C within 10 min and kept at reflux for an additional 5 min. After cooling to room temperature, the solution was acidified with 2 mL of 6 M aqueous HCl solution in order to obtain a pH <3, which was confirmed with a pH test strip. The acidified solution was reheated to 100 °C for 5 min under a stream of nitrogen to vaporise any excess bromine. The pH of the solution was gradually increased to 8.5 using 10 M aqueous NaOH solution. After dilution with 10 mL of H_2O, 5 mL of 50 % (w/w) ethanol and 10 mL of 10 % (w/w) barium acetate solution was added to the solution in order to precipitate any IP (Turner et al., 2012). The solution was then heated and boiled for 10 min and allowed to cool down overnight. The solution was subsequently transferred to a 50 mL centrifuge tube and a 10 mL aliquot of 50 % (w/w) ethanol was added, manually shaken and centrifuged at 1500 *g* for 15 min. The supernatant was removed, and a 15 mL aliquot of 50 % (w/w) ethanol was added to the precipitate, shaken and then centrifuged again as before. The supernatant was removed and the process repeated once more to further purify the pool of IP. Afterwards, the precipitate was transferred with 20 mL of H_2O into a 100 mL beaker that contained a 20 mL volume (equating to a mass of 15 g) of Amberlite® IR-120 cation exchange resin beads in the H^+ form (Sigma-Aldrich, product no. 06428). The suspension was stirred for 15 min and then passed through a Whatman no. 42 filter paper. A 9 mL aliquot of the filtrate was frozen at -80°C and then lyophilised prior to NMR analysis. This resulted in 18-26 mg of lyophilised material across all soils. Concentrations of total P in solutions were obtained using inductively coupled plasma-optical emission spectrometry (ICP-OES). Concentrations of molybdate reactive P (MRP) were obtained using the malachite green method of Ohno and Zibilske (1991). The difference in concentrations of total P and MRP in solution is molybdate unreactive P (MUP), which is predominantly organic P for these samples. To assess the effect of hypobromite oxidation on the stability of an IP_6, duplicate samples of the Cambisol (F) and the Gleysol were spiked with 0.1 mL of an 11 mM *myo*-IP_6 standard. The recovery of the added *myo*-IP_6 following hypobromite oxidation was calculated using Equation 1:

$$Spike\ recovery\ (\%) = \frac{C_{spiked}\left(\frac{mg}{L}\right) - C_{unspiked}\left(\frac{mg}{L}\right)}{C_{standard\ added}\left(\frac{mg}{L}\right)}, \quad (1)$$

where C_{spiked} and $C_{unspiked}$ are the concentrations of *myo*-IP_6 in NaOH-EDTA extracts following hypobromite oxidation of the spiked and unspiked samples, respectively. $C_{standard\ added}$ is the concentration of the added *myo*-IP_6 within the standard. As ^{31}P NMR

spectroscopy of the standard revealed impurities, the concentration of *myo*-IP_6 in the standard was calculated based on the ^{31}P NMR spectrum.

Sample preparation for solution ^{31}P NMR spectroscopy

The lyophilised material of the untreated soil extracts was prepared for solution ^{31}P NMR spectroscopy based on a modification of the methods of Vincent et al. (2013) and Spain et al. (2018). Briefly, 120 mg of lyophilised material was taken and dissolved in 600 µL of 0.25 M NaOH-0.05 M EDTA solution (ratio of 1:5). However, for the Cambisol (F) sample, this ratio resulted in an NMR spectrum that exhibited significant line broadening. Therefore, this was repeated on a duplicate sample but at a smaller lyophilised-material-to-solution ratio (ratio of 1:7.5), as suggested in Cade-Menun and Liu (2014), which resolved the issue of poor spectral quality. The suspension was stored overnight to allow for complete hydrolysis of phospholipids and RNA (Doolette et al., 2009; Vestergren et al., 2012) and was then centrifuged at 10621 *g* for 15 min. A 500 µL aliquot of the supernatant was taken, which was subsequently spiked with a 25 µL aliquot of a 0.03 M methylenediphosphonic acid standard made in D_2O (Sigma-Aldrich, product no. M9508) and a 25 µL aliquot of sodium deuteroxide at 40 % (w/w) in D_2O (Sigma-Aldrich, product no. 372072). The solution was then mixed and transferred to a 5 mm diameter NMR tube.

A similar procedure was used for the soil extracts that had undergone hypobromite oxidation, except the total mass of lyophilised material (18-26 mg) was dissolved with 600 µL of a 0.25 M NaOH-0.05 M EDTA solution. However, for the Cambisol (F) sample, the NMR spectrum exhibited considerable line broadening, and an additional 400 µL aliquot of NaOH-EDTA solution was added to the NMR tube, mixed and then returned to the NMR spectrometer. This resolved the issue of poor spectral quality.

Solution ^{31}P NMR spectroscopy

Solution ^{31}P NMR analyses were carried out on all untreated and hypobromite oxidised soil extracts at the NMR facility of the Laboratory of Inorganic Chemistry (Hönggerberg, ETH Zürich). All spectra were obtained with a Bruker AVANCE III MD 500 MHz NMR spectrometer equipped with a cryogenic probe (CryoProbe™ Prodigy) (Bruker Corporation; Billerica, MA). The ^{31}P frequency for this NMR spectrometer was 202.5 MHz and gated broadband proton decoupling with a 90° pulse of 12 µs was applied. The spectral resolution under these conditions for ^{31}P was <1 Hz. Longitudinal relaxation (T_1) times were determined for each sample with an inversion recovery experiment (Vold et al.,

1968). This resulted in recycle delays ranging from 8.7 to 30.0 s for the untreated extracts and 7.8 to 38.0 s for the hypobromite oxidised soil extracts. The number of scans for the untreated extracts was set to 1024 or 4096, depending on the signal-to-noise ratio of the obtained spectrum. All hypobromite oxidised spectra were acquired with 3700 to 4096 scans.

Processing of NMR spectra

All NMR spectra were processed with Fourier transformation, phase correction, and baseline adjustment within the TopSpin® software environment (version 3.5 pl 7, Bruker Corporation; Billerica, MA). Line broadening was set to 0.6 Hz. Quantification of NMR signals involved obtaining the integrals of the following regions: 1) up to four phosphonates (δ 19.8 to 16.4 ppm); 2) the added MDP (δ 17.0 to 15.8ppm) including its two carbon satellite peaks; 3) the combined orthophosphate and phosphomonoester region (δ 6.0 to 3.0 ppm); 4) up to four phosphodiesters (δ 2.5 to -3.0 ppm), and 5) pyrophosphate (δ -4.8 to -5.4 ppm). Due to overlapping peaks in the orthophosphate and phosphomonoester region, spectral deconvolution fitting (SDF) was applied as described in Reusser et al. (2020a). In brief, the SDF procedure involved the fitting of an underlying broad signal, based on the approach of Bünemann et al. (2008b) and McLaren et al. (2019). We carried out the SDF with a non-linear optimisation algorithm in MATLAB® R2017a (The MathWorks, Inc.) and fitted visually identifiable peaks by constraining their line widths at half-height as well as the lower and upper boundary of the peak positions along with an underlying broad signal in the phosphomonoester region. The sharp signals of high intensity (e.g. orthophosphate) and the broad peak were fitted using Lorentzian line shapes, whereas sharp signals of low intensity were fitted using Gaussian line shapes. The NMR observability of total P (P_{tot}) in NaOH-EDTA extracts was calculated using Equation 2 (Dougherty et al., 2005; Doolette et al., 2011a):

$$NMR\ observability\ (\%) = \frac{P_{tot}\ NMR}{P_{tot}\ ICP{-}OES} * 100\ \%\ , \qquad (2)$$

where P_{tot} NMR refers to the total P content in mg P/kg_{soil} detected in the soil extracts using solution ^{31}P NMR spectroscopy and P_{tot} ICP-OES refers to the total P concentration in mg P/kg_{soil} measured in the soil extracts prior to freeze-drying using ICP-OES.

Spiking experiments

To identify the presence of IP in hypobromite oxidised extracts, samples were spiked with a range of standards and then analysed again using NMR spectroscopy. This involved the addition of 5 to 20 µL aliquots of an IP standard solution directly into the NMR tube, which was then sealed with parafilm, manually shaken, and then allowed to settle prior to NMR analysis. Each sample extract was consecutively spiked with no more than four IP standards. The NMR spectra of soil extracts after spiking were overlaid with the NMR spectra of unspiked soil extracts to identify the presence of IP across all soil samples. This comparison of NMR spectra was possible due to negligible changes in the chemical shifts of peaks among soil samples. The IP standards used in this study are listed in Table 2.

Table 2. Standard solutions used for the spiking experiment of the hypobromite oxidised soil extracts. All standards were dissolved in 0.25 M NaOH and 0.05 M EDTA.

Standard	Product number	Company or origin	Concentration of standard in NaOH-EDTA (mg/mL)
myo-IP_6	P5681	Merck (Sigma-Aldrich)	8.10
L-*chiro*-IP_6	Collection of Dr Max Tate		2.39
-*chiro*-IP_6	CAY-9002341	Cayman Chemical	2.00
neo-IP_6	Collection of Dr Dennis Cosgrove[1]		4.62
D-*myo*-(1,2,4,5,6)-IP_5	CAY-10008452-1	Cayman Chemical	2.00
myo-(1,2,3,4,6)-IP_5	93987	Merck (Sigma-Aldrich)	2.00
D-*myo*-(1,3,4,5,6)-IP_5	CAY-10009851-1	Cayman Chemical	2.00
D-*myo*-(1,2,3,5,6)-IP_5	CAY-10008453-1	Cayman Chemical	2.00
scyllo-IP_5	Collection of Dr Dennis Cosgrove		2.64
L-*chiro*-IP_5	Collection of Dr Dennis Cosgrove		2.24
neo-IP_5	Collection of Dr Dennis Cosgrove		2.45
myo-IP_4	Collection of Dr Dennis Cosgrove		2.76
scyllo-IP_4	Collection of Dr Dennis Cosgrove		2.41
neo-IP_4	Collection of Dr Dennis Cosgrove		2.33

[1] Made up in 15 mM HCl

Transverse relaxation (T_2) experiments

Due to the presence of sharp and broad signals in the phosphomonoester region of NMR spectra on hypobromite oxidised soil extracts, transverse relaxation (T_2) experiments were carried out to probe their structural composition. The transverse relaxation (originally spin-spin relaxation) describes the loss of magnetisation in the x-y plane. This loss occurs due to magnetic field differences in the sample, arising either by instrumentally caused magnetic field inhomogeneities or by local magnetic fields in the sample caused by intramolecular and intermolecular interactions (Claridge, 2016b). Generally, small, rapidly tumbling molecules exhibit longer T_2 relaxation times compared to large, slowly tumbling molecules (McLaren et al., 2019).

Briefly, solution ^{31}P NMR spectroscopy with a Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence (Meiboom and Gill, 1958) was carried out on all hypobromite oxidised soil extracts, as described in McLaren et al. (2019). This involved a constant spin-echo delay (τ) of 5 ms, which was repeated for a total of eight iterations (spin-echo periods of 5, 50, 100, 150, 200, 250, 300 and 400 ms). A total of 4096 scans and a recycle delay of 4.75 s were used for all iterations. Transverse relaxation times for the aforementioned integral ranges were calculated using Equation 3 within the TopSpin® software environment. Due to overlapping peaks in the orthophosphate and phosphomonoester region, spectral deconvolution was carried out to partition the NMR signal, as described in McLaren et al. (2019). The T_2 times of the partitioned NMR signals were calculated using Equation 3 within RStudio© (version 1.1.442):

$$M(t) = M_0 * e^{(-t*{T_2}^{-1})}\ , \tag{3}$$

where M refers to the net magnetisation derived from the average angular momentum in the x–y plane, τ refers to the spin-echo delay in milliseconds (ms), and T_2 refers to the transverse relaxation time (ms).

Statistical analyses and graphics

Statistical analyses were carried out using Microsoft® Excel 2016 and MATLAB R2017a (©The MathWorks, Inc.). Graphics were created with Microsoft® Excel 2016 and MATLAB R2017a (©The MathWorks, Inc.). Solution (1D) ^{31}P NMR spectra were normalised to the peak intensity of MDP (δ 16.46 ppm). Spectra from the T_2 experiments were normalised to the peak intensity of *scyllo*-IP_6 (δ 3.22 ppm).

A one-way ANOVA was carried out in MATLAB R2017a (©The MathWorks, Inc.) with a subsequent multi-comparison of mean values using Tukey's honestly significant difference procedure based on the Studentised range distribution (Hochberg and Tamhane, 1987; Milliken and Johnson, 2009).

Results

Phosphorus concentrations in soil extracts

Concentrations of total soil P as determined by XRF ranged from 320 to 3841 mg P/kg_{soil} across all soils (Table 3). Concentrations of total P as estimated by the NaOH-EDTA extraction technique ranged from 160 to 1850 mg P/kg_{soil}, which comprised 28 % to 51 % of the total soil P as determined by XRF. Pools of organic P comprised 28 % to 72 % of the total P in NaOH-EDTA untreated soil extracts.

Concentrations of total P in NaOH-EDTA soil extracts following hypobromite oxidation ranged from 77 to 578 mg P/kg_{soil} (Table 3), which accounted for 31 % to 48 % (on average 38 %) of the total P originally present in the extracts. Similarly, pools of organic P in NaOH-EDTA extracts following hypobromite oxidation were lower, comprising 22 % to 48 % (on average 36 %) of that originally present in untreated NaOH-EDTA extracts across all soils.

Table 3. Concentrations of total P as measured by XRF and 0.25 M NaOH + 0.05 M EDTA extractable P before and after hypobromite oxidation of soil extracts. Concentrations of total P in NaOH-EDTA extracts were determined by ICP-OES, whereas the concentration of molybdate reactive P (MRP) was determined by the malachite green method of Ohno and Zibilske (1991). Concentrations of molybdate unreactive P (MUP) were calculated as the difference between total P and MRP.

Measure	Unit	Ferralsol	Vertisol	Cambisol (F)	Gleysol
XRF	P_{tot} (mg P/kg_{soil})	320	1726	3841	2913
NaOH-EDTA extractable P (untreated)	P_{tot} (mg P/kg_{soil})	160	484	1850	1490
	MRP (mg P/kg_{soil})	67	351	525	610
	MUP (P_{org}) (mg P/kg_{soil})	93	133	1326	880
NaOH-EDTA extractable P (hypobromite oxidised)	P_{tot} (mg P/kg_{soil})	77	158	580	578
	MRP (mg P/kg_{soil})	32	111	283	231
	MUP (P_{org}) (mg P/kg_{soil})	45	47	297	348

Solution ^{31}P NMR spectra of hypobromite oxidised soil extracts

The most prominent signal in the NMR spectra of untreated NaOH-EDTA soil extracts was that of orthophosphate at δ 5.25 (±0.25) ppm, followed by the phosphomonoester region ranging from δ 6.0 to 3.0 ppm (Figures 1a + b). There were also some minor signals due

to pyrophosphate δ -5.06 (±0.19) ppm (all soils), phosphodiesters ranging from δ 2.5 to - 2.4 ppm (not detected in the Vertisol), and phosphonates (not including the added MDP) at δ 19.8, 19.2 and 18.3 ppm (not detected in the Gleysol). However, these compounds comprised less than 8 % of the total NMR signal.

Following hypobromite oxidation of NaOH-EDTA extracts, the most prominent NMR signals were found in the orthophosphate (65 % of total NMR signal) and phosphomonoester (35 % of total NMR signal) region across all soils (Figures 1a + 1b). Phosphodiesters and pyrophosphates were removed following hypobromite oxidation in the Ferralsol, the Vertisol and the Cambisol (F). However, some signal remained in the Gleysol at low concentrations (0.4 % of the total NMR signal). Phosphonates were removed following hypobromite oxidation in the Ferralsol and the Vertisol, but a total of five sharp peaks in the phosphonate region were detected (δ 19.59, 18.58, 17.27 and 9.25 ppm) in the Cambisol (F). These peaks comprised 0.6 % of the total NMR signal.

The phosphomonoester region of NMR spectra on untreated NaOH-EDTA extracts exhibited two main features: 1) the presence of a broad signal centred at around δ 4.1 ppm (±0.1) with an average line width at half-height of 256.12 Hz; and 2) the presence of between 19 and 34 sharp signals. This was similarly the case for hypobromite oxidised extracts, except there was a decrease in the intensity of the broad signal and a change in the distribution and intensity of sharp signals. For the Cambisol (F) and Gleysol, the number of sharp signals in the phosphomonoester region approximately doubled (to 40 and 70 sharp signals, respectively) following hypobromite oxidation. In contrast, less than half of the sharp signals remained in the Ferralsol following hypobromite oxidation (i.e. 14 of the 30 peaks originally present in the untreated extract), whereas one peak was removed following hypobromite oxidation in the Vertisol. There was little change (0.23 ppm) in the chemical shifts of peaks between the untreated and hypobromite oxidised extracts.

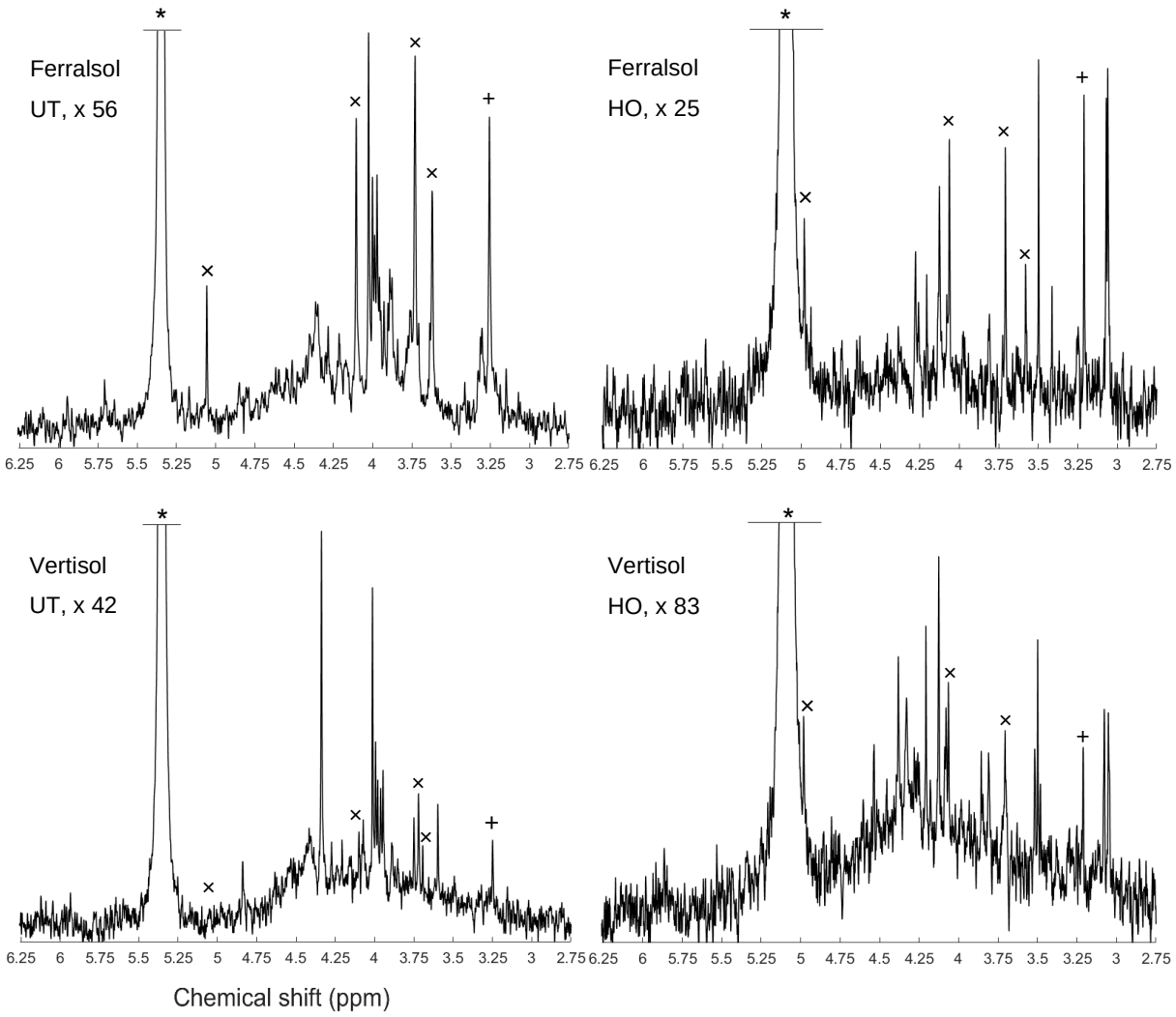


Figure 1a. Solution ^{31}P nuclear magnetic resonance (NMR) spectra (500 MHz) of the orthophosphate and phosphomonoester region on untreated (UT) and hypobromite oxidised (HO) 0.25 M NaOH + 0.05 M EDTA soil extracts Ferralsol and Vertisol. Signal intensities were normalised to the MDP peak intensity. The vertical axes were increased for improved visibility of spectral features, as indicated by a factor. The orthophosphate peak is marked with an asterisk. The symbol '×' marks the four individual peaks of *myo*-IP_6 and '+' the peak of *scyllo*-IP_6.

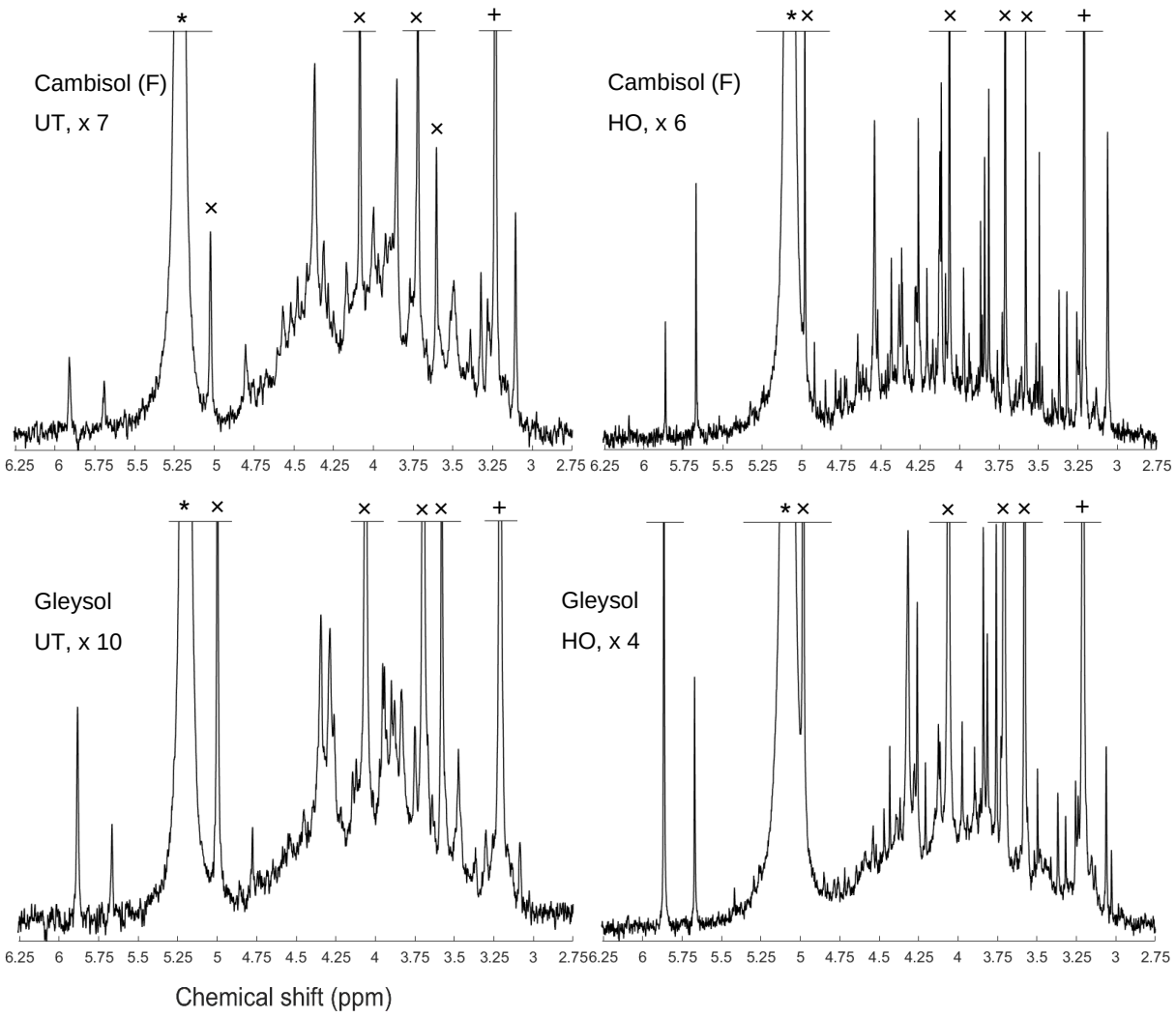


Figure 1b. Solution ^{31}P nuclear magnetic resonance (NMR) spectra (500 MHz) of the orthophosphate and phosphomonoester region on untreated (UT) and hypobromite oxidised (HO) 0.25 M NaOH + 0.05 M EDTA soil extracts Cambisol (F) and Gleysol. Signal intensities were normalised to the MDP peak intensity. The vertical axes were increased for improved visibility of spectral features, as indicated by a factor. The orthophosphate peak is marked with an asterisk. The symbol '×' marks the four individual peaks of *myo*-IP_6 and '+' the peak of *scyllo*-IP_6.

Table 4. Concentrations (mg P/kg_{soil}) of P compounds in solution ^{31}P NMR spectra of 0.25 M NaOH + 0.05 M EDTA soil extracts (Ferralsol, Vertisol, Cambisol (F) and Gleysol) before and after hypobromite oxidation (HO). Quantification was based on spectral integration and deconvolution fitting. The proportion of P detected in hypobromite oxidised extracts compared to that in untreated extracts is provided in brackets.

Phosphorus class		Ferralsol	Vertisol	Cambisol (F)	Gleysol
Phosphonates	before HO	1.0	2.6	14.5	-
	after HO	-	-	3.0	0.2
				(21)	
Orthophosphate	before HO	54.8	221.4	434.3	368.3
	after HO	32.0	116.6	329.3	243.4
		(58)	(53)	(76)	(66)
Phosphomonoester	before HO	36.3	39.1	501.1	399.2
	after HO	12.7	24.2	210.3	292.1
		(35)	(62)	(42)	(73)
Broad peak in phosphomonoester region	before HO	21.6	30.9	305.8	216.7
	after HO	8.3	19.3	99.2	108.4
		(39)	(63)	(32)	(50)
Phosphodiester	before HO	5.1	-	28.2	26.9
	after HO	-	-	-	2.0
					(8)
Pyrophosphate	before HO	1.9	1.8	12.9	23.9
	after HO	-	-	-	-

Identification and quantification of inositol phosphates (IP_6, IP_5 and IP_4) in soil extracts

Detailed views of the phosphomonoester regions of spiked samples are shown in Figure SI-1 to SI-5 of the supporting information. The number of identified sharp peaks in the phosphomonoester region ranged from 7 (Vertisol) to 33 (Gleysol). *myo*- and *scyllo*-IP_6 were identified in the hypobromite oxidised extracts of all soils (Table 5). On average, 72 % of *myo*-IP_6 and 56 % of *scyllo*-IP_6 present in the untreated extracts remained in the hypobromite oxidised extracts (Table SI-1 in the supporting information). *neo*-IP_6 was identified in the 2-equatorial-4-axial and 4-equatorial-2-axial conformations, and c*hiro*-IP_6 in the 2-equatorial-4-axial confirmation, of the oxidised extracts in the Cambisol (F) and Gleysol but was absent in the Ferralsol and the Vertisol (Figures SI-4 and SI-5).

The *myo*, *scyllo*, *chiro* and *neo* stereoisomers of IP_5 were identified in various hypobromite oxidised extracts (Table 5). Two isomers of *myo*-IP_5 were identified in some extracts, which included *myo*-(1,2,4,5,6)-IP_5 and *myo*-(1,3,4,5,6)-IP_5. In addition, *scyllo*-IP_4 was detected

in all soils except that of the Vertisol. There was insufficient evidence for the presence of *myo*-IP_4 in these soil samples, as only one of the two peaks of this compound was present in the NMR spectra of untreated extracts. This could possibly be due to the partial dephosphorylation of *myo*-IP_4 during the hypobromite oxidation procedure. The reason for the reduced resistance of lower-order IP to hypobromite oxidation compared to IP_{5+6} might be due to their reduced steric hindrance and charge density, as fewer phosphate groups are bound to the inositol ring.

Concentrations of total IP ranged from 1.4 to 159.3 mg P/kg_{soil} across all soils, which comprised between 1 % (Vertisol) and 18 % (Gleysol) of the organic P in untreated NaOH-EDTA extracts (Table 3). Pools of IP_6 were the most abundant form of IP, which ranged from 0.9 to 144.8 mg P/kg_{soil} across all soils (Table 5). The proportion of IP_6 stereoisomers across all soils was on the order of *myo* (61 %, SD=12), *scyllo* (29 %, SD=3), *chiro* (6 %, SD=8) and *neo* (4 %, SD=5). Similarly, the *myo* and *scyllo* stereoisomers were also the most predominant forms of IP_5 but comprised between 83 % (Cambisol (F)) and 100 % (Ferralsol and Vertisol) of total IP_5 (Table 5). Trace amounts of *scyllo*-IP_4 were also detected in three of the four soils. The ratio of total IP_6 to IP_5 differed across all soils (Figure 2).

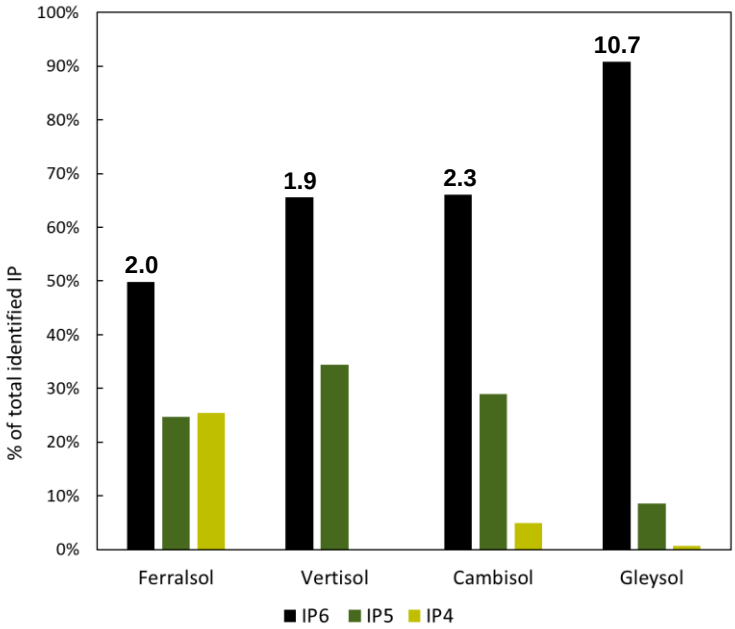


Figure 2. The proportion of total identifiable pools of inositol hexakisphosphates (IP_6), -pentakis-phosphates (IP_5) or -tetrakisphosphates (IP_4) in the total pool of identifiable IP, as determined by solution ^{31}P NMR spectroscopy on four soil extracts (Ferralsol, Vertisol, Cambisol (F) and Gleysol) following hypobromite oxidation. Values located above the IP_6 bar are the ratio of total identifiable IP_6 to total identifiable IP_5 in each soil sample.

Table 5. Concentrations of identified inositol phosphates (IP) in hypobromite oxidised 0.25 M NaOH + 0.05 M EDTA soil extracts (Ferralsol, Vertisol, Cambisol (F) and Gleysol). Concentrations were calculated from solution ^{31}P NMR spectra using spectral deconvolution fitting including an underlying broad signal. When no concentration is given, the IP compound was not detected in the respective soil extract. Chemical shift positions are based on the NMR spectrum of the Cambisol (F) extract (Figure SI-7). Peak positions varied up to +0.018 ppm (Gleysol). Conformation equatorial (eq) and axial (ax) according to Turner et al. (2012).

Phosphorus compound	Chemical shift δ ppm	Concentrations (mg P/kg$_{soil}$)			
		Ferralsol	Vertisol	Cambisol (F)	Gleysol
***myo*-IP$_6$**	4.97, 4.06, 3.70, 3.57	1.1	0.6	26.3	85.0
***scyllo*-IP$_6$**	3.20	0.4	0.3	15.6	41.1
***neo*-IP$_6$ 4-eq-2-ax**	5.86, 3.75	-	-	1.4	8.8
***neo*-IP$_6$ 2-eq-4-ax**	4.36, 4.11	-	-	4.0	1.3
D-*chiro*-IP$_6$ 2-eq-4-ax	5.66, 4.25, 3.83	-	-	9.4	8.6
***myo*-(1,2,4,5,6)-IP$_5$**	4.42, 3.97, 3.72, 3.36, 3.25	-	-	7.0	4.1
***myo*-(1,3,4,5,6)-IP$_5$**	4.12, 3.60, 3.23	-	-	2.8	1.3
***scyllo*-IP$_5$**	3.81, 3.31, 3.05	0.7	0.5	10.8	6.1
***neo*-IP$_5$**	4.64, 4.27, 4.01, 3.87, 3.13	-	-	3.3	2.1
***chiro*-IP$_5$**	4.61, 3.39	-	-	0.9	-
***scyllo*-(1,2,3,4)-IP$_4$**	4.12, 3.25	0.8	-	4.3	1.0
Total IP		3.0	1.4	85.9	159.3

If sharp peaks arising from IP were identified in the NMR spectra on hypobromite oxidised extracts, a comparison was made with that of their corresponding untreated extracts. The sharp peaks of all stereoisomers of IP_6 were present in the untreated extracts. The five peaks of *myo*-(1,2,4,5,6)-IP_5 and the three peaks of *scyllo*-IP_5 were also identified. However, it was not possible to clearly identify other IP_5 compounds in untreated extracts due to overlapping signals. In the Gleysol, all three peaks of *scyllo*-IP_5 were detected, but only two of the possible five peaks could be clearly assigned to *myo*-(1,2,4,5,6)-IP_5. In the Ferralsol, both peaks of *scyllo*-IP_4 were present in the untreated extract, but only two of the three possible peaks could be assigned to *scyllo*-IP_5. In the Vertisol, no IP_5 was identified. Concentrations of IP in untreated extracts assessed by spectral deconvolution fitting were generally double that measured in hypobromite oxidised extracts. Recoveries

of added *myo*-IP_6 in the Gleysol and Cambisol (F) following hypobromite oxidation were 47 % and 20 %, respectively.

Spin-echo analysis of selected P compounds

Due to the presence of sharp and broad signals in hypobromite oxidised soil extracts, the structural composition of phosphomonoesters was probed. A comparison of the NMR spectra at the lowest (1*τ) and highest (80*τ) pulse delays revealed a fast-decaying broad signal for all hypobromite oxidised soil extracts, which was particularly evident in the Gleysol (Figure 3). Calculated T_2 times of all IP_6 stereoisomers were longer than those of the broad signal (Table 6). The T_2 times of *scyllo*-IP_6 (on average 175.8 ms, SD=49.7) were generally the longest of all stereoisomers of IP_6. The T_2 time of the orthophosphate peak was the shortest, which was on average 11.5 ms (SD=4.9).

The average (n=4) T_2 times of the broad peak was significantly different than that of *scyllo*- and *myo*-IP_6 ($p < 0.05$). Significant differences in the T_2 times of *neo*- and D-*chiro*-IP_6 were not tested, as these compounds were not detected in the Ferralsol and the Vertisol.

Table 6. Transversal relaxation times (T_2) of various P species in the orthophosphate and phosphomonoester regions as determined by solution ^{31}P nuclear magnetic resonance (NMR) spectroscopy and a Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence on hypobromite oxidised soil extracts.

Phosphorus compound	T_2 (ms)			
	Ferralsol	Vertisol	Cambisol (F)	Gleysol
***myo*-IP_6**	163	140	139	121
***scyllo*-IP_6**	250	155	154	144
***neo*-IP_6**	-	-	203	102
***D-chiro*-IP_6**	-	-	108	132
Orthophosphate	14	9	17	6
Broad peak	44	69	89	62

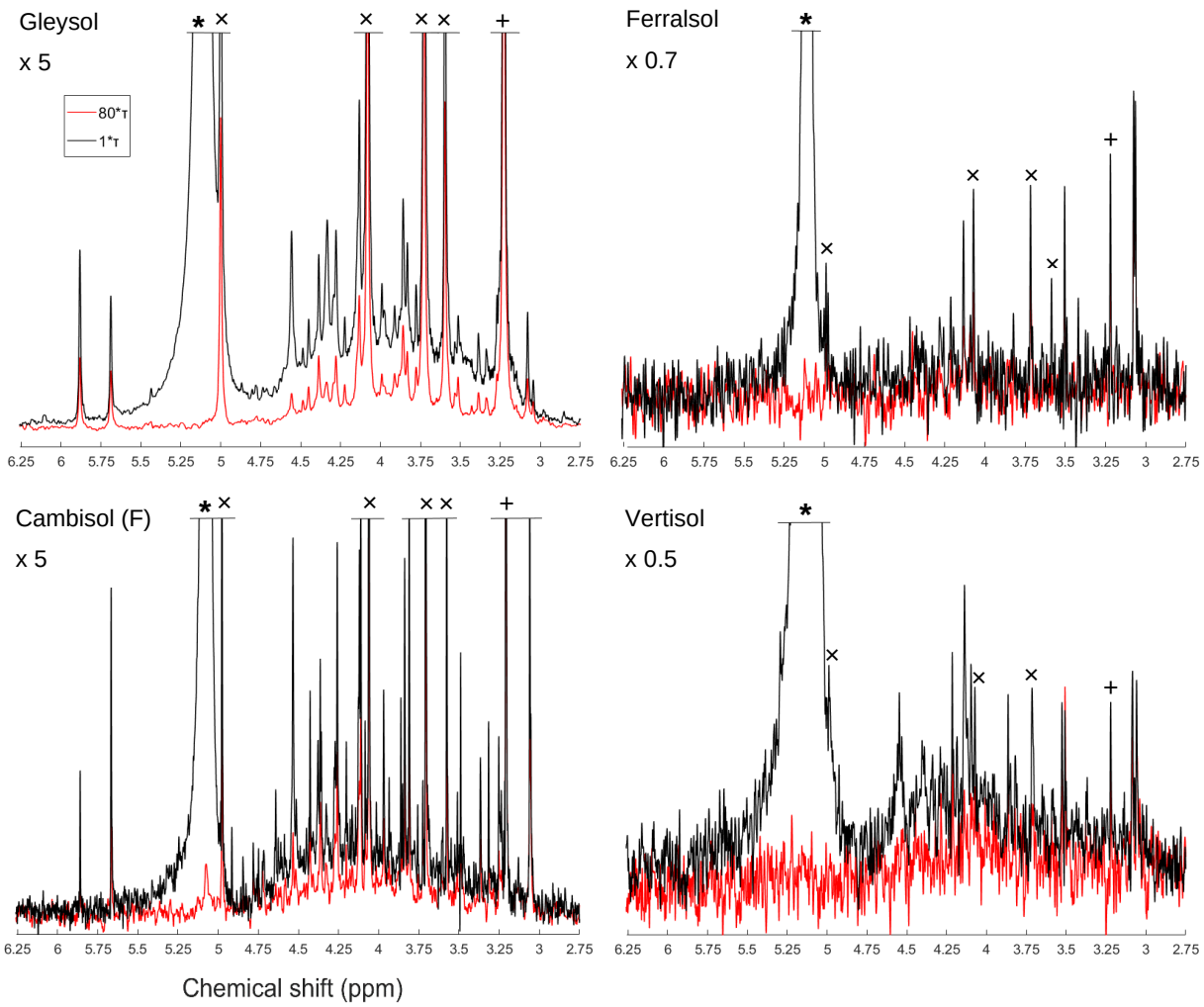


Figure 3. Solution ^{31}P NMR spectra of hypobromite oxidised soil extracts acquired with a CPMG pulse sequence with 1*τ (black) and 80*τ (red) spin-echo delays. The orthophosphate (*), *scyllo*-IP$_6$ (+) and *myo*-IP$_6$ peaks (×) are marked accordingly. Spectra were normalised to the maximum *scyllo*-IP$_6$ peak intensity in the 1*τ spectrum for each soil. The vertical axes were increased/decreased for better visualisation by an indicated factor.

Discussion

Pools of phosphorus in untreated and hypobromite oxidised soil extracts

On average, 44 % of total P (as measured with XRF) was extracted by NaOH-EDTA, which is consistent with previous studies (Turner, 2008; Li et al., 2018; McLaren et al., 2019). The non-extractable pool of P is likely to comprise inorganic P as part of insoluble mineral phases but could also contain some organic P (McLaren et al., 2015a). Nevertheless, the NaOH-EDTA extraction technique is considered to be a measure of total organic P in soil, which can be subsequently characterised by solution ^{31}P NMR spectroscopy (Cade-Menun and Preston, 1996).

Hypobromite oxidation resulted in a decrease in the concentration of inorganic and organic P in NaOH-EDTA extracts across all soils. The decrease of organic P is consistent with previous studies (Turner and Richardson, 2004; Turner et al., 2012; Almeida et al., 2018). However, Almeida et al. (2018) reported an overall increase in the concentration of inorganic P following hypobromite oxidation, which the authors proposed to be caused by the degradation of organic P forms not resistant to hypobromite oxidation. A decrease in the concentration of organic P in NaOH-EDTA extracts following hypobromite oxidation was expected based on the oxidation of organic molecules containing P. The products of hypobromite oxidation are most probably carbon dioxide, simple organic acids from the oxidative cleavage of the phosphoesters and orthophosphate (Irving and Cosgrove, 1981; Sharma, 2013).

Overall, hypobromite oxidation of NaOH-EDTA soil extracts resulted in a considerable increase in the number of sharp peaks and a decrease in the broad underlying peak in the phosphomonoester region compared to untreated soil extracts. This was particularly the case for the Cambisol (F) and the Gleysol, which had high concentrations of extractable organic P. Since the broad peak is thought to be closely associated with the SOM (Dougherty et al., 2007; Bünemann et al., 2008b; McLaren et al., 2015b), its decrease in soil extracts following hypobromite oxidation is consistent with that observed for other organic compounds (Turner et al., 2012). Our results indicate that a majority of sharp peaks present in the phosphomonoester region of untreated soil extracts are stable to hypobromite oxidation and are therefore likely to be IP.

Across all soils, 5 to 15 peaks in the phosphomonoester region were removed following hypobromite oxidation compared to those in untreated extracts, which are likely due to the

oxidation of: α-and β-glycerophosphate (Doolette et al., 2009; McLaren et al., 2015b), RNA mononucleotides (8 peaks) (Vincent et al., 2013), glucose 6-phosphate, phosphocholine, glucose 1-phosphate, or phosphorylethanolamine (Cade-Menun, 2015).

Phosphorus assignments of sharp peaks in hypobromite oxidised extracts

The detection of *myo-, scyllo*-, *chiro*, and *neo*-IP_6 in untreated and hypobromite oxidised soil extracts is consistent with previous studies using chromatography (Irving and Cosgrove, 1982; Almeida et al., 2018) and NMR (Turner and Richardson, 2004; Doolette et al., 2011b; Vincent et al., 2013; Jarosch et al., 2015; McLaren et al., 2015b). Turner et al. (2012) suggested that hypobromite oxidised extracts only contained *neo*-IP_6 in the 4-equatorial-2-axial conformation due to the absence of signals from the 2-equatorial-4-axial conformation. In the current study, both conformations could be identified in two of the four soil extracts, which is likely due to improved spectral resolution and sensitivity. The relative abundances of the four identified stereoisomers of IP_6 in soil extracts were similar to previous studies (Irving and Cosgrove, 1982; Turner et al., 2012).

Several studies have shown overlap of peaks relating to RNA mononucleotides and those of α-and β-glycerophosphate, which are the alkaline hydrolysis products of RNA and phospholipids, respectively. However, in the current study, several sharp peaks were present in hypobromite oxidised extracts which are in the chemical shift range of RNA mononucleotides and α-and β-glycerophosphate. Whilst a peak at δ 4.36 ppm would be assigned to α-glycerophosphate based on spiking experiments in the untreated extracts of the Cambisol (F) and the Gleysol, hypobromite oxidation revealed the presence of the 2-equatorial-4-axial C2,5 peak of *neo*-IP_6 at δ 4.37 ppm, and also an unidentified peak at δ 4.36 ppm in the Cambisol (F). Therefore, the assignment and concentration of α-glycerophosphate may be unreliable in some soils of previous studies.

For the first time, we identified lower-order IP (IP_5 and IP_4) in soil extracts using solution ^{31}P NMR spectroscopy. Smith and Clark (1951) were the first to suggest the presence of IP_5 in soil extracts using anion-exchange chromatography, which was later confirmed (Anderson, 1955; Cosgrove, 1963; McKercher and Anderson, 1968b). Halstead and Anderson (1970) reported the presence of all four stereoisomers (*myo*, *scyllo*, *neo* and *chiro*) in the lower ester fractions (IP_2-IP_4) as well as in the higher ester fractions (IP_5, IP_6) isolated from soil, with the *myo* stereoisomer being the main form in all fractions. In the current study, all four stereoisomers of IP_5 could be detected in the hypobromite oxidised soil extracts, of which the *myo* and *scyllo* stereoisomers were the most abundant. The

relative abundances of IP_5 stereoisomers are consistent with the findings of Irving and Cosgrove (1982) using gas-liquid chromatography on the combined IP_6 + IP_5 fraction. The detection of all four stereoisomers of IP_5 in NMR spectra provides direct spectroscopic evidence for their existence in soil extracts.

In addition to the four stereoisomers of IP_5, we were able to identify the presence of two isomers of *myo*-IP_5 in the Cambisol (F) and Gleysol, i.e. *myo*-(1,2,4,5,6)-IP_5 and *myo*-(1,3,4,5,6)-IP_5. These two isomers have not yet been detected in soil extracts. A distinction of different *myo*-IP_5 isomers was not reported in earlier studies using chromatographic separation. In non-soil extracts, *myo*-(1,2,4,5,6)-IP_5 was detected by Doolette and Smernik (2016) in grapevine canes and *myo*-(1,3,4,5,6)-IP_5 as the thermal decomposition product of a phytate standard (Doolette and Smernik, 2018). It is possible that an abiotic transformation of *myo*-IP_6 to *myo*-(1,3,4,5,6)-IP_5 occurs, which could then be adsorbed by soil constituents. Stephens and Irvine (1990) reported *myo*-(1,3,4,5,6)-IP_5 as an intermediate in the synthesis of IP_6 from *myo*-IP in the cellular slime mould *Dictyostelium*. Therefore, *myo*-(1,3,4,5,6)-IP_5 could have been biologically added to the soil. Furthermore, *myo*-(1,3,4,5,6)-IP_5 was present in different animal feeds and manures (Sun and Jaisi, 2018). Sun et al. (2017) reported *myo*-(1,3,4,5,6)-IP_5 and *myo*-(1,2,4,5,6)-IP_5 as intermediates in the minor pathways and major pathways, respectively, of *Aspergillus niger* phytase and acid phosphatase (potato) phytate degradation. The presence of *myo*-(1,2,3,4,6)-IP_5 could not be confirmed as NMR analyses on the compound itself exhibited a broad NMR signal (Figure SI-6). This is because in solutions with a pH of 9.5 or above, the 1-axial-5-equatorial and 5-axial-1-equatorial forms of *myo*-(1,2,3,4,6)-IP_5 are in a dynamic equilibrium, which can cause broadening (Volkmann et al., 2002). According to Turner and Richardson (2004) and Chung et al. (1999), the two identified *scyllo*-IP_4 peaks (signal pattern 2:2) can be attributed to the *scyllo*-(1,2,3,4)-IP_4 isomer. However, the two peaks of *scyllo*-IP_4 overlapped in the Cambisol (F) and Gleysol with the peak at the furthest upfield chemical shift of *myo*-(1,2,4,5,6)-IP_5 at δ 3.25 ppm, and with the peak at the furthest downfield chemical shift of *myo*-(1,3,4,5,6)-IP_5 at δ 4.12 ppm. Turner and Richardson (2004) reported NMR-signals for two other *scyllo*-IP_4 isomers, which could not be tested for in this study due to the lack of available standards.

Whilst on average 48 % of the sharp peaks in the phosphomonoester region of soil extracts following hypobromite oxidation could be attributed to IP_6, IP_5 and *scyllo*-IP_4, the identities of many sharp peaks remain unknown. An unidentified peak at δ 4.33 ppm is present in all soil samples except in the Ferralsol, with concentrations of up to 10 mg P/kg_{soil} (Cambisol

(F)). Other unidentified peaks at δ 3.49, 3.86, 4.20 and 3.91 ppm were detected in all soils, with concentrations ranging from 1 to 2 mg P/kg_{soil}. Interestingly, two peaks upfield of *scyllo*-IP_6 became more prominent (at δ 3.08 and 3.05 ppm) following hypobromite oxidation, which was particularly the case in the Vertisol soil. The diversity of organic P species in the Vertisol soil appears to be much greater than previously reported (McLaren et al., 2014). We hypothesise that many of these unidentified peaks arise from other isomers of *myo*- and *scyllo*-IP_5, based on the higher abundance of their IP_6 counterparts.

The ratio of IP_6 to lower-order IP varied across soils, which ranged in decreasing order: Gleysol ≫ Cambisol (F) > Vertisol > Ferralsol. McKercher and Anderson (1968a) found a higher ratio of IP_6 to IP_5 in some Scottish soils (ratio 1.8 to 4.6) compared to some Canadian soils (0.9 to 2.4). The authors attributed this difference to the greater stabilisation of IP_6 relative to lower esters in the Scottish soils, possibly due to climatic reasons or effects of different soil properties. In a subsequent study, McKercher and Anderson (1968b) observed increased IP contents with increasing total organic P content. Studies of organic P speciation along chronosequences found that *myo*-IP_6 concentrations declined in older soils (McDowell et al., 2007; Turner et al., 2007a; Turner et al., 2014). Similarly, Baker (1976) found that the IP_6 + IP_5 concentrations in the Franz Josef chronosequence increased until reaching 1000 years, followed by a rapid decline. In our soil samples, the highest IP_6-to-IP_5 ratio was found in the soil with the highest SOM content, suggesting a possible stabilisation of IP_6 due to association with SOM (Borie et al., 1989; Makarov et al., 1997). In contrast, the Ferralsol sample containing high amounts of Fe and Al showed the smallest IP_6 to IP_5 ratio, even though IP_6 is known to strongly adsorb to sesquioxides (Anderson and Arlidge, 1962; Anderson et al., 1974). However, the production, input and mineralisation rates of IP_6 and IP_5 are not known for our soil samples. Further research is needed to understand the mechanisms controlling the flux of lower-order IP in soil.

In the Ferralsol and the Cambisol (F), there was an overall decrease in the concentration of IP_6 and IP_5 following hypobromite oxidation compared to the untreated extracts. Since the main cause of resistance of IP to hypobromite oxidation is that of steric hindrance, which generally decreases with a decreasing phosphorylation state and conformation of the phosphate groups (axial vs. equatorial), we assume that low recoveries of added *myo*-IP_6 are due to losses of precipitated P_{org} compounds during the precipitation and dissolution steps. This is supported by the decrease in the concentration of orthophosphate following hypobromite oxidation compared to untreated extracts.

Therefore, quantities of IP as reported in the current study should be considered as conservative.

Structural composition of phosphomonoesters in hypobromite oxidised soil extracts

The NMR spectra on hypobromite oxidised soil extracts revealed the presence of sharp and broad signals in the phosphomonoester region. Transverse relaxation experiments revealed a rapid decay of the broad signal compared to the sharp peaks of IP_6, which support the hypothesis that the compounds causing the broad signal arise from P compounds other than IP. These findings are consistent with those of McLaren et al. (2019), who probed the structural composition of phosphomonoesters in untreated soil extracts. Overall, measured T_2 times in the current study on hypobromite oxidised extracts were markedly longer compared to those of untreated extracts reported in McLaren et al. (2019). This could be due to removal of other organic compounds by hypobromite oxidation in the matrix and therefore a decrease in the viscosity of the sample. This would result in an overall faster tumbling of the molecules and hence an increased T_2 relaxation time. As reported by McLaren et al. (2019), calculations of the broad signal's line width based on the T_2 times were considerably lower compared to those of the standard deconvolution fitting (SDF). When applying the same calculations to our samples, the line width of the broad signal at half-height is on average 5.2 Hz based on the T_2 times. In contrast, the line widths acquired from the SDF average 256.1 Hz. McLaren et al. (2019) suggested that the broad signal is itself comprised of more than one compound. Our results are consistent with this view and therefore it is likely that the main cause of the broad signal is a diversity of P molecules of differing chemical environments within this region, rather than the slow tumbling of just one macromolecule. Nebbioso and Piccolo (2011) reported that high molecular weight material of organic matter in soil is an association of smaller organic molecules. We suggest that these associations would still cause a broad signal in the phosphomonoester region of soil extracts due to the reduced tumbling of the large molecular structure with resulting short T_2 time and enhanced line width. Furthermore, the incorporation into a larger association could be a reason that some organic molecules containing P are protected from hypobromite oxidation.

Since a portion of the broad signal is resistant to hypobromite oxidation, this suggests the organic P is complex and in the form of polymeric structures. The chemical resistance of the broad signal to hypobromite oxidation may also indicate a high stability in soil (Jarosch

et al., 2015). Annaheim et al. (2015) found that concentrations of the broad signal remained unchanged across three different organic fertiliser strategies after 62 years of cropping. In contrast, the organic P compounds annually added with the fertilisers were completely transformed or lost in the slightly acidic topsoil of the field trial. The large proportion of the broad signal in the total organic P pool demonstrates its importance in the soil P cycle.

Unexpectedly, the transverse relaxation times of orthophosphate were shorter than those of the broad signal. This was similarly the case in an untreated NaOH-EDTA extract of a forest soil with the same origin as the Cambisol (F) as reported in McLaren et al. (2019). The authors hypothesised that this might be due to the sample matrix (i.e. high concentration of metals and organic matter). Whilst these factors are likely to affect T_2 times, they do not appear to be the main cause as the hypobromite oxidised extracts in the current study contained low concentrations of organic matter and metals as a consequence of the isolation procedure. The fast decay of orthophosphate was found across all four soil extracts with a diverse array of organic P concentrations and compositions of organic P in the phosphomonoester region. Therefore, another possible explanation could be a matrix effect or an association with large organic P compounds causing the broad signal (McLaren et al., 2019). It is known that dynamic intramolecular processes, such as ring inversion, and intermolecular processes, such as binding of small-molecule ligands to macromolecules, can cause a broadening or a doubling of resonances (Claridge, 2016b). When the smaller molecule is bound to the larger molecule, it experiences slower tumbling in the solution and hence a shorter T_2 time. It is possible that a chemical exchange of the orthophosphate with a compound in the matrix or an organic P molecule could result in the short T_2 time of the orthophosphate peak. We carried out a T_2 experiment on a pure solution of monopotassium phosphate (described in the supporting information), in which the matrix effects should be considerably reduced compared to the soil extracts. We found that the T_2 time of orthophosphate (203 ms) in the pure solution was considerably longer than that reported in soil extracts following hypobromite oxidation.

Conclusion

Inositol phosphates are an important pool of organic P in soil, but information on the mechanisms controlling their flux in soil remains limited due in part to an inability to detect them using solution ^{31}P NMR spectroscopy. For the first time, we identified six different lower-order IP in the solution ^{31}P NMR spectra on soil extracts. Solution ^{31}P NMR spectra on hypobromite oxidised extracts revealed the presence of up to 70 sharp peaks, of which about 50 % could be identified. Our results indicate that a majority of the sharp peaks in solution ^{31}P NMR soil spectra were resistant to hypobromite oxidation, and therefore suggest the presence of diverse IP. Our study highlights the great diversity and abundance of IP in soils and therefore their importance in terrestrial P cycles. Further research on the mechanisms and processes involved in the cycling of this wide variety of IP in soil will have implications for our understanding of organic P turnover as well as plant availability, and possibly help improve fertiliser strategies in agricultural systems.

Furthermore, we provide new insight into the large pool of phosphomonoesters represented by the broad signal, of which a considerable portion was resistant to hypobromite oxidation. Further research is needed to understand the chemical composition of the broad signal, and the mechanisms controlling its flux in terrestrial ecosystems.

Supporting information

NMR observability

Measures of NMR observability were calculated for the untreated and the hypobromite oxidised extracts of all soils. Measures of NMR observability refer to the percentage of total P detected using NMR compared to that by ICP-OES. For the untreated soil extracts, measures of NMR observability ranged from 52 % (Gleysol) to 89 % (Ferralsol), with an average NMR observability of 66 %. For the hypobromite oxidised extracts, measures of NMR observability ranged from 58 % (Ferralsol) to 94 % (Cambisol (F)), with an average value of 83 %.

Inositol hexakisphosphate concentrations before and after hypobromite oxidation

Table SI-1. Concentrations of inositol hexakisphosphates in 0.25 M NaOH + 0.05 M EDTA soil extracts before and after hypobromite oxidation (HO). Quantification was based on spectral integration and deconvolution fitting of solution ^{31}P NMR spectra. The proportion of P (%) detected in hypobromite oxidised extracts compared to that in untreated extracts is provided in brackets.

Concentrations (mg P/kg$_{soil}$)		**Ferralsol**	**Vertisol**	**Cambisol (F)**	**Gleysol**
***myo*-IP$_6$**	before HO	4.4	0.6	46.2	90.4
	after HO	1.1	0.6	26.3	85.0
		(25)	(111)	(57)	(94)
***scyllo*-IP$_6$**	before HO	2.5	0.4	34.9	42.6
	after HO	0.4	0.3	15.6	41.1
		(14)	(68)	(45)	(97)
***neo*-IP$_6$ 4-eq-2-ax**	before HO	-	-	4.2	7.0
	after HO	-	-	1.4	8.8
				(33)	(126)
D-*chiro*-IP$_6$	before HO	-	-	7.2	6.7
	after HO	-	-	9.4	8.6
				(130)	(128)

Solution ^{31}P NMR spectra of spiked hypobromite oxidised soil extracts

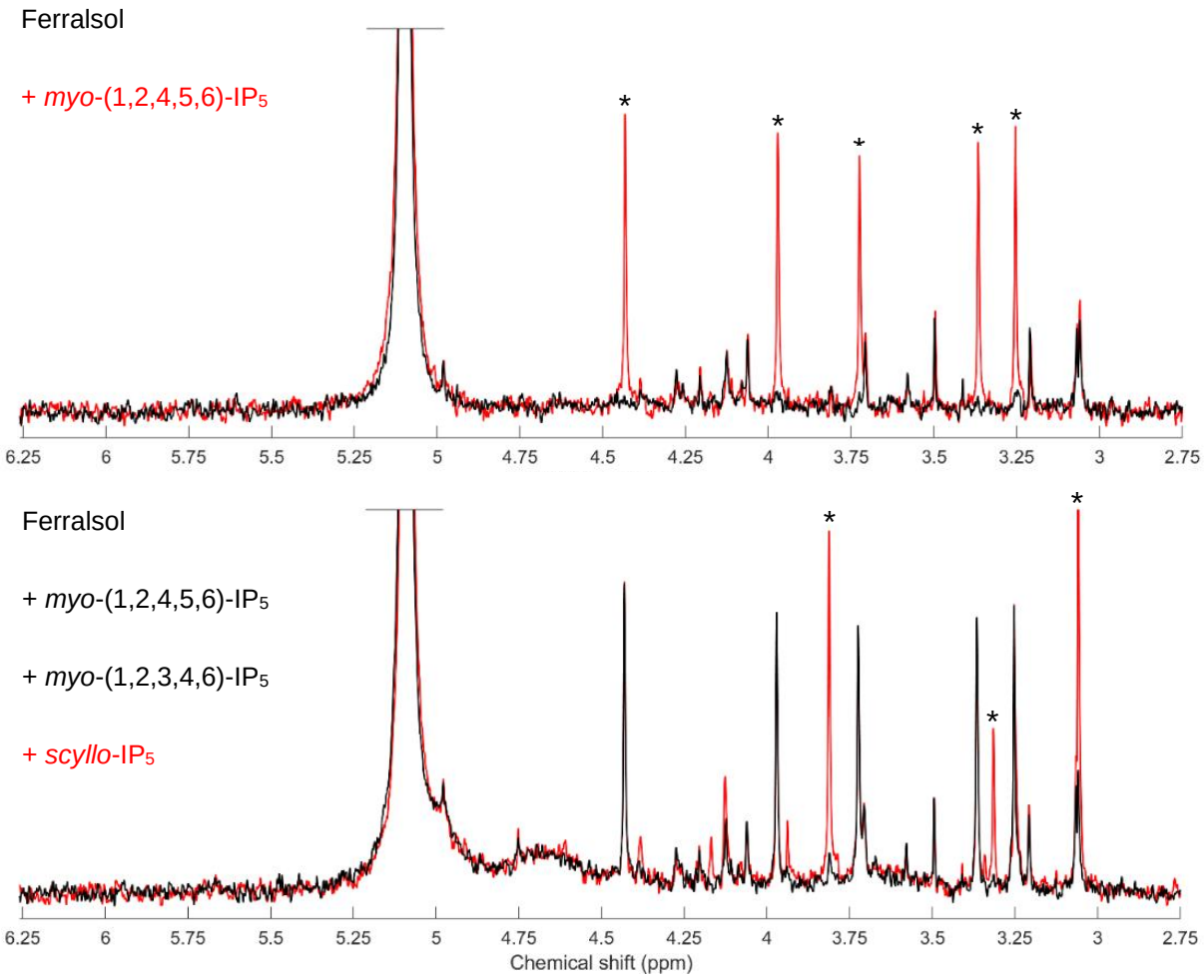


Figure SI-1. Solution ^{31}P NMR spectra of the orthophosphate and phosphomonoester region on Ferralsol extract following hypobromite oxidation (black trace), and also that following a spike with an IP standard (red trace). Peaks assigned to the IP standard marked with *.

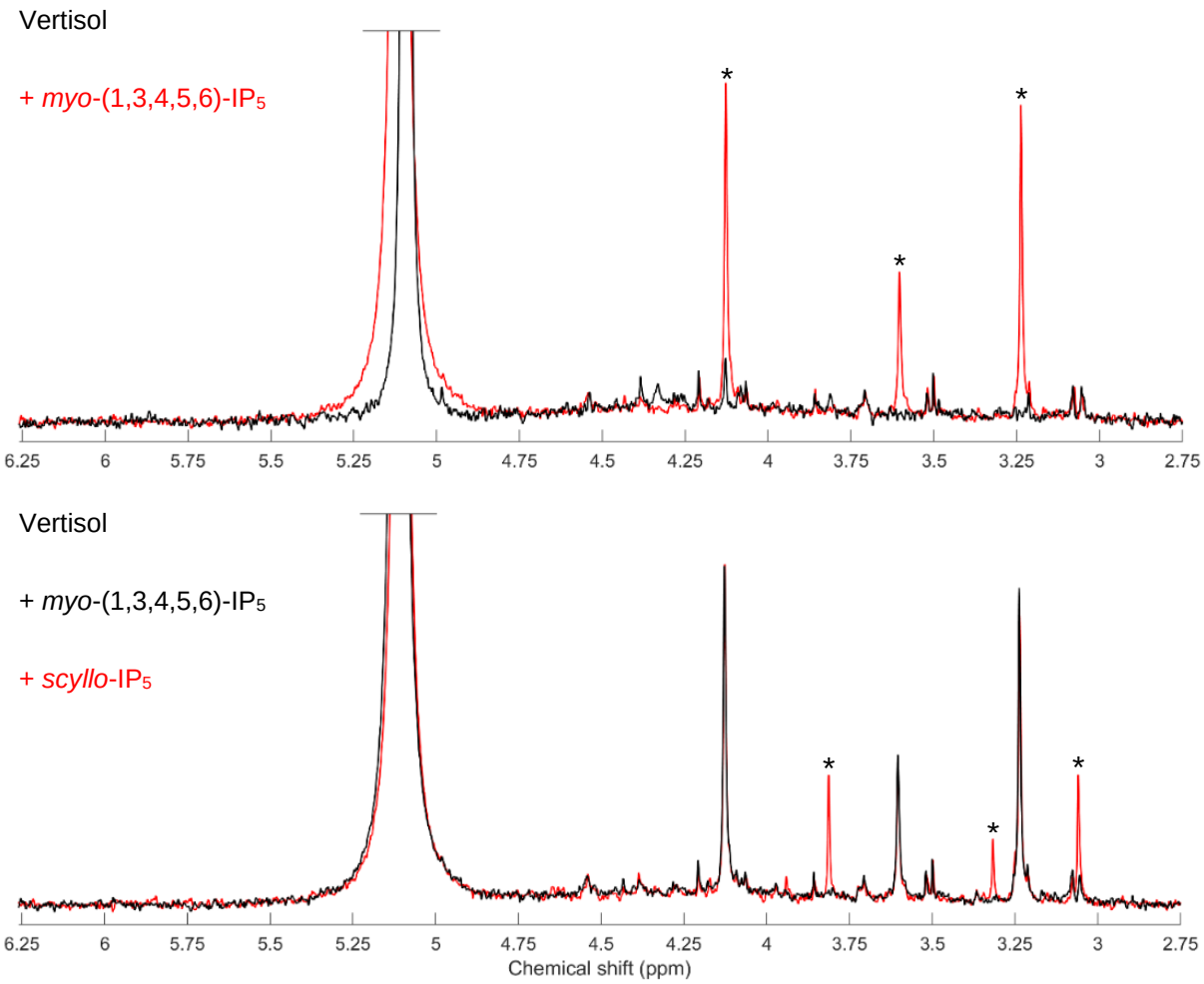


Figure SI-2. Solution ^{31}P NMR spectra of phosphomonoester region of hypobromite oxidised 0.25 M NaOH + 0.05 M EDTA Vertisol extract. Spiked spectrum with indicated standard in red. Peaks assigned to standard marked with *.

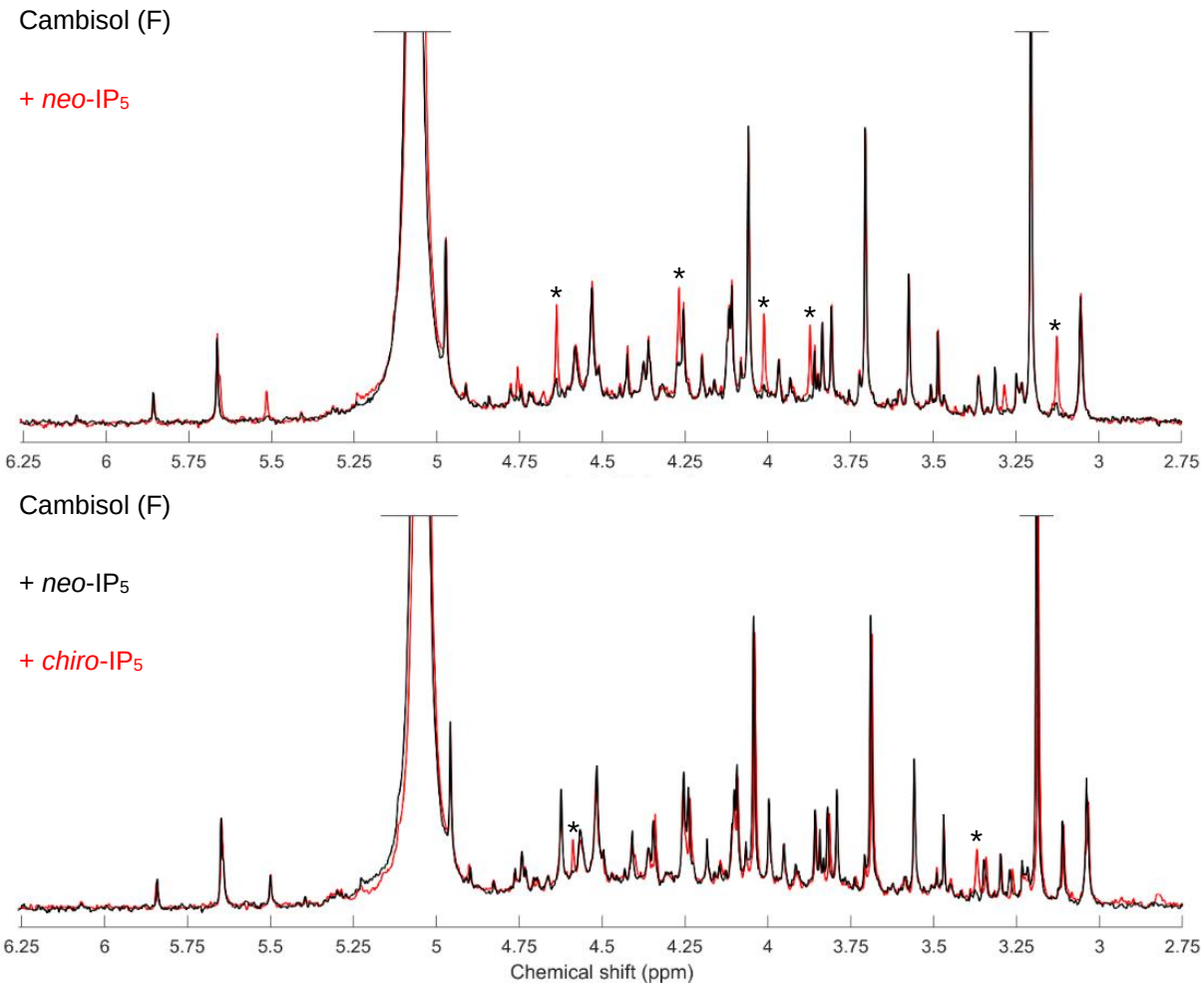


Figure SI-3. Solution ^{31}P NMR spectra of phosphomonoester region of hypobromite oxidised 0.25 M NaOH + 0.05 M EDTA Cambisol (F) extract. Spiked spectrum with indicated standard in red. Peaks assigned to standard marked with *.

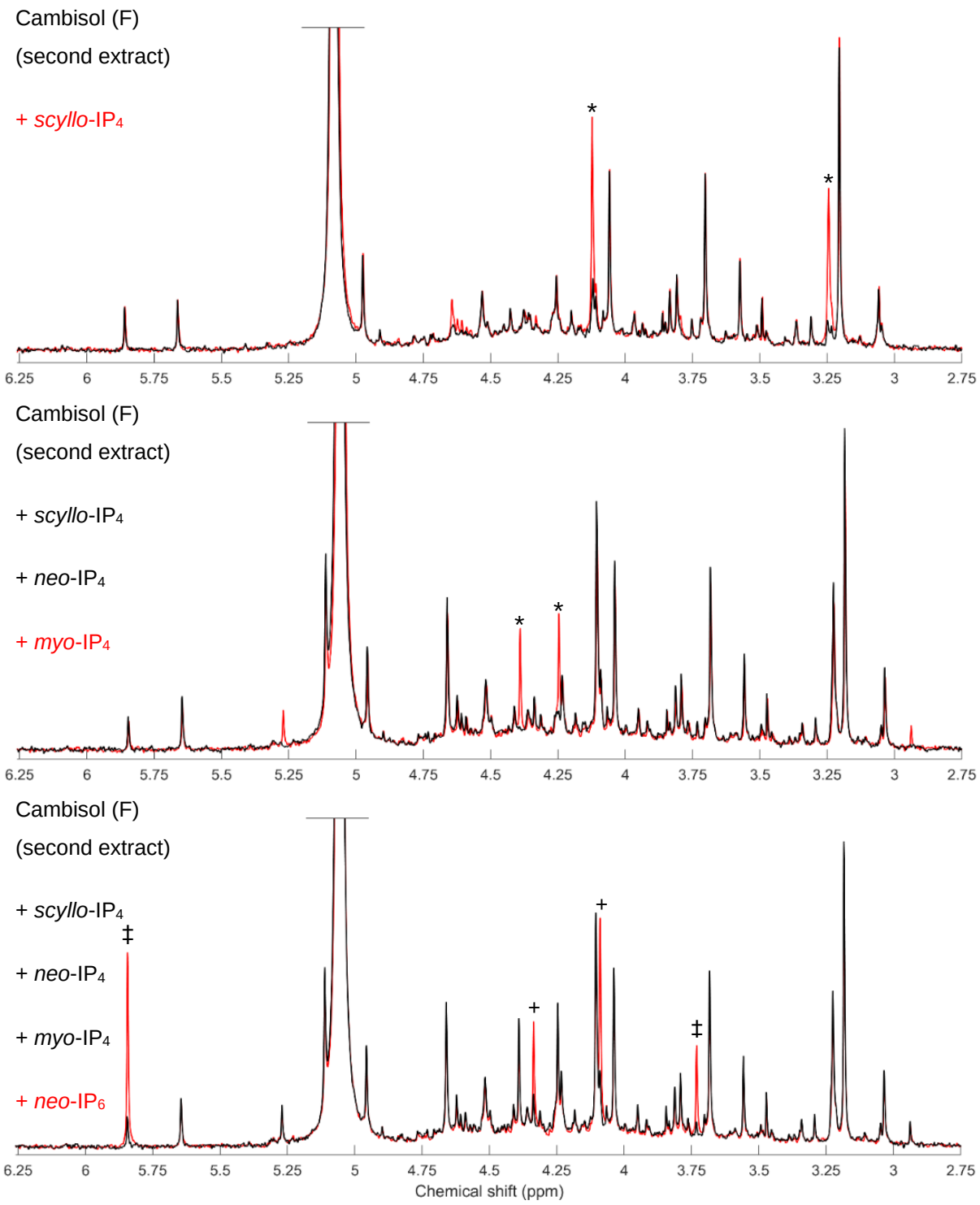


Figure SI-4. Solution ^{31}P NMR spectra of phosphomonoester region of hypobromite oxidised 0.25 M NaOH + 0.05 M EDTA Cambisol (F) extract. Spiked spectrum with indicated standard in red. Peaks assigned to 4-equatorial-2-axial conformation marked with ‡, peaks assigned to 2-equatorial-4-axial conformation marked with +.

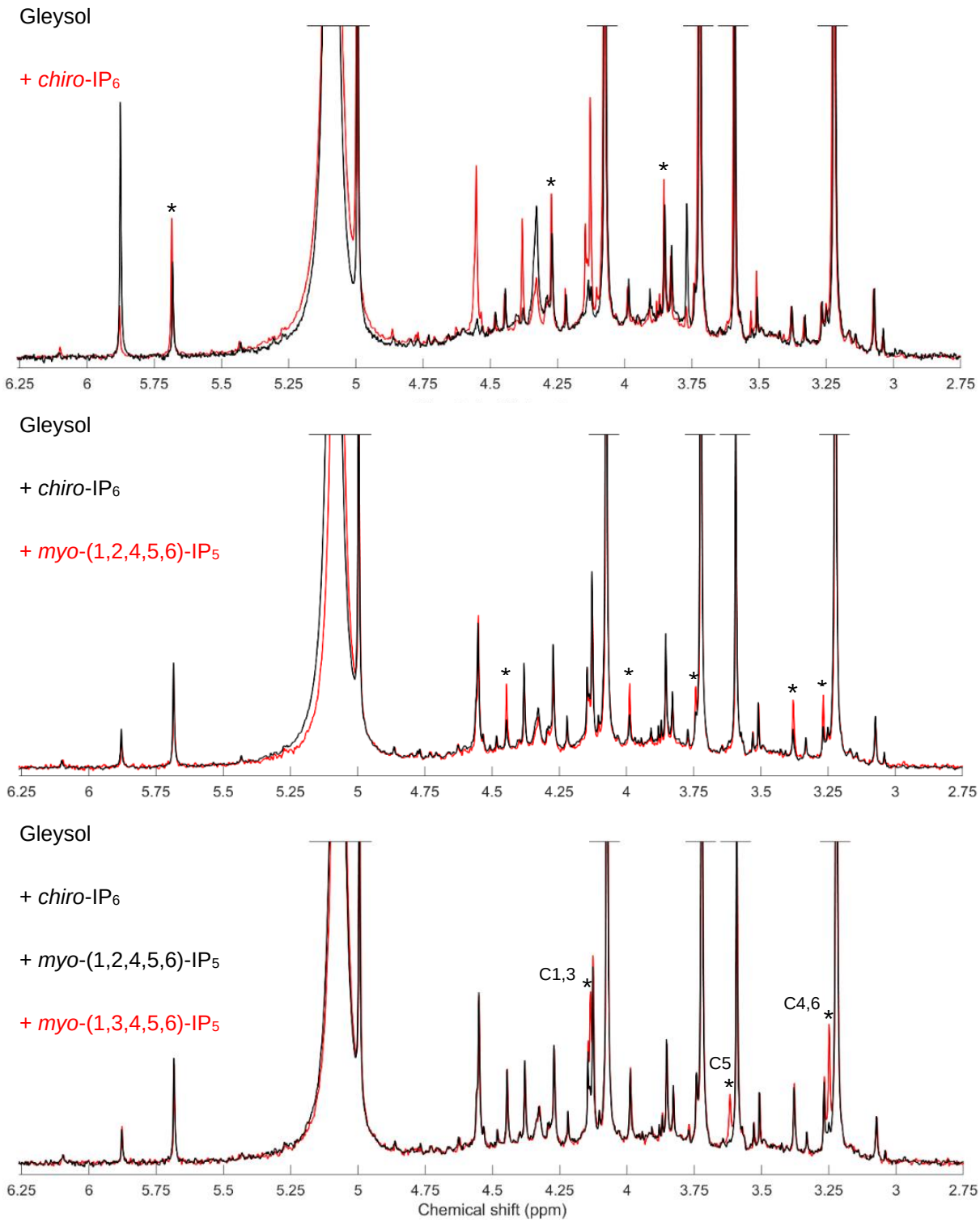


Figure SI-5. Solution ^{31}P NMR spectra of phosphomonoester region of hypobromite oxidised 0.25 M NaOH + 0.05 M EDTA Gleysol extract. Spiked spectrum with indicated standard in red. Peaks assigned to standard marked with *. For *myo*-(1,3,4,5,6)-IP_5, the respective phosphorylated carbon nuclei of the inositol have been marked based on the ^{31}P NMR spectrum prediction of the program Mnova 11.0.4 (©Mestrelab Research).

Transverse relaxation time of an orthophosphate solution

The analysis of a 0.25 M NaOH + 0.05 M EDTA solution containing 910 mg KH_2PO_4/L resulted in a single orthophosphate peak in the NMR spectrum (δ 5.09 ppm) with a line width at peak half-height of 0.56 Hz. Transverse relaxation experiments were carried out (similar to that previously described) on the solution, which resulted in a T_2 time of 203 ms for orthophosphate.

NMR analysis of purchased myo-(1,2,3,4,6)-IP_5 standard

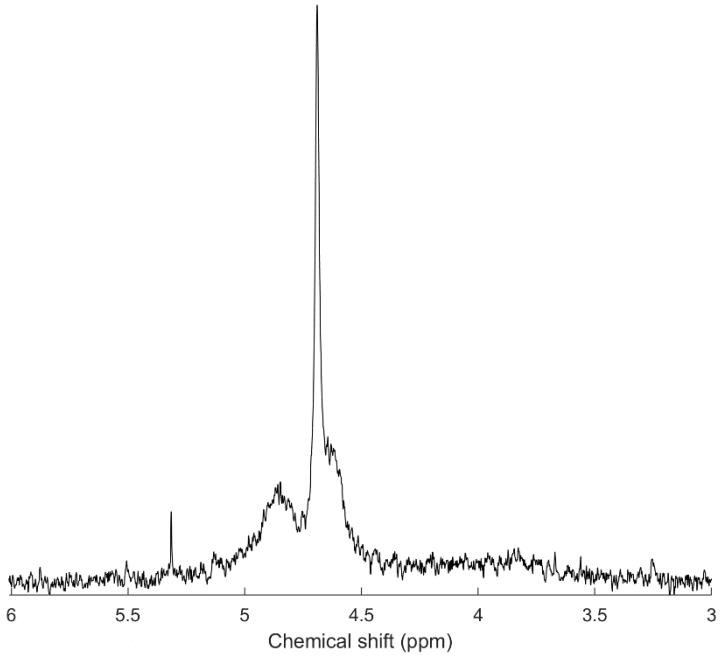


Figure SI-6. Solution ^{31}P NMR spectrum of phosphomonoester region of purchased *myo*-(1,2,3,4,6)-IP_5 standard dissolved in 0.25 M NaOH + 0.05 M EDTA.

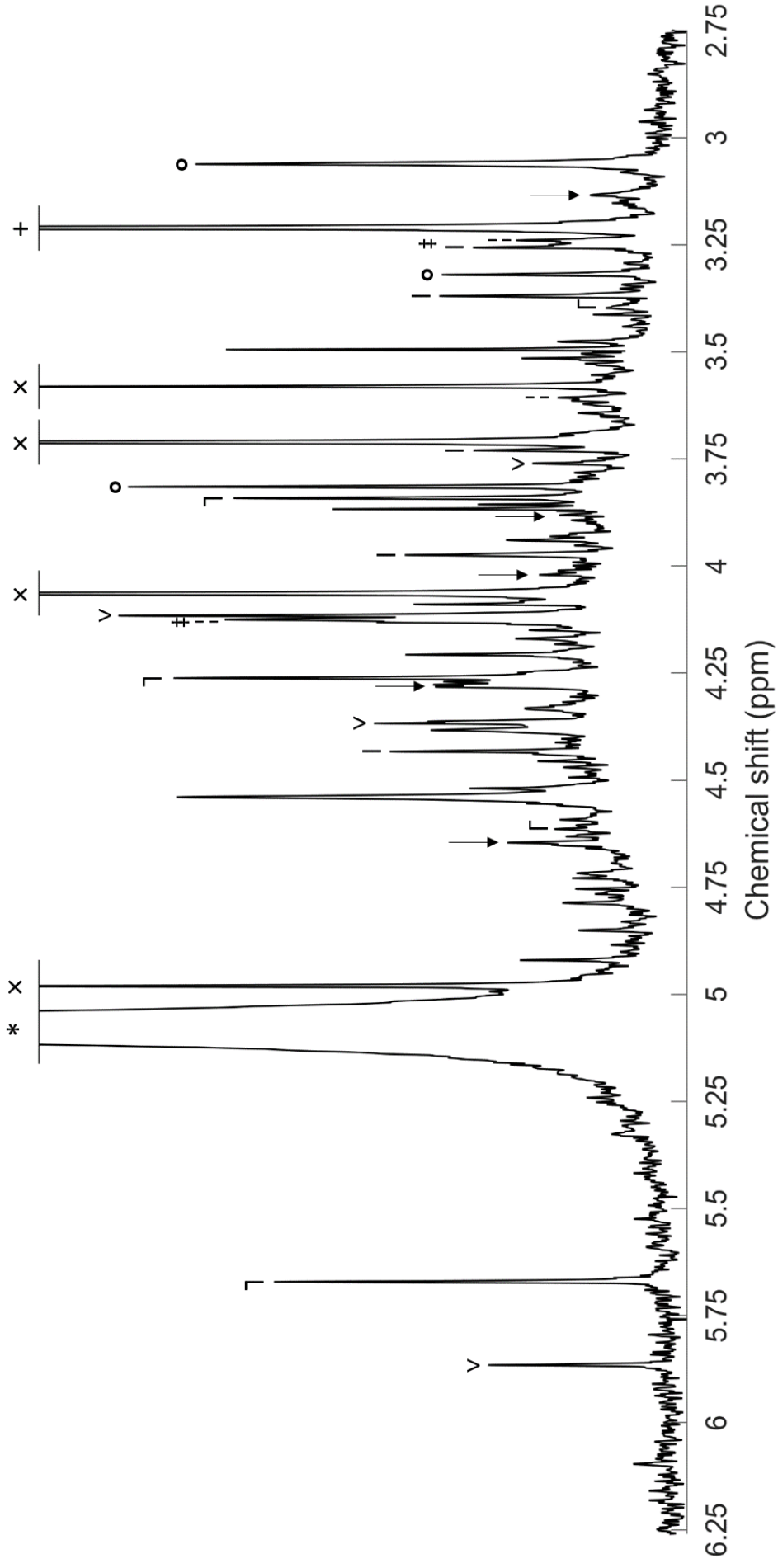


Figure SI-7. Solution ^{31}P NMR spectra of phosphomonoester region of hypobromite oxidised 0.25 M NaOH + 0.05 M EDTA Cambisol (F) extract. All identified peaks are marked: orthophosphate (*), *myo*-IP_6 (×), *scyllo*-IP_6 (+), *neo*-IP_6 (v), *chiro*-IP_6 (Π), *myo*-(1,2,4,5,6)-IP_5 (I), *myo*-(1,3,4,5,6)-IP_5 (¡), *scyllo*-IP_5 (°), *neo*-IP_5 (↓), *chiro*-IP_5 (Γ), *scyllo*-(1,2,3,4)-IP_4 (‡). The chemical shifts in ppm of all identified peaks are listed in Table 5.

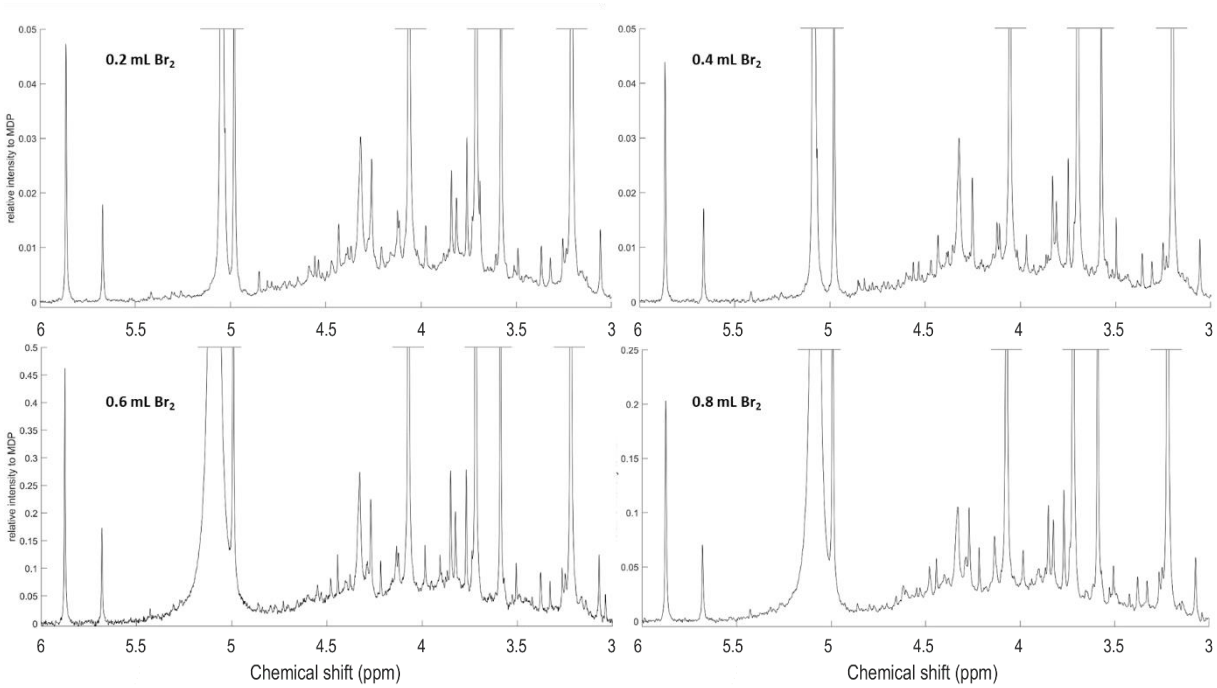


Figure SI-8. Solution ^{31}P nuclear magnetic resonance (NMR) spectra (500 MHz) of the orthophosphate and phosphomonoester region of hypobromite oxidised 0.25 M NaOH + 0.05 M EDTA Gleysol extract, using 0.2 mL, 0.4 mL, 0.6 mL and 0.8 mL Br_2 in the hypobromite oxidation procedure. Signal intensities were normalised to the MDP peak (intensity of 1 on y-axes).

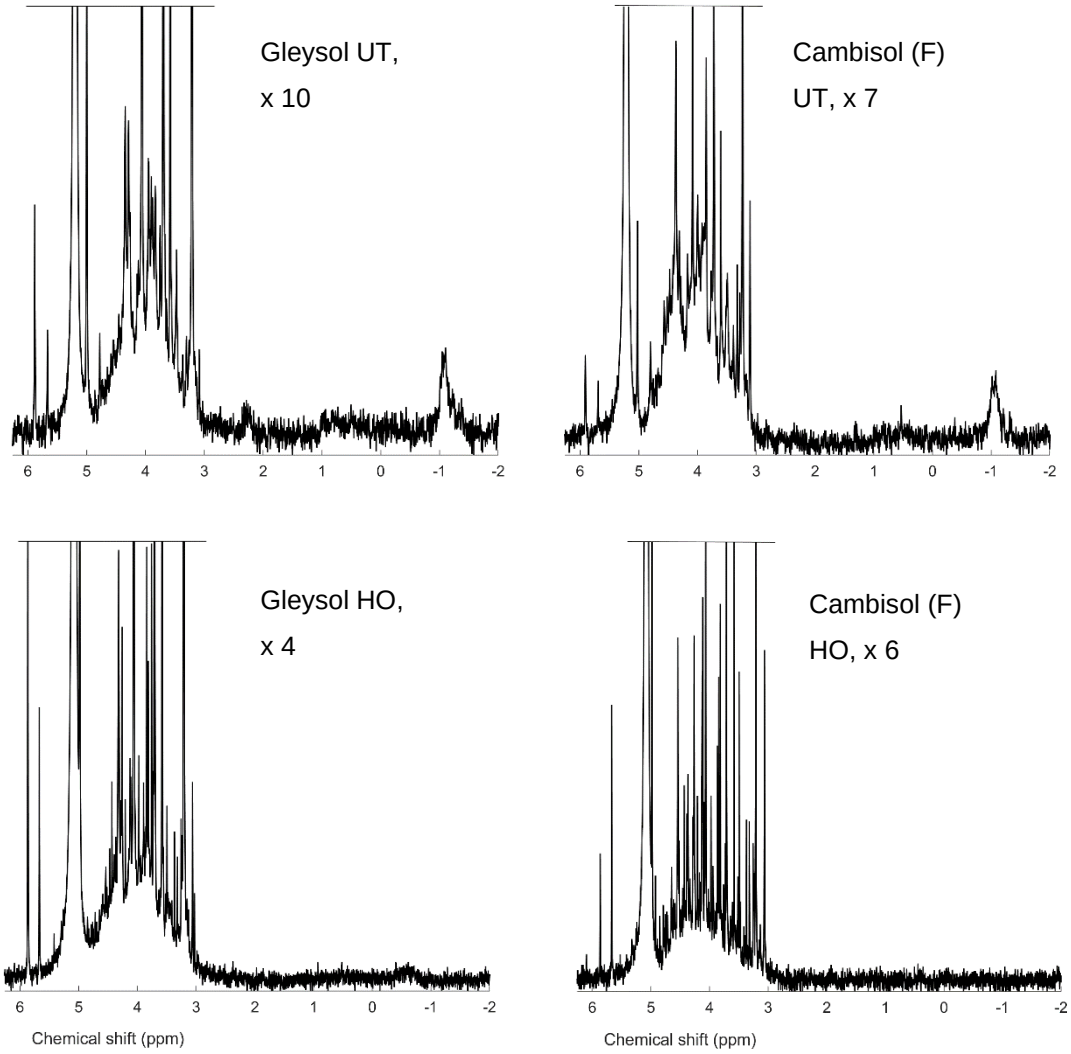


Figure SI-9. Solution ^{31}P nuclear magnetic resonance (NMR) spectra (500 MHz) of the orthophosphate, phosphomonoester and phosphodiester region on untreated (UT, on top) and hypobromite oxidised (HO, below) 0.25 M NaOH + 0.05 M EDTA soil extracts of the Gleysol (left) and Cambisol (F) (right). Signal intensities were normalised to the MDP peak intensity. The vertical axes were increased for improved visibility of spectral features, as indicated by a factor.

Chapter 3

The molecular size continuum of soil organic phosphorus and its chemical associations

Under review as:

Reusser, J.E., Tamburini, F., Neal, A.L., Verel, R., Frossard, E., McLaren, T.I. The molecular size continuum of soil organic phosphorus and its chemical associations. Geoderma, The Global Journal of Soil Science, Elsevier.

Abstract

The chemical nature of most organic P (P_{org}) in soil remains 'unresolved' but can be accounted for by a broad signal in the phosphomonoester region of solution ^{31}P nuclear magnetic resonance (NMR) spectra. The molecular size range of this broad signal and its molecular structure remain unclear. The aim of this study was to elucidate the chemical nature of P_{org} with increasing molecular size in soil extracts combining size exclusion chromatography (SEC) with solution ^{31}P NMR spectroscopy. Gel-filtration SEC was carried out on NaOH-EDTA extracts of four soils (range 238-1135 mg P_{org}/kg_{soil}) to collect fractions with molecular sizes of <5, 5-10, 10-20, 20-50, 50-70, and >70 kDa. These were then analysed by NMR spectroscopy.

Organic P was detected across the entire molecular size continuum from <5 to >70 kDa. Concentrations of P_{org} in the >10kDa fraction ranged from 107 to 427 mg P/kg_{soil} and exhibited on average three to four broad signals in the phosphomonoester region of NMR spectra. These broad signals were most prominent in the 10-20 and 20-50 kDa fractions, accounting for on average 77 % and 74 % of total phosphomonoesters, respectively. Our study demonstrates that the broad signal is present across a large continuum of molecular sizes and comprises multiple components of varying NMR peak line width (20 to 250 Hz).

The stereoisomers *myo*- and *scyllo*-inositol hexakisphosphates (IP_6) were also present across multiple molecular size ranges but were predominant in the 5-10 kDa fraction. The proportion of IP associated with large molecular size fractions >10 kDa was on average 23 % (SD=39 %) of total IP across all soils. These findings suggest that stabilisation of IP in soil includes processes associated with the organic phase.

Introduction

Phosphorus (P) is an innate constituent of soil organic matter. Pools of organic P (P_{org}) typically comprise between 20 % and 80 % of total P (P_{tot}) in most soils (Anderson, 1980; Harrison, 1987). Inositol phosphates (IP) are an abundant and relatively stable pool of P_{org} in soil (Turner et al., 2002): the *myo* stereoisomer of IP_6 is typically the most abundant form and comprises on average one third of the soil P_{org} (McLaren et al., 2020). *myo*-IP_6 accumulates in soil, largely due to its high sorption affinity to soil components, e.g. iron (Fe) and aluminium (Al) oxides, calcium (Ca) carbonates, clay surfaces and organic matter (Celi and Barberis, 2007). Such sorption as well as complexation with metal cations (Martin and Evans, 1987) limit the accessibility of IP to enzymatic hydrolysis (Celi and Barberis, 2005, 2007). Other identifiable pools of P_{org} include phosphonates (e.g. 2-aminoethylphosphonate), nucleic acids (e.g. DNA and RNA) and phospholipids. These generally account for less than 10 % of soil P_{tot} (Anderson and Malcolm, 1974; Dalal, 1977) due to their rapid microbial degradation (Auten, 1923; Harrison, 1982). The chemical nature of the majority of soil P_{org} remains unresolved and consequently the mechanisms controlling its flux in the soil-plant system are still poorly understood (McLaren et al., 2020).

Several studies have shown that large proportions of P_{org} in soil extracts occur in large molecular weight material. Moyer and Thomas (1970) reported that 31 % of extractable P_{org} from a Luvisol had a molecular weight range of 1-50 kDa, whereas 36 % of the extractable P_{org} had a molecular weight range greater than 50 kDa. Steward and Tate (1971) reported that the majority of extractable P_{org} from several alkaline soils had a molecular weight range greater than 30 kDa. Large pools of P_{org} present in high molecular weight material have been reported in several review papers (Anderson, 1980; Kögel-Knabner and Rumpel, 2018; McLaren et al., 2020). However, there is limited information on the chemical nature of P_{org} with increasing molecular weight. Some studies provide evidence supporting the existence of IP in high molecular weight material (Moyer and Thomas, 1970; Hong and Yamane, 1981; Borie et al., 1989), despite the molecular size of IP being about 660 Da or less. These studies often suggest that the presence of IP in high molecular weight material could be due to a polymeric form of IP, or IP bound to (or within) other organic compounds. In this latter case, it is proposed that IP might be complexed to organic compounds via polyvalent metals, such as Fe and Al (Lévesque and Schnitzer, 1967; Lévesque, 1969; Veinot and Thomas, 1972).

In general, the largest pool of soil P_{org} is observed as an underlying broad signal in the phosphomonoester region of solution ^{31}P NMR spectra on soil extracts (Doolette et al.,

2011a; Jarosch et al., 2015; McLaren et al., 2015b). Compounds responsible for the broad signal typically comprise 40-70 % of total P_{org} and are of an apparent high molecular weight. Jarosch et al. (2015) used Sephadex G-25 gel to separate NaOH-EDTA extracts according to molecular size from several soils. The authors found that P_{org} was detected in both low (<5 kDa) and high (>5 kDa) size fractions. Pools of P_{org} in the latter fraction were largely stable to enzyme hydrolysis and correlated with pools of phosphomonoesters exhibiting a broad NMR signal. At the same time, McLaren et al. (2015b) used ultrafiltration to separate NaOH-EDTA extracts from diverse soils into <10 kDa and >10 kDa fractions. These authors reported that about 32 % of P_{org} was measured in the larger fraction, and was dominated by a broad NMR signal in the phosphomonoester region. However, the authors also reported the presence of a broad signal in the phosphomonoester region of NMR spectra of <10 kDa fractions.

The structural composition of the broad signal in the phosphomonoester region of solution ^{31}P NMR spectra on soil extracts was later investigated by McLaren et al. (2019) based on transverse relaxation (T_2) experiments. The authors found that T_2 times associated with the broad signal were considerably shorter than those of sharp signals (e.g. *myo*-IP_6) in the phosphomonoester region, suggestive of the former having a large molecular size (Bloembergen et al., 1948; Keeler, 2010). This was similarly the case for T_2 experiments carried out on soil extracts following hypobromite oxidation (Reusser et al. (2020b). McLaren et al. (2019) compared the line width of the broad signal at half peak intensity based on spectral deconvolution with that derived from T_2 experiments. The authors reported that the line width at half peak intensity of the broad signal based on the former approach was significantly larger than that based on the latter approach, suggesting the broad signal itself is derived from multiple components. The number of components and their chemical composition remains unknown.

Organic compounds containing phosphate in soil are known for their capacity to form complexes with metal cations, such as Al and Fe (Parfitt, 1979; Stevenson, 1994; Turner et al., 2002; Vincent et al., 2012). Such complexes of soil organic matter (SOM) and P_{org} (i.e. IP) with metals contribute to the stabilisation and possible accumulation of compounds in soil (Stevenson, 1994; Celi and Barberis, 2005). Gerke (2010) reported that orthophosphate is not directly bound to an organic moiety of SOM, as would be the case for P_{org} compounds, but indirectly associated with the SOM via metal bridges. The authors identified Al^{3+} and Fe^{3+} as the predominant metal linkages in humic-metal-P complexes.

The aim of this study was to provide new insight on the chemical nature of P_{org} at increasing molecular size in soil extracts, combining size exclusion chromatography (SEC) and solution ^{31}P NMR spectroscopy. We hypothesised that: 1) phosphomonoesters in large molecular size material are dominated by several broad signals in the phosphomonoester region, but also contain several sharp signals due to IP; 2) phosphomonoesters in small molecular size material are dominated by sharp signals primarily from IP; 3) the composition, and hence the peak shape and intensity of the broad signal, will change among molecular size fractions; and 4) concentrations of Fe and Al in large molecular size material will correlate with that of IP, indicating the importance of metals in bridging IP with large molecular size compounds.

Materials and methods

Soil collection and preparation

The A horizon of calcareous Cambisol arable (A) and pasture (P) soils from Switzerland, a Cambisol forest (F) soil from Germany, and a Gleysol pasture soil from Switzerland (WRB, 2014) were investigated in this study. Detailed information of sampling sites, cropping systems and properties of the four soils can be found in Reusser et al. (2020a). Additional information on oxalate extractable concentrations of total Fe, Al, and P are reported in Table 1. For each, 1 g of soil was extracted with 100 mL of 0.2 M acid oxalate solution (Tamm-reagent at pH 3) according to the modified method of Schwertmann (1964), as described in Pansu and Gautheyrou (2007). Tubes were shaken end-over-end for 4 h at 11 revolutions per min in the absence of light at 20 °C, centrifuged at 6480 *g* for 10 min, and the supernatant analysed for total Al, Fe and P using inductively coupled plasma-optical emission spectroscopy (ICP-OES).

Table 1. Oxalate extractable aluminium, iron and phosphorus of the soil samples reported in this study.

Parameter	Unit	Cambisol (P)	Cambisol (F)	Gleysol	Cambisol (A)
Al_{ox}	g/kg_{soil}	1.5	7.2	3.2	0.8
Fe_{ox}	g/kg_{soil}	3.4	48.8	3.5	2.0
P_{ox}	mg/kg_{soil}	643.4	1795.0	989.3	337.0

Concentrations of total phosphorus in soil

Total concentrations of soil P were obtained by X-ray fluorescence (XRF) spectroscopy (SPECTRO XEPOS ED-XRF, AMETEK®) using 4.0 g of soil sample ground to powder mixed with 0.9 g of wax (CEREOX Licowax, FLUXANA®). The XRF instrument was calibrated using commercially available reference soils. Microwave acid digestion was used to determine the concentration of P_{tot} in soil based on the method of Fioroto et al. (2017). In brief, 200 mg of dried and ground soil was digested in 4 mL of a 14 M nitric acid solution using a turboWAVE® MRT microwave digestion system (MILESTONE Srl, Sorisole, Italy) at 250 °C for 35 min. The digestate was then diluted to 10 mL with deionised water and analysed for P using the malachite green method of Ohno and Zibilske (1991).

Extraction of soil organic phosphorus

Extraction of P_{org} from soil was based on the method of Cade-Menun et al. (2002). Briefly, 4 g of soil was extracted with 40 mL of 0.25 M NaOH + 0.05 M EDTA (NaOH-EDTA) solution and shaken for 16 h. The tube was then centrifuged at 4643 *g* for 10 min and the supernatant filtered through a Whatman no. 42 filter paper. Concentrations of P_{tot} in soil filtrates were determined by ICP-OES. Concentrations of molybdate reactive P (MRP) in soil filtrates were measured using the malachite green method. The difference in concentrations of P_{tot} and MRP is referred to as molybdate unreactive P (MUP), conventionally considered to be P_{org} (Bowman and Moir, 1993).

Initially, we carried out high-pressure liquid (HPL)-SEC with in-line inductively coupled plasma-mass spectrometry (ICP-MS) on lyophilised NaOH-EDTA soil extracts based on the method of Dell'Aquila et al. (2020) at Rothamsted Research (supporting information). Separation of samples was performed via SEC using an HPLC (PerkinElmer LC 200 Series HS, Seer Green, Bucks, UK) composed of an injector, a Flexar UV/VIS detector operating at 280 nm and a high-pressure peristaltic pump equipped with PEEK tubing (0.17 mm internal diameter), operated at a flow rate of 0.6 mL/min. Samples were analysed using an ICP-MS (PerkinElmer NexION 300XX, Seer Green, Bucks, UK) equipped with a glass Meinhard nebuliser, a quadrupole mass spectrometer and a collision cell. The column was a Superdex Peptide 10/300 GL (10 mm diameter x 300 mm length, GE Healthcare Bio-Sciences, Sweden). This revealed well resolved chromatograms with a continuum of P spanning the <1 to >20 kDa molecular size range (Figures SI-2 and SI-3). Concentrations of P in collections of discrete molecular size fractions were too low for P characterisation using NMR spectroscopy. An alternative preparative SEC method (i.e. high-resolution SEC) was used to fractionate NaOH-EDTA soil extracts according to molecular size which yielded sufficient concentrations of P in collected solution volumes for characterisation using NMR spectroscopy.

Size-exclusion chromatography

To prepare the soil extracts for size-fraction collection by high-resolution SEC, salts were first removed. This preparatory desalting procedure resulted in separation of soil filtrates into two size fractions with molecular sizes of <5 kDa and >5 kDa, based on a modified version of the method of Jarosch et al. (2015). Soils were extracted with NaOH-EDTA as described above. A 15 mL aliquot of soil filtrate was diluted with 35 mL of deionised water and filtered through a 0.22 µm syringe filter: the Cambisol (F) was filtered using a 0.45 µm

syringe filter due to blockage of 0.22 µm syringe filters. Desalting of the filtrate was carried out on an ÄKTAprime plus system (GE Healthcare), using a 4.5 cm (diameter) x 16 cm (length) column filled with Sephadex™ G-25 (GE Healthcare) with an exclusion limit of 5 kDa. The injected sample was eluted with deionised water at a flow rate of 1 mL/min and an upper-limit pressure of 0.2 MPa. Runtime was about 360 min and carried out under constant temperature (4 °C). The eluate was collected in 9 mL fractions. The cut-off at 5 kDa was defined based on the conductivity (peak indicating eluted orthophosphate ions), the absorbance (two major peaks) and measured concentrations of P_{inorg} and P_{org} in the collected fractions (Tamburini et al., 2018). The absorbance and conductivity curves with the determined cut-off are plotted in Figures SI-5 and SI-6 in the supporting information. Subsamples with molecular sizes greater than 5 kDa were combined (total volume of 126-135 mL) and lyophilised. Similarly, subsamples with molecular sizes less than 5 kDa were combined (total volume of 117-135 mL) and lyophilised.

High-resolution SEC was carried out to collect fractions of P_{org} across a wide molecular size distribution, which could then be analysed using solution ^{31}P NMR spectroscopy. The lyophilised material of the desalted soil extracts (>5 kDa fraction) was redissolved in 10 mL of deionised water. A 4 mL aliquot was then injected into the ÄKTAprime plus system of GE Healthcare equipped with a 1.6 cm (diameter) x 68.2 cm (length) column containing Superdex™ 75 prep grade resin beads (GE Healthcare) with a fractionation range of ~3 to 70 kDa (globular proteins). The sample was eluted with degassed 0.2 M ammonium nitrate (pH approx. 7.6) at a flow rate of 0.5 mL/min and a pump pressure of 0.3 MPa. Runtime for each sample was about 300 min and carried out under constant temperature (4 °C). As soon as the UV/VIS detector measured an increase in absorbance, 1 mL subsamples were collected.

The sensitivity curve of the column was calculated in the same way as described for the HPL-SEC analysis in the supporting information (Figure SI-1) but using a standard mixture of vitamin B_{12} (0.24 mg/mL), aprotinin (0.92 mg/mL), myoglobin (0.98 mg/L), albumin (1.44 mg/L) and blue dextran (1.16 mg/mL). Based on the sensitivity curve, and information provided from HPL-SEC analysis, elution times corresponding to molecular size cut-offs of 5 kDa, 10 kDa, 20 kDa, 50 kDa and 70 kDa were chosen. Subsamples were combined accordingly, resulting in molecular size fractions of <5 kDa (28-30 mL), 5-10 kDa (14 mL), 10-20 kDa (14 mL), 20-50 kDa (18 mL), 50-70 kDa (7 mL) and >70 kDa (14-18 mL). For the Cambisol (P) soil, the 50-70 kDa and >70 kDa fractions were combined due to a preparation error. A 450 µL aliquot was taken from each fraction in

order to determine the concentration of MRP using the malachite green method (Ohno and Zibilske, 1991), and concentrations of total P, Fe, Al, Mn, Ca and K were determined using ICP-OES. The remaining volume of collected fractions was lyophilised prior to NMR sample preparation, resulting in 90 to 440 mg of lyophilised material.

Sample preparation for solution ^{31}P NMR spectroscopy

Sample preparation for NMR analyses was carried out based on the methods of Spain et al. (2018) and Reusser et al. (2020a). Lyophilised material from each fraction was redissolved in 600 µL of NaOH-EDTA (a concentration factor of 13.55), except for the 10-20 kDa fraction of the Cambisol (F) which had a concentration factor of 18.55, when compared to the pre-lyophilisation volume. The solution was stored overnight to allow for complete hydrolysis of RNA and phospholipids (Vestergren et al., 2012), and then centrifuged at 10621 *g* for 15 min at 21 °C. A 500 µL aliquot of the supernatant was transferred to a 1.5 mL microcentrifuge tube and spiked with 25 µL sodium deuteroxide (NaOD) at 40 % (w/w) in D_2O (Sigma-Aldrich, product no. 372072) and 25 µL of a 0.03 M methylenediphosphonic acid standard (MDP) made up in D_2O (Sigma-Aldrich, product no. M9508). The prepared solution was then homogenised with a vortex mixer and transferred to a 5 mm NMR tube.

The fraction removed by the desalting process (<5 kDa), which was not subjected to high-resolution SEC, was prepared by redissolving 150 mg of lyophilised material with 750 µL of NaOH-EDTA. This was then stored and centrifuged as above. A 125 µL aliquot of the supernatant was diluted with 9.875 mL of deionised water and subsequently analysed for P and metals as described above. A 500 µL aliquot of the remaining supernatant was transferred to a 1.5 mL microcentrifuge tube. The supernatant was further prepared for NMR analysis as described above, except that 35 µL of NaOD was used instead of 25 µL to adjust the chemical shift of the orthophosphate peak. Furthermore, the Cambisol (F) and Gleysol samples were diluted with an additional 300 µL of NaOH-EDTA in the NMR tube due to considerable line broadening, which resolved the issue.

Solution ^{31}P NMR analysis and processing of spectra

Detailed information on the solution ^{31}P NMR analyses and processing of NMR spectra can be found in McLaren et al. (2019) and Reusser et al. (2020a). Spectra were acquired using a Bruker 500 MHz NMR spectrometer (^{31}P frequency of 202.5 MHz) equipped with a Prodigy CryoProbe™ (Bruker Corporation, Billerica, USA). Inverse gated proton

decoupling (90° pulse of 12 µs) was applied. Longitudinal relaxation (T_1) times were determined for each sample based on a preliminary inversion recovery experiment (Vold et al., 1968), which involved a recycle delay of 5 s and 24 scans per experiment. T_1 values were calculated from 10 individual experiments with increasing time periods ranging from 5 to 400 ms between the applied pulses (total analysis time was 56 min per sample). These experiments indicated recycle delays (i.e. five times T_1 of the slowest NMR signal) ranging from 5.5 to 33.0 s across all samples. Solution 1D ^{31}P NMR analyses were carried out with 4096 scans for each sample.

NMR spectra were processed (phase correction, baseline adjustment and integration) using the TopSpin® software environment (Bruker Corporation, Billerica, USA). Spectra are presented with a common line broadening of 0.6 Hz. For quantification of P species, spectral integration was performed. Relative proportions of peak regions compared to the net peak area of the MDP standard (δ 17.45 to 16.06 ppm), which has a known P concentration, were calculated for all samples (Turner, 2008; Doolette et al., 2011a). Peak regions included: 1) phosphonates (δ 20.14 to 16.41 ppm); 2) a combined orthophosphate and phosphomonoester region (δ 6.39 to 3.05 ppm); 3) phosphodiesters (δ 1.08 to -2.78 ppm), and 5) pyrophosphates (δ -4.12 to -5.48 ppm). The NMR observability for P_{tot} in NaOH-EDTA extracts was determined for all samples by comparing the proportion of P_{tot} as measured by NMR spectroscopy with that measured by ICP-OES (Dougherty et al., 2005; Doolette et al., 2011a).

Spectral deconvolution fitting procedure

Spectral deconvolution fitting (SDF) was applied to the combined orthophosphate and phosphomonoester region to separate overlapping peaks as described in Reusser et al. (2020a). Sharp peaks were fitted simultaneously with underlying broad signals using a non-linear optimisation algorithm in MATLAB® R2017a (The MathWorks, Inc.). Multiple broad signals were fitted based on a visual assessment of the peak distributions, to further minimise the residuals of the fitting procedure, and supporting evidence for multiple components in the literature (McLaren et al., 2019; Reusser et al., 2020a; Reusser et al., 2020b).

Statistical analysis and graphics

Figures of chromatograms and element plots were created in Microsoft® Excel 2016, whereas all NMR spectra were created using MATLAB® R2017a of The MathWorks Inc. Statistical analyses, including calculation of mean values, standard deviations (SD) and coefficients of determination (R^2) of certain measured parameters, were carried out in Microsoft® Excel 2016. Pearson correlation coefficients r were calculated in R, version 4.0.3 (R Core Team, 2020).

Results

Soil phosphorus

Concentrations of P_{tot} in soil ranged from 864 to 2260 mg P/kg_{soil} (Table 2). Pools of NaOH-EDTA extractable P_{tot} ranged from 41 % (Cambisol (A)) to 95 % (Cambisol (P)) of P_{tot} in each soil. Concentrations of P_{org} in the extracts ranged from 238 mg P/kg_{soil} (Cambisol (A)) to 1135 mg P/kg_{soil} (Cambisol (F)), which comprised on average 67 % of P_{tot} in NaOH-EDTA extracts.

Table 2. Concentrations of P_{tot} in soil determined by XRF and acid digestion, and pools of 0.25 M NaOH + 0.05 M EDTA extractable P determined using the method of Cade-Menun et al. (2002). For the NaOH-EDTA extractable P, mean values (*n*=3) are listed with standard deviations in parentheses.

Soil	XRF	Digestion	NaOH-EDTA extraction		
	P_{tot} mg P/kg_{soil}	**P_{tot}** mg P/kg_{soil}	**P_{tot}** mg P/kg_{soil}	**MRP[a]** mg P/kg_{soil}	**MUP[b]** mg P/kg_{soil}
Cambisol (P)	2553	913	863 ±13	319 (14)	544 (20)
Cambisol (F)	3841	2260	1626 ±13	491 (5)	1135 (17)
Gleysol	2913	1763	1620 ±49	508 (18)	1113 (66)
Cambisol (A)	1724	864	356 ±11	118 (2)	238 (9)

[a] Molybdate reactive P (MRP), considered to be P_{inorg} (Ohno and Zibilske, 1991).

[b] Molybdate unreactive P (MUP), considered to be P_{org} (Ohno and Zibilske, 1991; Hens and Merckx, 2001).

Phosphorus concentrations in molecular size fractions

Concentrations of P_{tot} in the combined molecular size fractions ranged from 189 (Cambisol (A)) to 706 mg P/kg_{soil} (Gleysol), as determined by high-resolution SEC and ICP-OES (Table 3). This accounted for an average 48 % of P_{tot} assessed in unfractionated soil extracts. Concentrations of P_{org} in the combined molecular size fractions ranged from 182 (Cambisol (A)) to 678 mg P/kg_{soil} (Gleysol), as determined from the differences of P_{tot} and MRP. Overall, the distribution of P_{org} across the molecular size fractions can be broadly separated into three main groups (Figure SI-7): 1) concentrations of P_{org} were predominant in molecular size fractions <10 kDa (Cambisol (P) and Gleysol); 2) concentrations of P_{org} were evenly distributed across the molecular size fractions (Cambisol (A)); and

3) concentration of P_{org} were predominant in molecular size fractions >70 kDa (Cambisol (F)).

Concentrations of P_{inorg} in the combined molecular size fractions were low (<45 mg P/kg_{soil}), as determined by high-resolution SEC and NMR spectroscopy (Table 3). The majority of P_{inorg} was measured in the <5 kDa fraction of the soil filtrate following the desalting preparation process (Figure SI-8, Table SI-1). The predominant P species in this fraction was orthophosphate, comprising between 85 % and 91 % of P_{tot}. This fraction also contained minor concentrations of P_{org} as phosphomonoesters, which comprised between 7 % and 13 % of P_{tot}.

Organic phosphorus speciation in molecular size fractions

Diverse NMR signals were evident across all molecular size fractions (Figure 1 and Figure 2, and Figure SI-9), the majority being found in the phosphomonoester region (δ 6.0 to 3.0 ppm) (Figure 1, Figure 2). This region exhibited two main spectral features: a multitude of sharp peaks arising largely from IP, which were predominant in the <5 and 10 kDa fractions; and several broad signals arising from unidentified P_{org} compounds, which were predominant in the >10 kDa fractions.

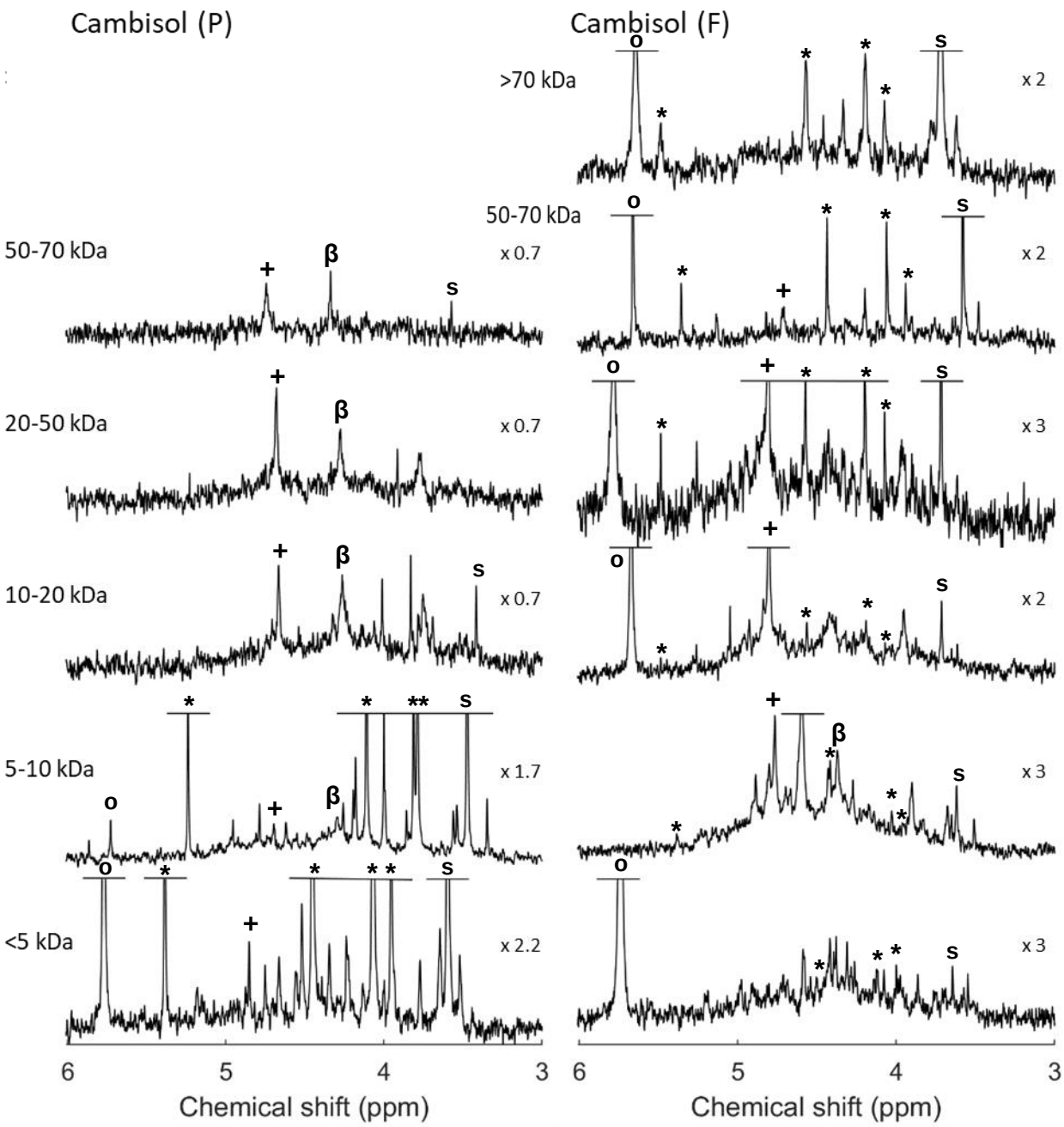


Figure 1. Solution ^{31}P nuclear magnetic resonance (NMR) spectra (500 MHz) of the combined orthophosphate (o) and phosphomonoester region for each of six molecular size fractions generated from 0.25 M NaOH + 0.05 M EDTA soil extracts of the Cambisol (P) and Cambisol (F) soils (Superdex™ 75 prep grade column). Peaks associated with α-glycerophosphate/unknown high molecular weight P_{org} (McLaren et al., 2015b) (**+**), β-glycerophosphate (**β**) as well as *myo*- (*****) and *scyllo*- (**s**) IP_6 are marked. Signal intensities were normalised to a methylenediphosphonic acid standard peak intensity. The vertical axes were scaled for improved visibility of spectral features, as indicated by a factor. The molecular size fractions >70 kDa and 50-70 kDa of the Cambisol (P) were pooled, shown as the >50 kDa spectrum.

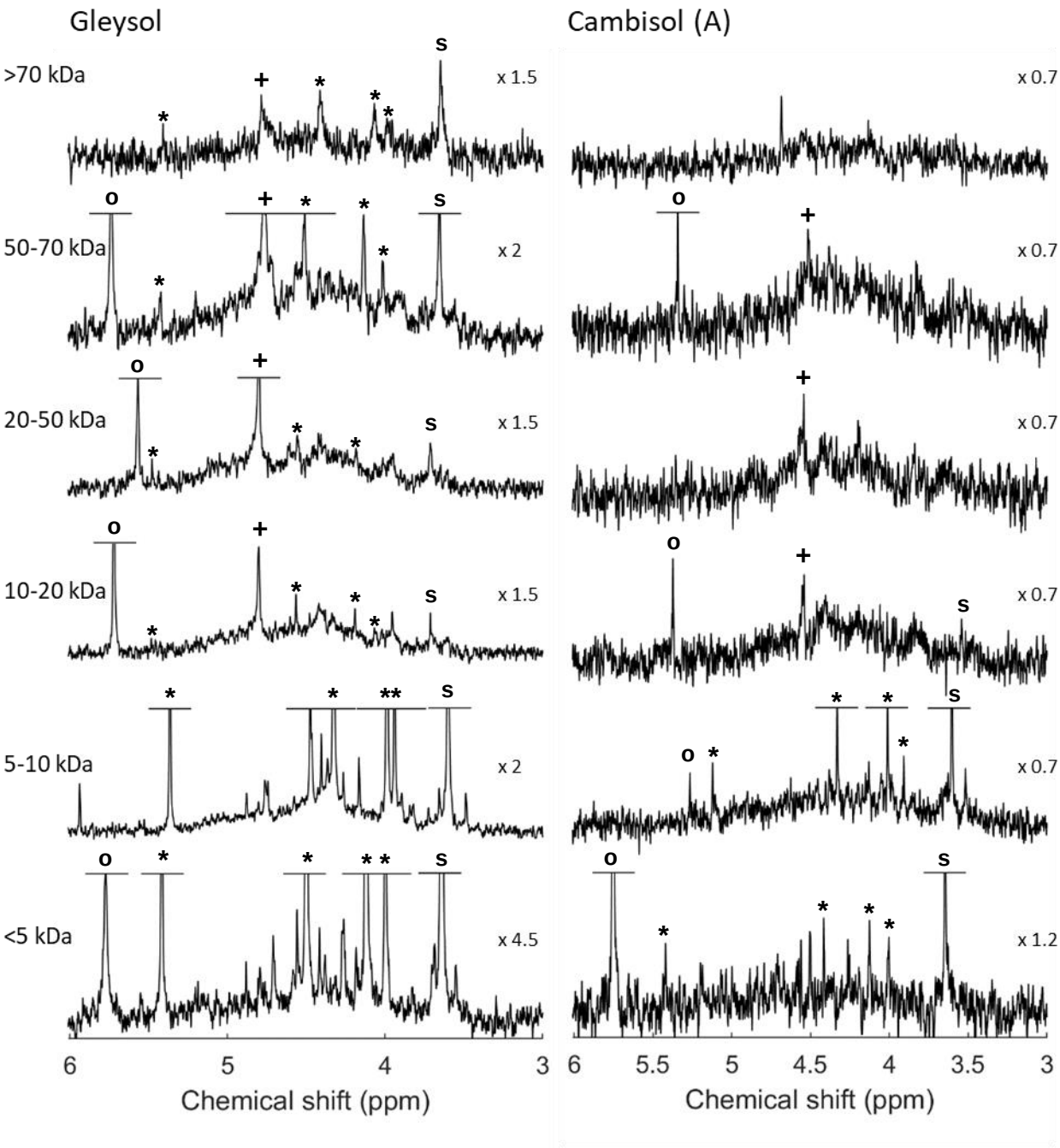


Figure 2. Solution ^{31}P nuclear magnetic resonance (NMR) spectra (500 MHz) of the combined orthophosphate (o) and phosphomonoester region on the molecular size fractions (Superdex™ 75 prep grade column) of 0.25 M NaOH + 0.05 M EDTA soil extracts of the Gleysol and Cambisol (A). The peaks of α-glycerophosphate/unknown high molecular weight P_{org} (McLaren et al., 2015b) (**+**) as well as *myo*- (*****) and *scyllo*- (**s**) IP_6 are marked. Signal intensities were normalised to a methylenediphosphonic acid standard peak intensity. The vertical axes were scaled for improved visibility of spectral features, as indicated by a factor.

Concentrations of phosphomonoesters were highest in the <10 kDa fractions and ranged from 19 (<5 kDa fraction Cambisol (A)) to 120 mg P/kg_{soil} (<5 kDa fraction Gleysol) (Table 3). Sharp signals due to *myo*- and *scyllo*-IP_6 were detected in the <5 and 5-10 kDa fractions of the Cambisol (P) and Cambisol (A) (Table 4). In contrast, these two IP were prominent in several fractions of >20 kDa for the Cambisol (F) and similarly for the Gleysol, but at lower concentrations (Table 4).

Concentrations of lower-IP in the combined molecular size fractions ranged from 2 (Cambisol (A)) to 16 mg P/kg_{soil} (Cambisol (P)) across all soils (Table 4). This included signals of *myo*-(1,2,4,5,6)-IP_5 in the <5 and 5-10 kDa fractions of the Gleysol, *scyllo*-IP_5 in the <5 kDa fractions of all soils and in all molecular size fractions of the Cambisol (F). Furthermore, traces (1.6 to 5.2 mg P/kg_{soil}) of *scyllo*-IP_4 were detected in the <5 and 5-10 kDa fractions of the Cambisol (P) and in the 5-10 kDa of the Cambisol (F).

Several broad signals were present in the phosphomonoester region of all molecular size fractions and represented on average 64 % (SD=20 %) of the total concentration of phosphomonoesters (Table 3, Table 4 and Figure SI-10). These broad signals were particularly prominent in the 10-20 and 20-50 kDa fractions, which comprised an average 77 % and 74 % of the total concentration of phosphomonoesters of each fraction, respectively. In contrast, their contribution to the total pool of phosphomonoesters was lowest in the <5 kDa fraction but were still an important component (on average 44 %). Unlike the other soil samples, the broad signals dominated all size fractions of the Cambisol (A) soil extract. In this sample, on average 3.7 (SD=1.9) broad peaks were fitted in the SDF procedure across all size fractions, whereas for other soil samples this was on average 2.0 (SD=0.7) broad peaks. The average line width at half-height for each broad peak was 88 Hz (SD=53). The half widths of the individual broad peaks decreased with increasing number of fitted broad peaks and were lowest in the Cambisol (A) (67, SD=60), for which most underlying broad peaks were fitted.

Table 3. Concentrations (mg P/kg_{soil}) of total P (P_{tot}) in each molecular size fraction as determined by ICP-OES, and the main P classes (orthophosphate, phosphomonoester, phosphodiester and other

compounds, such as polyphosphates, pyrophosphates and phosphonates) as determined by solution ^{31}P NMR spectral deconvolution fitting. Molecular size fractions were obtained on NaOH-EDTA soil extracts using high-resolution SEC (Superdex™ 75 prep grade column). The proportion (%) of P to P_{tot} in each molecular size fraction is reported in parentheses. The <5 kDa fractions do not account for the fraction removed by the desalting process (Table SI-1).

	P_{tot} ICP-OES (mg P/kg_{soil})	P_{tot} NMR (mg P/kg_{soil})	Ortho-P (mg P/kg_{soil})	P-monoester (mg P/kg_{soil})	P-diester (mg P/kg_{soil})	other P (mg P/kg_{soil})
Cambisol (P)						
<5 kDa	190.4	143.3	18.9 (13)	115.9 (81)	0.0 (0)	8.4 (6)
5-10 kDa	124.7	81.2	0.0 (0)	79.3 (98)	0.2 (0)	1.7 (2)
10-20 kDa	46.4	29.4	0.0 (0)	26.9 (91)	1.2 (4)	1.3 (5)
20-50 kDa	39.7	23.0	0.0 (0)	20.7 (90)	1.1 (5)	1.2 (5)
50->70 kDa[2]	33.5	15.4	0.0 (0)	13.8 (90)	0.0 (0)	1.6 (10)
Cambisol (F)						
<5 kDa	108.9	68.5	24.5 (36)	42.4 (62)	0.0 (0)	1.6 (2)
5-10 kDa	133.5	87.4	0.0 (0)	80.1 (92)	4.1 (5)	3.2 (4)
10-20 kDa	86.2	49.9	5.2 (10)	39.9 (80)	3.4 (7)	1.6 (3)
20-50 kDa	93.4	51.9	5.0 (10)	36.4 (70)	9.0 (17)	1.6 (3)
50-70 kDa	50.7	22.5	3.2 (14)	15.5 (69)	2.0 (9)	1.8 (8)
>70 kDa	229.9	54.7	6.7 (12)	43.1 (79)	2.6 (5)	2.2 (4)
Gleysol						
<5 kDa	73.5	135.8	7.9 (6)	119.9 (88)	0.0 (0)	8.0 (6)
5-10 kDa	73.4	95.4	0.0 (0)	93.5 (98)	0.5 (1)	1.5 (2)
10-20 kDa	103.8	42.2	4.9 (12)	35.5 (84)	1.6 (4)	0.2 (0)
20-50 kDa	81.4	51.6	2.9 (6)	42.9 (83)	5.4 (10)	0.4 (1)
50-70 kDa	172.8	33.1	1.7 (5)	23.7 (72)	6.6 (20)	1.1 (3)
>70 kDa	201.5	20.1	0.0 (0)	14.9 (74)	5.2 (26)	0.0 (0)
Cambisol (A)						
<5 kDa	43.4	25.4	6.0 (24)	19.4 (76)	0.0 (0)	0.0 (0)
5-10 kDa	38.3	25.4	0.4 (2)	24.9 (98)	0.0 (0)	0.1 (0)
10-20 kDa	21.8	14.4	0.4 (3)	13.7 (96)	0.0 (0)	0.2 (2)
20-50 kDa	31.8	16.6	0.0 (0)	16.5 (99)	0.0 (0)	0.1 (1)
50-70 kDa	18.1	9.9	0.3 (3)	8.7 (87)	0.5 (5)	0.6 (6)
>70 kDa	36.0	9.4	0.0 (0)	7.4 (79)	1.8 (20)	0.1 (2)

Table 4. Concentrations (mg P/kg_{soil}) of P species in each molecular size fraction as determined by solution ^{31}P NMR spectroscopy and spectral deconvolution fitting, as described in (Reusser et al.,

[2] The 50-70 kDa and >70 kDa fractions of Cambisol (P) have been pooled together. The given values represent the concentrations measured in the pooled fraction.

2020a). Molecular size fractions were obtained on NaOH-EDTA soil extracts using high-resolution SEC (Superdex™ 75 prep grade column). Pools of 'other IP_6' include the *neo* and *chiro* stereoisomers of IP_6. Pools of lower-order IP (lo. order IP) include *scyllo*-IP_5, *myo*-(1,2,4,5,6)-IP_5 and *scyllo*-IP_4. The concentrations of glycerophosphate (glycerol-P) include the unknown high molecular weight (HMW) P_{org} compound described by McLaren et al. (2015b), whose peak overlaps with α-glycerophosphate. The proportion (%) of P to that of the P_{tot} in each molecular size fraction is reported in parentheses. The <5 kDa fractions do not include the removed fraction by the desalting process (Table SI-1).

	***myo*-IP_6**	***scyllo*-IP_6**	**other IP_6**	**lo. order IP**	**broad peaks**	**glycero-P + HMW P_{org}**
	(mg P/kg_{soil})	**(mg P/kg_{soil})**	**(mg P/kg_{soil})**	**(mg P/kg_{soil})**	**(mg P/kg_{soil})**	**(mg P/kg_{soil})**
Cambisol (P)						
<5 kDa	44.6	17.2	6.0	11.4	26.8 (23)	1.4
5-10 kDa	20.3	8.8	2.4	4.6	32.1 (41)	2.6
10-20 kDa	-	0.6	-	-	18.0 (67)	3.3
20-50 kDa	-	-	-	-	13.7 (66)	3.2
50->70 kDa[3]	-	0.3	-	-	9.1 (66)	3.4
Cambisol (F)						
<5 kDa	1.0	0.7	-	1.4	26.3 (62)	-
5-10 kDa	0.1	1.4	-	3.4	51.5 (64)	6.2
10-20 kDa	0.6	0.6	-	0.3	30.4 (76)	3.2
20-50 kDa	2.9	2.1	-	0.9	23.2 (64)	2.4
50-70 kDa	3.0	2.1	-	0.8	6.6 (43)	0.2
>70 kDa	8.1	9.4	-	3.1	21.5 (50)	-
Gleysol						
<5 kDa	50.5	23.1	7.8	7.5	25.7 (21)	-
5-10 kDa	33.1	13.1	1.7	5.0	33.3 (36)	-
10-20 kDa	0.7	1.6	-	-	27.5 (77)	2.5
20-50 kDa	1.2	2.4	-	-	34.0 (79)	4.1
50-70 kDa	1.9	1.5	0.2	-	16.6 (70)	2.8
>70 kDa	2.5	1.7	-	-	9.6 (65)	1.0
Cambisol (A)						
<5 kDa	1.9	2.8	-	-	13.7 (71)	-
5-10 kDa	2.0	1.7	-	2.0	17.6 (71)	-
10-20 kDa	-	0.1	-	-	11.8 (86)	1.7
20-50 kDa	-	-	-	-	14.5 (88)	2.0
50-70 kDa	-	-	-	-	8.1 (93)	2.4
>70 kDa	-	-	-	-	7.0 (95)	-

Concentrations of phosphodiesters were highest in molecular size fractions above 10 kDa (Table 3). The peak associated with DNA was most prominent in the 20-50 kDa fraction of

[3] The 50-70 kDa and >70 kDa fractions of Cambisol (P) have been pooled together. The given values represent the concentrations measured in the pooled fraction.

the Cambisol (F) and in the 50-70 kDa and >70 kDa size fraction of the Gleysol, which comprised more than 25 % of the total NMR signal (Table 3). Other P compounds including phosphonates and pyrophosphates comprised less than 10 % of the total NMR signal (Table 3).

The NMR observability in the fraction removed by the desalting process (<5 kDa) was on average 91 %. In contrast, the NMR observability of the analysed molecular size fractions differed markedly: on average 27 % (SD=1 %) in the >70 kDa fraction, 48 % (SD=4 %) in the 50-70 kDa fraction, 54 % (SD=3 %) in the 20-50 kDa fraction, 60 % (SD=5 %) in the 10-20 kDa fraction, 63 % (SD=5 %) in the 5-10 kDa fraction and 66 % (SD=6 %) in the <5 kDa fraction.

Metal concentrations in molecular size fractions

In general, concentrations of Fe were highest in the <5 and >70 kDa molecular size fractions of all soils (Figure 3). The ratio of Fe to Al in NaOH-EDTA extracts differed between soils and molecular size fractions. For the Cambisol (P), Al dominated the lowest size fraction, whereas Fe dominated the highest molecular size fraction. For the Gleysol, increasing concentrations of Fe were measured with increasing molecular size, not including the <5 kDa fraction. The Cambisol (F) exhibited a similar distribution pattern of the Fe and Al concentrations across the molecular size fractions. For the Cambisol (A), Al was only detected in the highest molecular size fraction. This soil was alkaline with a pH of 7.7 measured in water (Meyer et al., 2017). In contrast, the pH of the other soil samples ranges from 3.6 (Cambisol (F)) to 5.1 (Cambisol (P)) (Reusser et al., 2020a). Concentrations of IP across all size fractions and soil samples were not significantly correlated with the measured Fe (r=0.048, p=0.829) or Al (r=0.123, p=0.576) concentrations. However, considering all molecular size fractions of the individual soils, IP concentrations were significantly positively correlated with Fe concentrations in the Cambisol (F) sample (r=0.847, p=0.033) and with Al concentrations in the Gleysol sample (r=0.823, p=0.044). These correlations were much less apparent in the other soils.

Calcium concentrations were highest in the <5 kDa fraction for all soils and decreased rapidly with increasing molecular size.

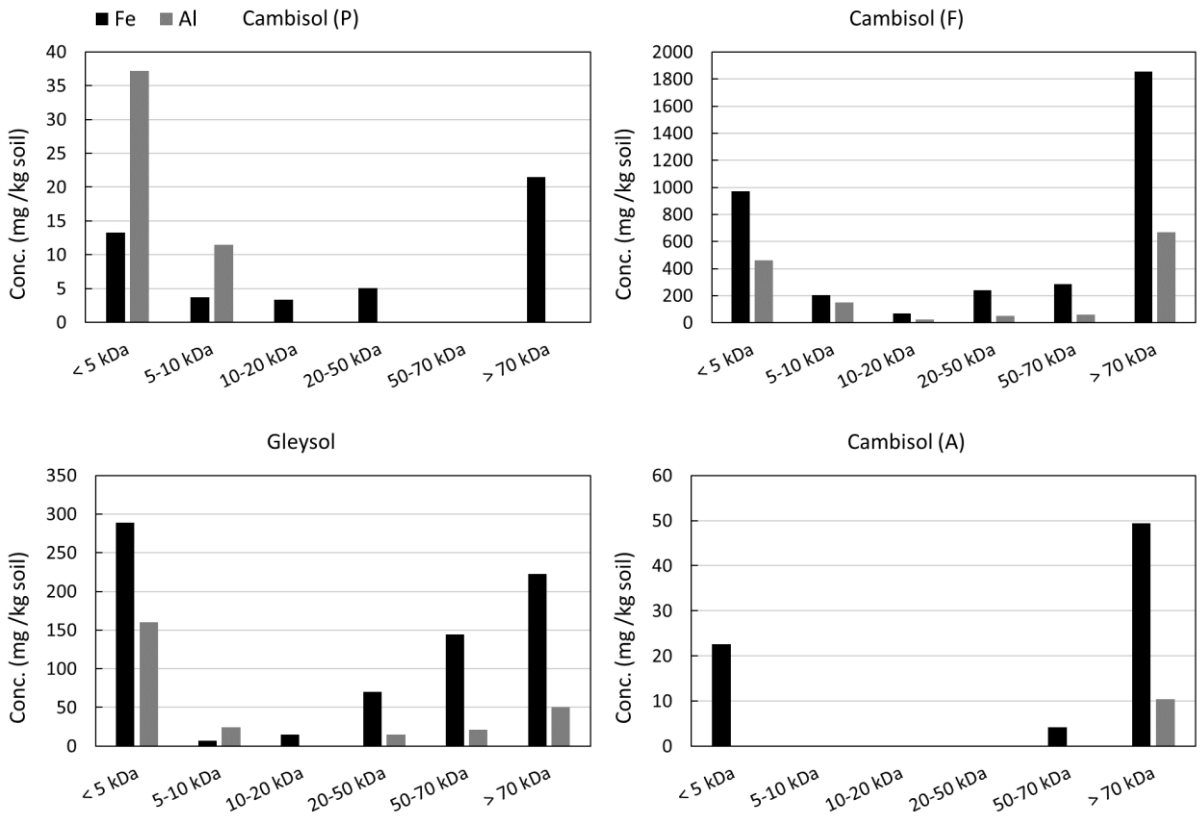


Figure 3. Concentrations of aluminium (Al, grey) and iron (Fe, black) in 0.25 M NaOH + 0.05 M EDTA soil extracts following passage through a Superdex™ 75 prep grade column, which resulted in several technically defined fractions with molecular sizes of <5, 5-10, 10-20, 20-50, 50-70, and >70 kDa. Concentrations in the molecular size fractions were analysed by ICP-OES.

Discussion

Phosphorus concentrations in molecular size fractions

The molecular weight distribution of P_{org} in soils has typically been studied using technically defined molecular weight fractions using cut-offs of either 5 kDa (Jarosch et al., 2015) or 10 kDa (McLaren et al., 2015b). We present the continuous molecular size distributions of extracted soil P. Our reported molecular size distribution of P ranges from small molecules, i.e. orthophosphate, to large molecular size fractions approaching nanoparticles (1-100 nm (Gottselig et al., 2017)). This study bridges research on small identifiable P_{org} molecules, such as lower-order IP (Reusser et al., 2020b), with research on colloid-associated P (Missong et al., 2016). Furthermore, our quantitative results of P_{inorg} and P_{org} concentrations in different molecular size fractions confirm the previously conjectured continuum of P along the molecular size distribution.

Pools of P_{org} exhibited a wide concentration range in the investigated soils (238 to 1135 mg P/kg_{soil}). Concentrations of NaOH-EDTA extractable P were similar to those reported in previous studies (Jarosch et al., 2015; McLaren et al., 2019). SEC chromatograms revealed a broad molecular size distribution of P in NaOH-EDTA extracts from soil. Furthermore, the majority of P_{org} was present in molecular size fractions >5 kDa across all soils investigated here, despite differences in their chemical and physical properties. Several studies using SEC have reported that a large portion of P_{org} in soil extracts is detected in molecular size fractions above 10 kDa (Moyer and Thomas, 1970; Steward and Tate, 1971; Veinot and Thomas, 1972). The consistency and abundance of this high molecular size P_{org} pool in different soil systems suggests its importance in the soil P cycle.

We identified three P_{org} distribution patterns in high molecular size material: 1) concentrations of P_{org} were predominant in molecular size fractions <10 kDa (Cambisol (P) and Gleysol); 2) concentrations of P_{org} were evenly distributed across all molecular size fractions (Cambisol (A)), and; 3) concentrations of P_{org} were predominant in molecular size fractions >70 kDa (Cambisol (F)). As discussed in the following sections, each specific distribution of P_{org} in high molecular size material could be attributed to differing contents of SOM, metals and soil minerals.

Speciation of organic P compounds in molecular size fractions

Solution ^{31}P NMR spectroscopy on soil extracts revealed a large diversity of P compounds across all molecular size fractions and soil samples. Our results demonstrate a much greater diversity of organic P compounds along the molecular size continuum than was evident in previous studies (Jarosch et al., 2015; McLaren et al., 2015b).

The molecular size material below 5 kDa was dominated by orthophosphate in the fraction removed by the desalting process and by phosphomonoesters in the <5 kDa fraction of the subsequent high-resolution SEC. In the Cambisol (P) and Gleysol samples, the <5 kDa fraction consisted mostly of IP, which were probably in the free form or desorbed from soil mineral surfaces (Moyer and Thomas, 1970; Celi and Barberis, 2007). The fraction removed by the desalting process (cut-off=5 kDa) was less enriched in IP compared to the <5 kDa fraction of the subsequent high-resolution SEC of the desalted extract, despite the molecular size of IP being 660 Da or less. Nebbioso and Piccolo (2012) reported that chromatographic separation weakens the supramolecular architecture of SOM. This may indicate that associations of IP with higher molecular size compounds were weakened during the desalting chromatography step. The liberated IP may then have been eluted in the <5 kDa fraction in the subsequent high-resolution SEC step.

Spectra derived from molecular size fractions greater than 5 kDa were dominated by poorly resolved broad ^{31}P NMR signals. Compounds causing these broad NMR signals are apparently of large molecular size. These findings are consistent with the studies of Jarosch et al. (2015) and McLaren et al. (2015b). We also employed multiple analytical approaches to demonstrate that the underlying broad signal in the NMR phosphomonoester region is not derived from a single molecule with a discrete molecular size, but rather from a continuum of unidentified P_{org} molecules spanning a wide molecular size distribution. This is in agreement with observations of McLaren et al. (2019) and Reusser et al. (2020b). In the four studied soils, broad signals were most prominent in the 10-20 and 20-50 kDa fractions rather than the highest molecular size fractions.

The number of underlying peaks necessary to account for the broad feature by spectral deconvolution fitting varied between soils and molecular size fractions. Previously, a single broad signal has been used to represent the ‘unresolved’ P_{org} pool (Reusser et al., 2020a). However, this approach results in an imperfect Lorentzian/Gaussian distribution, explaining the high variations of the C5 peak of *myo*-IP_6 in spectral deconvolution fitting results reported by Reusser et al. (2020a). The line width at half-height of 88 Hz in this

study provides evidence for the existence of more than one underlying broad peak, suggesting the presence of several compounds. McLaren et al. (2019) reported line widths at half-height of the fitted broad signal were on average 196 Hz, however transverse relaxation times suggest the line widths are closer to 11 Hz. In addition, the underlying broad peak at chemical shift region δ 4.5-4.0 ppm appears to be more pronounced compared to the fitted broad peaks more to the side of the phosphomonoester region. This could suggest that 1) the net peak area of sharp peaks arising from other phosphomonoesters in this region is influenced by the underlying broad signal. If this is the case, these peaks may be more susceptible to overestimation when underlying broad signals are not accounted for in SDF procedures compared to other compounds at the side of the chemical shift region; and 2) the restriction of the chemical shift region where most of the signal of the unresolved P_{org} pool appears will help to identify P containing molecular structures resonating in this specific region. This could improve our understanding of the chemical nature of this pool in future studies.

Furthermore, the results of this study show that at least 32 % (Cambisol (P)) and up to 93 % (Cambisol (F)) of total detected IP was present in molecular size fractions greater than 5 kDa. These findings are in agreement with other studies as Veinot and Thomas (1972), Hong and Yamane (1981) and Borie et al. (1989). Moreover, we provide direct spectroscopic evidence that IP are present in large molecular size fractions of soil extracts as well as information on their speciation.

The mechanisms by which IP become associated with large molecular size material is assumed to be due to a close association of IP with compounds of large molecular size, such as SOM and/or relatively small (nano-) sized soil minerals (Moyer and Thomas, 1970; Omotoso and Wild, 1970; Borie et al., 1989; Celi and Barberis, 2005). Possible mechanisms involve metal bridges, presumably with Fe or Al (Veinot and Thomas, 1972; Vincent et al., 2012), or nucleophilic addition of free amine groups to aromatic rings (Borie et al., 1989). The occurrence of the latter mechanism during phosphorylation of model humic polymers was reported by Brannon and Sommers (1985). Cosgrove (1977) suggested that high molecular weight material consists of complexes of Fe(III) and nitrogenous organic matter containing IP_6. Another possibility could be that the phosphate groups of IP associate with SOM *via* ester-bonds, forming part of the SOM suprastructure model introduced by Piccolo (2001). Jørgensen et al. (2015) reported that the relative importance of the main binding sites (SOM matrix, amorphous metal oxides or clay minerals) in stabilising IP_6 can vary markedly between soils and that the association of IP_6

with SOM can be important in soils despite high contents of amorphous iron and aluminium oxides. The Cambisol (F) studied here not only has high organic carbon contents but also the highest amounts of oxalate extractable (amorphous) Fe and Al. Such findings emphasise that the mechanisms of incorporation of IP into large organic molecules and a possible role of IP in stabilising SOM, as proposed by Tipping et al. (2016), warrant further investigation. For example, targeted degradation of the soil IP pool for the purpose of plant-nutrition could result in a concomitant destabilisation of SOM structure. The results of this study show that IP associated with large molecular size material are mainly in the *myo* and *scyllo* forms, suggesting these stereoisomers are more prone to association with high molecular size compounds and reduced enzymatic hydrolysis compared to *neo* and *chiro* forms.

We present evidence that molecular size fractions greater than 5 kDa consist not only of IP_6, as previously proposed, but also of lower-order IP. *scyllo*-IP_5 was detected in all molecular size fractions of the Cambisol (F) soil extract, suggesting similar mechanisms of association to IP_6. Association of a proportion of IP with organic phases may limit not only the accessibility of IP_6 to enzymatic hydrolysis in soils, but also a fraction of lower-order IP. However, this is likely to depend on the stereochemistry of the associations, including the spatial protection conferred to IP, as well as whether IP is linked to the SOM matrix by metal bridges or ester bonds, which might be more readily hydrolysed by phosphatases.

Our findings regarding phospholipids and phosphodiesters in large molecular size material are comparable with the studies of Missong et al. (2016) on forest soil colloids and Wang et al. (2017) using DOSY NMR spectroscopy of NaOH-EDTA soil extracts. The presence of glycerophosphates in large molecular size fractions could be explained by hydrolysis of large phospholipid molecules by NaOH (Doolette et al., 2011a) or associations of glycerophosphates, which were intrinsically present in the soil sample (Wang et al., 2021), with large molecular size compounds. However, McLaren et al. (2015b) reported the presence of a high molecular weight P_{org} compound at a similar chemical shift to the α-glycerophosphate peak. We could not resolve these two peaks in our spectra, hence the concentrations of glycerophosphate and this unknown high molecular weight compound are combined. We assume that this compound, and not α-glycerophosphate, is mainly present in the Cambisol (F), Gleysol and Cambisol (A) soil samples due to the lack of the corresponding β-glycerophosphate peak.

Gerke (1992) proposed that a dominant proportion of soil P is present in the form of P_{inorg} (as orthophosphate) associated with organic molecules via metal bridges, i.e. Fe^{3+} and Al^{3+}, forming large molecules. The author states that often over 50 % of these associations may be hydrolysed during spectrophotometric determination of P_{org} and P_{inorg} using the phosphomolybdate method. Hence, P in these associations would be misattributed to the free orthophosphate pool. In our study, orthophosphate is present in large molecular size fractions but to a much lesser extent compared to the <5 kDa fractions. The predominant proportion of P in large molecular size fractions is present as phosphomonoesters, and to a lesser extent phosphodiesters. We propose that 1) IP complexed by SOM and large molecular size material through metal bridges (Borie et al., 1989) and 2) P linked via ester bonds to the SOM organic molecules (McLaren et al., 2015b; McLaren et al., 2020) contribute to the majority of large molecular size soil P.

Nuclear magnetic resonance observability

Few studies on solution ^{31}P NMR spectroscopy of soil samples report the NMR observability as recommended by Doolette et al. (2011a). In our previous studies on the same soil samples, NMR observability values ranged from 52% to 94% (Reusser et al., 2020a; Reusser et al., 2020b). Reduction of the NMR observability with increasing molecular size in this study could be related to an increase of precipitates at the bottom of the microcentrifuge tubes after centrifugation at 10621 *g*. Missong et al. (2016) used ultracentrifugation at 14000 *g* for 60 min to separate water-dispersible colloids from the electrolyte phase. In our study, precipitation of colloids containing P not removed by filtration but by centrifugation could account for the low NMR observability of large molecular size material. To test this, we measured total P contents in the precipitates by total digestion and subsequent P analysis by ICP-OES. Results indicate that on average only 7 % (SD=6 %) of P_{tot}, representing 4.5 mg P/kg_{soil} (SD=3.6 mg P/kg_{soil}), was lost to precipitation. However, an exception was the >70 kDa fraction of the Cambisol (F), which had a concentration of 121.3 mg P/kg_{soil} in the precipitate, representing 58 % of P_{tot}. This soil sample contains a large amount of amorphous and complexed Fe and Al oxides (Table) as well as a large amount of Fe in the largest molecular size fraction. The increase in the pH during sample preparation for NMR may have resulted in complex formation of IP and other P_{org} compounds with metals and organic matter, decreasing their solubility (Martin and Evans, 1987; Celi and Barberis, 2005; Hiradate et al., 2006). P concentrations in precipitates were correlated with NMR observability in the Cambisol (F) and Cambisol (A) but not in the Cambisol (P) and the Gleysol (Figure SI-11). Decreasing NMR

observability with increasing molecular size due to precipitation suggests that quantification and identification of P compounds in large molecular size material by solution ^{31}P NMR spectroscopy could be underestimated in some soils.

Metal concentrations in molecular size fractions

Multivalent cations, including Al^{3+} and Fe^{3+}, are known to form complexes with organic compounds possessing free electron pairs such as amines, carboxylates and ethers, as well as phosphate anions (Parfitt, 1979; Stevenson, 1994; Turner et al., 2002; Gerke, 2010; Vincent et al., 2012). In our study, Al was present mostly at elution volumes in the range of 1 kDa (single peak at K_{av}=0.7, Figure SI-4). This large signal is assumed to arise from metal-EDTA complexes, EDTA being present in the extractant. We assume that most Zn and Al was either present in the free ion form or was liberated from soil compounds during extraction.

Concentrations of Fe were highest in the <5 kDa and >70 kDa molecular size fractions: the first pool is likely to be dominated by Fe-EDTA complexes; the second by Fe-oxyhydroxides or Fe bound to/complexed with oxygen containing organic compounds (Stevenson, 1994), based on their molecular sizes (Lead and Wilkinson, 2006). When converting molecular size in Da to the nm scale, a spherical molecule with a molecular size of 70 kDa would exhibit a minimum diameter of approximately 5.5 nm (Erickson, 2009). Fe-oxyhydroxides range from 1 nm to 1 µm in size (Lead and Wilkinson, 2006). The high Fe concentration in the >70 kDa fraction of the Cambisol (F) could also indicate the presence of P_{org} adsorbed to Fe oxyhydroxide nanoparticles (Ognalaga et al., 1994; Lead and Wilkinson, 2006; Missong et al., 2016) or P_{org} complexed with Fe^{3+}. These hypotheses are supported by the positive correlation of IP concentrations in the size fractions of the Cambisol (F) sample with measured Fe. Celi and Barberis (2005) reported that indirect adsorption of P_{org} through bi- or trivalent cations with the resulting formation of ternary organic matter-metal-P complexes is one abiotic process which contributes to incorporation of P_{org} into SOM. Furthermore, Fe oxides were identified as important sorbents of organic matter and thus playing a major role in controlling long-term SOM preservation (Kögel-Knabner et al., 2008).

The differing molecular size distributions of Fe and Al apparent from our chromatograms indicate that total soil concentrations of Al and Fe do not reflect the abundance of larger molecular size complexes containing P and that the association of P_{org} with Al in large molecular size complexes is less important than its association with Fe. Further research

is needed to understand the association of P_{inorg} and P_{org} with Fe and Al in large molecular size material to elucidate the stability and cycling of these complexes in soil systems.

Conclusion

The majority of P_{org} in soil extracts are associated with molecular sized material above 5 kDa. For the first time, we report the presence of the *myo* and *scyllo* stereoisomers of IP_6, as well as some lower-order IP species, in molecular size material above 5 kDa by combining high-resolution size exclusion chromatography and solution ^{31}P NMR spectroscopy. Our findings also show that the largest pool of P_{org} in large molecular size material is generally represented by three or four underlying broad signals in the phosphomonoester region of NMR spectra. We provide further evidence for an unresolved P_{org} pool in soil being responsible for poorly resolved broad signals in the phosphomonoester region of soil extracts, and that this is comprised of a continuum of P_{org} compounds having a broad range of molecular sizes. This unresolved pool of phosphomonoesters contains multiple components rather than single polymers of discrete molecular sizes. The corresponding NMR signal is comprised of several overlapping broad peaks with varying peak shapes and line widths, indicating diverse compositions of the unresolved P_{org} pool. Furthermore, our results suggest that this unresolved P_{org} pool is represented by P_{org} associated with SOM rather than orthophosphate adsorbed to organic molecules via metal bridges. Nevertheless, metals, particularly Fe, appear to be an important bridging component of P_{org} associations in large molecular size material. The findings of this study will forward our understanding of P_{org} - SOM associations and their importance in the soil P cycle.

Supporting information

High-pressure liquid size exclusion chromatography with in-line ICP-MS analysis

High-pressure liquid (HPL) size exclusion chromatography (SEC) with in-line inductively coupled plasma-mass spectrometry (ICP-MS) was carried out on soil filtrates to acquire highly resolved P and metal chromatograms. The SEC-ICP-MS analyses were carried out based on the method of Dell'Aquila et al. (2020) at Rothamsted Research in Harpenden (UK). A 20 mL aliquot of the NaOH-EDTA soil filtrate was frozen at -80 °C and lyophilised. This resulted in a freeze-dried mass of on average 647 mg across all soils. The lyophilised material was redissolved in 3 mL of deionised water, or 30 mL for the Gleysol soil, and then filtered with a 0.2 µm syringe filter. A preliminary study showed that the viscosity of the samples was too high for the SEC. Hence, the redissolved extracts were diluted further with a dilution factor (DF)=4 for the Cambisol (P) and the Gleysol, a DF=10 for the Cambisol (F) and a DF=2 for the Cambisol (A) soil sample.

The HPL-SEC was carried out on the soil filtrates using the Flexar™ HPLC system (PerkinElmer Inc., Waltham, USA) and the Superdex™ Peptide 10/300 GL column (GE Healthcare, Chicago, USA). A 100 µL aliquot of soil filtrate was injected and subsequently eluted with degassed 0.2 M ammonium-nitrate (pH approx. 7.6) at a flow rate of 0.6 mL/min and a pump pressure of 1.2 MPa. The runtime was set at 40 min and carried out under constant temperature (25 °C). Furthermore, an in-line UV/VIS detector (wavelength=280 nm) was connected with a NexION™ 300 ICP-MS (PerkinElmer Inc., Waltham, USA), which determined concentrations of Al, Co, Cu, Fe, Mn, Zn and P with elution time.

To evaluate the void volume (V_o), and calculate the partition coefficient (K_{av}) of organic compounds, an experiment was carried out using compounds with known molecular sizes. This involved a mixture of standards (150 µL each) containing vitamin B_{12} (0.2 mg/mL), triglycine (10 mg/mL), cytochrome c (1 mg/mL) and blue dextran, which was passed through the column. Equation SI-1 was used to calculate the K_{av} (Ó'Fágáin et al., 2011):

$$K_{av} = \frac{V_e - V_o}{V_t - V_o}, \qquad \text{(SI-1)}$$

where V_o is the void volume in mL (i.e. elution volume of blue dextran), V_e is the elution volume in mL (i.e. the volume from sample injection to the peak maximum of a compound), and V_t is the total volume of the packed bed in mL.

The values of K_{av} for the standards were plotted against their molecular size in Dalton, which resulted in a selectivity curve of the column (Figure SI-1). Consequently, the elution times of compounds in the soil filtrates are measured and their molecular size estimated. These calculations are based on the assumption that all compounds have the same symmetrical shape (Ó'Fágáin et al., 2011).

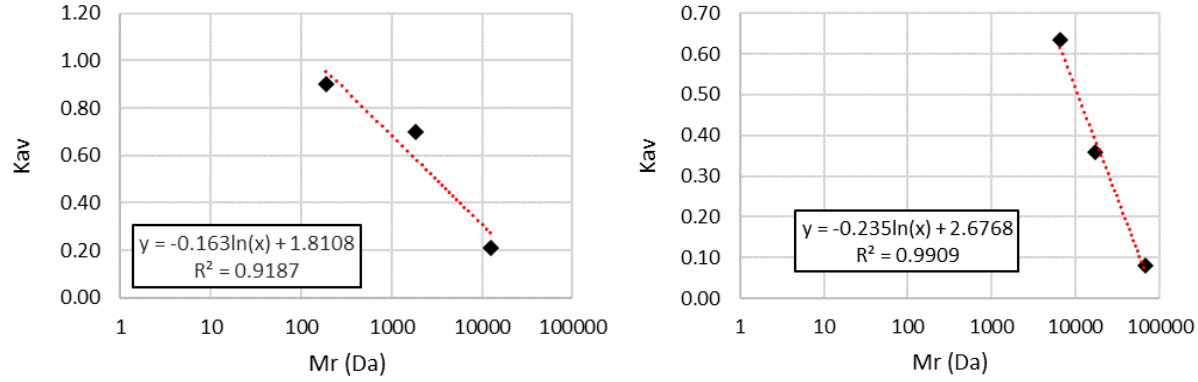


Figure SI-1. Selectivity curves of standards for Superdex™ Peptide 10/300 GL (left) and Superdex™ 75 prep grade (right). The calculated partition coefficient K_{av} is plotted against the molecular size Mr.

All chromatograms were baseline corrected by subtracting the average number of counts per second in the first 5 minutes of elution from the overall chromatogram. Calibration curves were used to determine the concentration of P_{tot} and metals in solution. The coefficients of determination (R^2) of the calibration curves were 0.9998 for P, 0.9986 for Al, 0.5987 for Mn, 0.8239 for Fe and 0.7656 for Zn.

Results

The P chromatograms of NaOH-EDTA extracts were dominated by the orthophosphate signal at molecular size <1 kDa in all six soils, as determined by HPL-SEC with in-line ICP-MS (Figures SI-2 and SI-3). In addition, three main features can be distinguished in the P chromatograms: 1) the presence of P signals corresponding to molecular sizes of >20 kDa, which were particularly dominant in the Cambisol (F), Gleysol and Cambisol (A) soils; 2) a continuum of P signals corresponding to molecular sizes ranging from 1 to <20 kDa; and 3) a broad signal corresponding to a molecular size range of 1 to 5 kDa, which was particularly dominant in the Cambisol (P), Cambisol (F) and Gleysol.

The ultraviolet-visible (UV-VIS) chromatograms of the SEC-ICP-MS measurements were partitioned into four main regions (from left to right, Figures SI-2 and SI-3). This included: 1) a sharp signal at K_{av}=0 corresponding to organic compounds with molecular sizes >20 kDa; 2) a continuum of signals corresponding to organic molecules with molecular sizes between 1 kDa and <20 kDa; 3) a sharp signal corresponding to organic compounds with a molecular size of ~1 kDa, and attributed to the previously described EDTA complexes, and; 4) one to four sharp signals corresponding to organic compounds with molecular sizes <1 kDa. This information was used to carry out further size separation of soil filtrates using high-resolution SEC, which provided samples with sufficient concentrations of P_{org} in each molecular weight range for subsequent characterisation using solution ^{31}P NMR spectroscopy.

Concentrations of Fe, Al, manganese (Mn) and zinc (Zn) were highest in the Gleysol and Cambisol (F), as determined by ICP-MS (Figure SI-4). Concentrations of Fe were highest in molecular size fractions above 20 kDa for all soils. The chromatograms of Mn, Al and Zn were all dominated by a peak at K_{av}=0.7, corresponding to a molecular size of 1 kDa (Figure SI-4). In addition, there was elevated concentrations of Mn in fractions with a molecular size of 1 kDa.

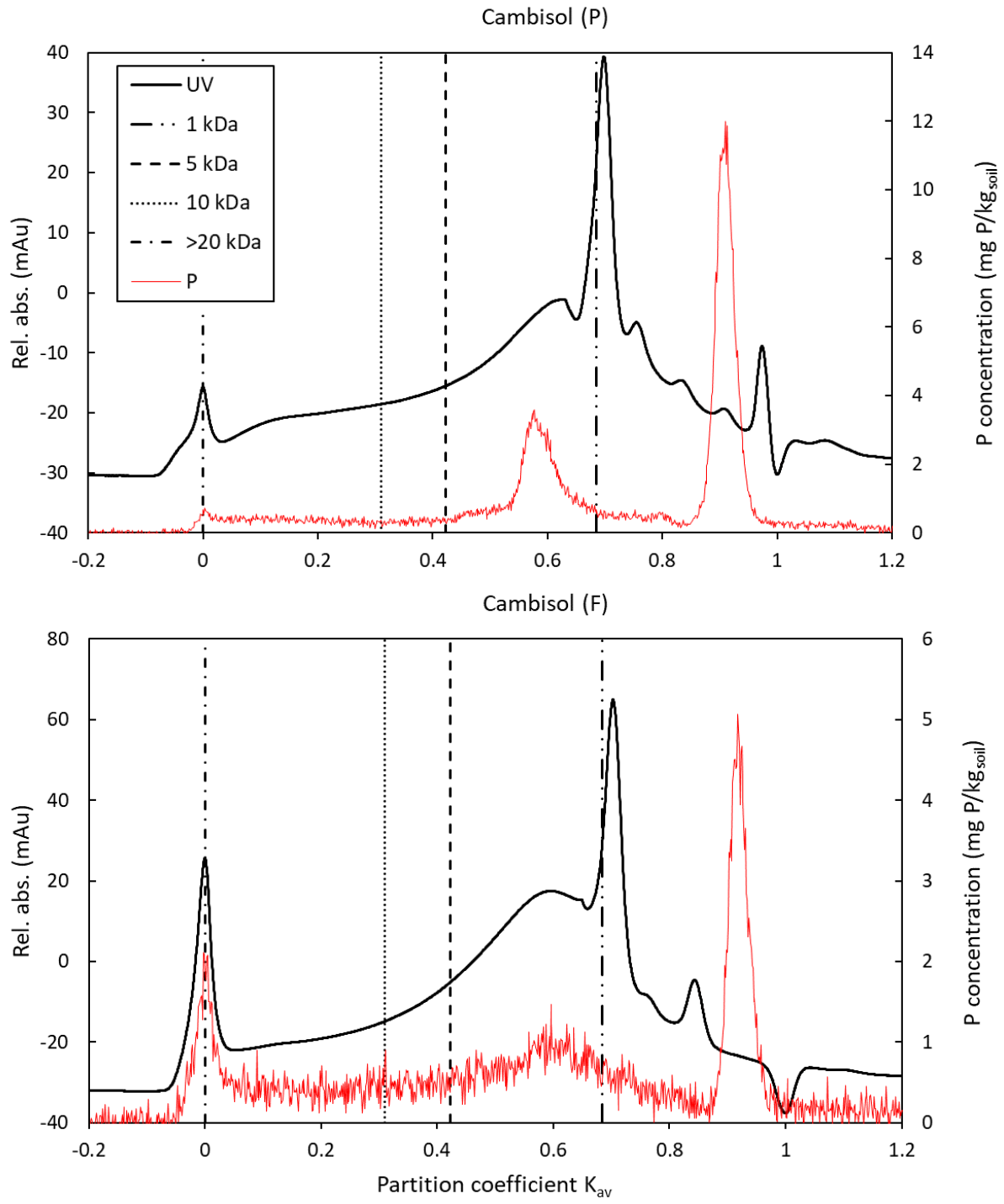


Figure SI-2. The UV-VIS absorbance (black) and phosphorus (P, red) chromatograms of the HPL SEC-ICP-MS measurements of the 0.25 M NaOH + 0.05 M EDTA soil extracts. The partition coefficients at molecular sizes of 1, 5 10 and >20 kDa are marked with dashed lines.

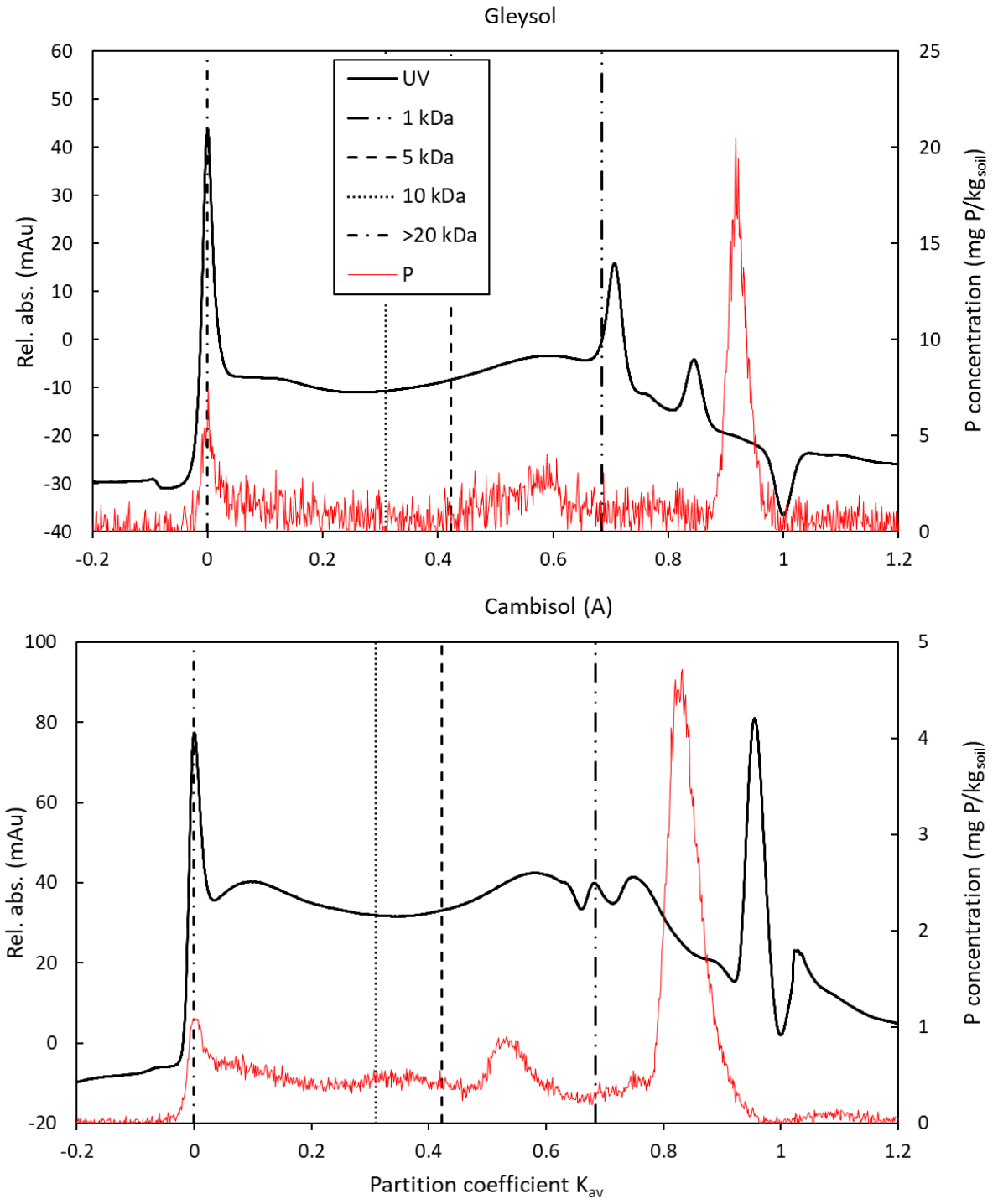


Figure SI-3. The UV-VIS (black) and P (red) chromatograms of the HPL SEC-ICP-MS measurements of the 0.25 M NaOH + 0.05 M EDTA soil extracts. The partition coefficients at molecular sizes of 1, 5 10 and >20 kDa are marked with dashed lines.

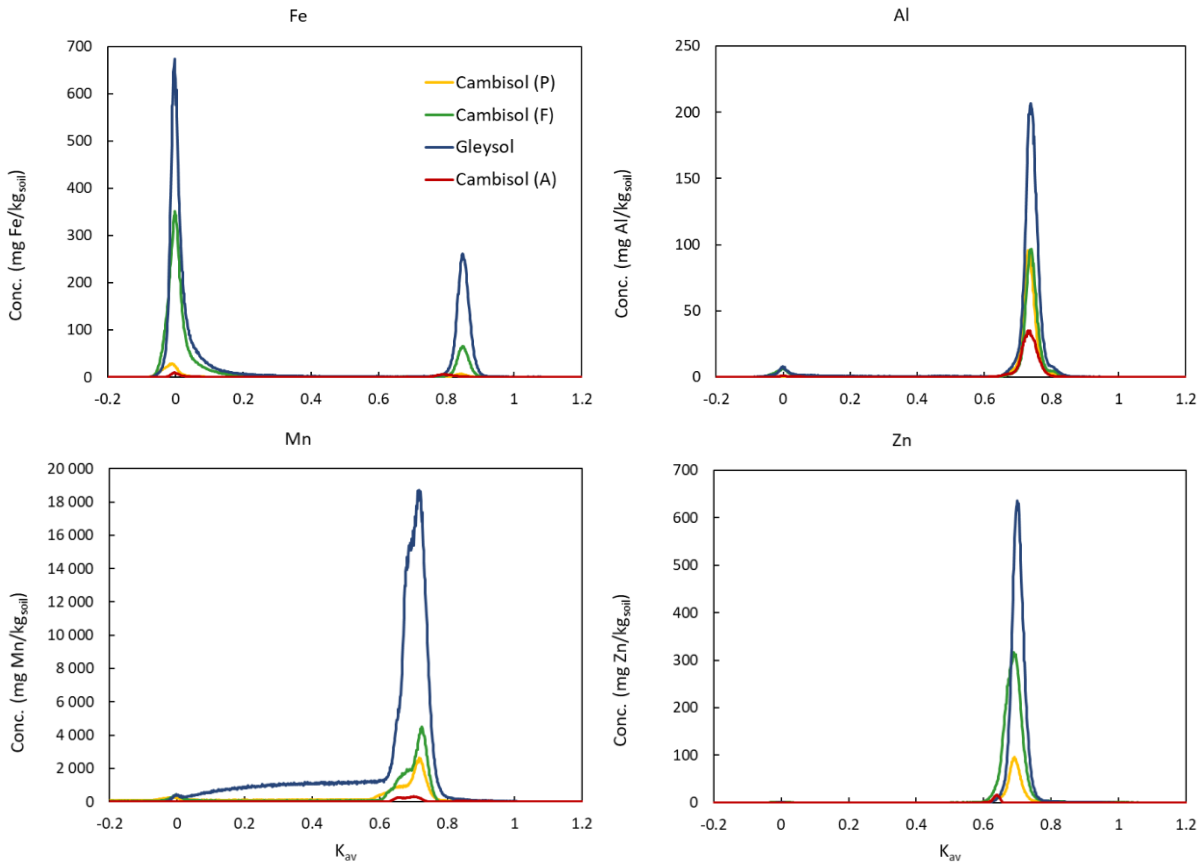


Figure SI-4. Iron (Fe), aluminium (Al), manganese (Mn) and zinc (Zn) chromatograms of the HPL-SEC ICP-MS measurements of the 0.25 M NaOH + 0.05 M EDTA soil extracts. The metal concentrations in mg/kgsoil are plotted against the partition coefficient K_{av}.

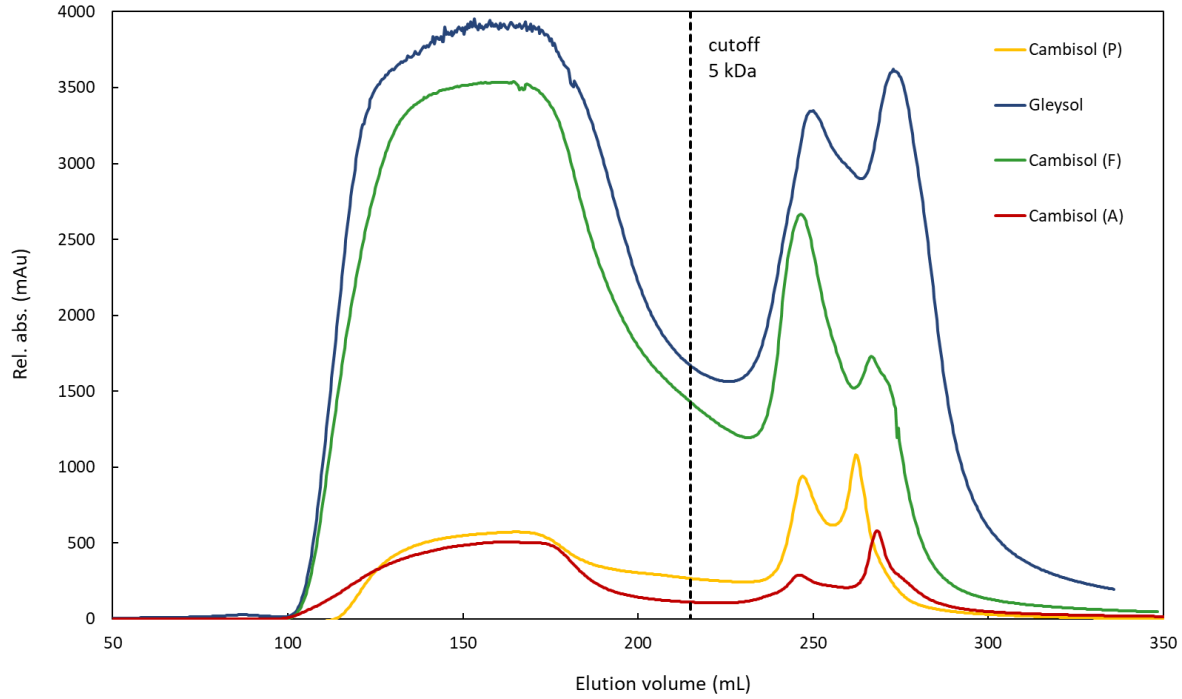


Figure SI-5. UV-VIS absorbance of the desalting process of NaOH-EDTA soil extracts eluted in Sephadex™ G-25 gel media. Dashed black line: cut-off of 5 kDa.

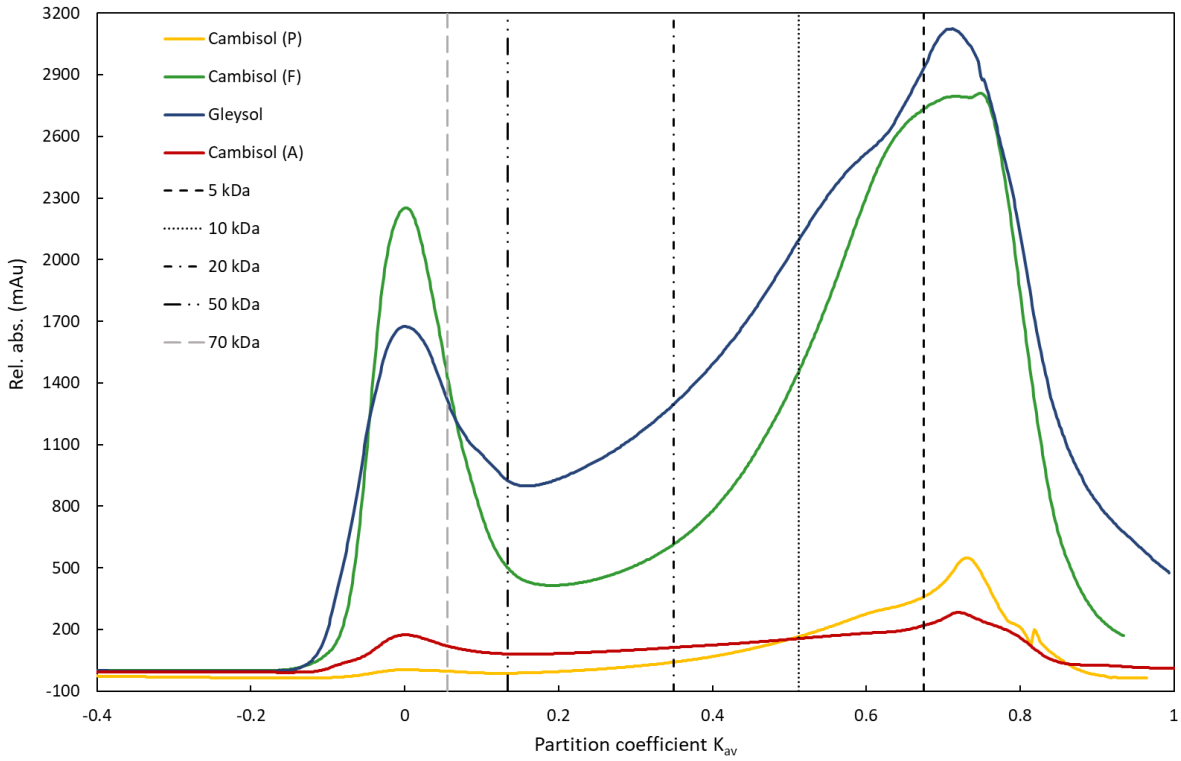


Figure SI-6. UV-VIS chromatograms of NaOH-EDTA soil extracts eluted in Superdex™ 75 prep grade gel media. The partition coefficients related to molecular sizes of 5, 10, 20, 50 and 70 kDa are marked with dashed lines.

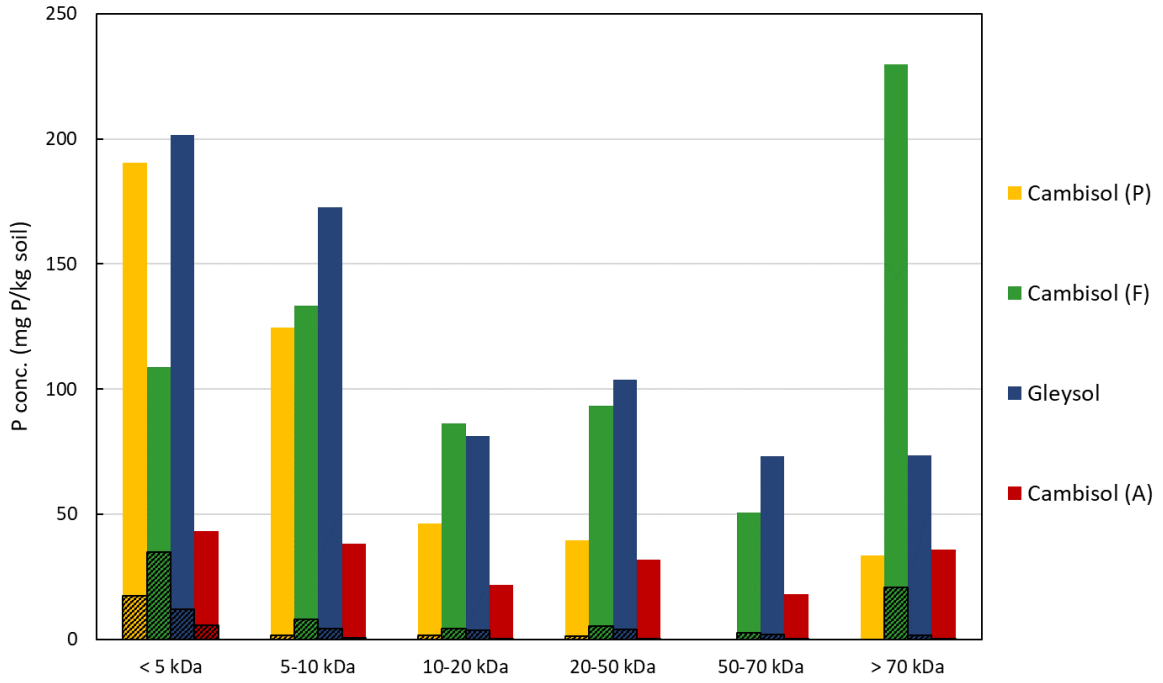


Figure SI-7. Molecular size distribution of total P measured in 0.25 M NaOH + 0.05 M EDTA soil extracts after passing through the Superdex™ 75 prep grade column. The proportion of molybdate reactive P to the total P measured by ICP-OES is indicated by the shaded region of the bars.

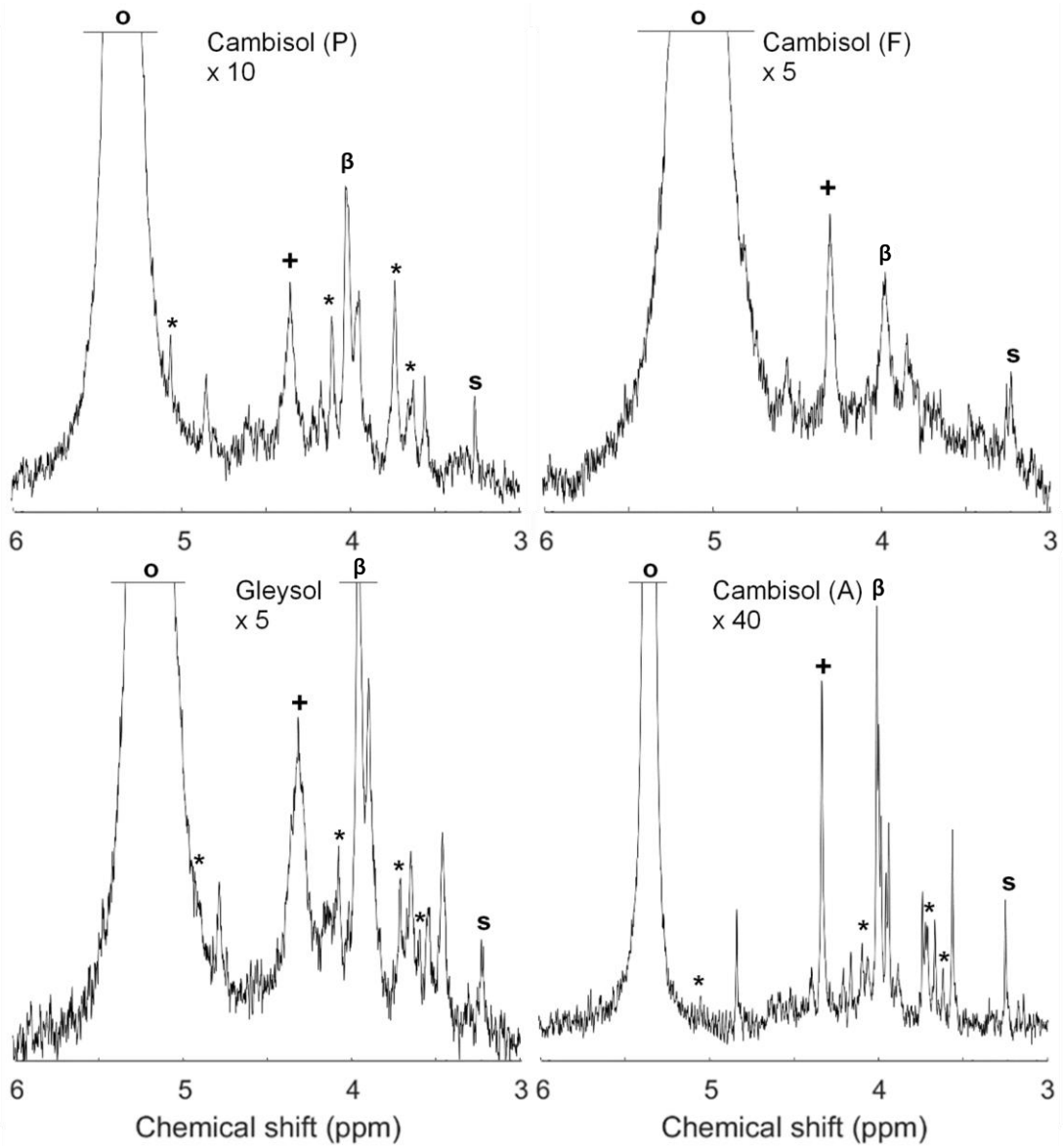


Figure SI-8. Solution ^{31}P nuclear magnetic resonance (NMR) spectra (500 MHz) of the combined orthophosphate (o) and phosphomonoester region on the fraction of 0.25 M NaOH + 0.05 M EDTA soil extracts (<5 kDa) removed by the desalting process (Sephadex™ G-25 gel media). The peaks of α-glycerophosphate/unknown high molecular weight P_{org} (McLaren et al., 2015b) (**+**), β-glycerophosphate (**β**) as well as *myo*- (*****) and *scyllo*- (**s**) IP_6 are marked. Signal intensities were normalised to the MDP peak intensity. The vertical axes were increased for improved visibility of spectral features, as indicated by a factor.

Table SI-1. Concentrations of total P (P_{tot}) determined by ICP-OES and main P classes (orthophosphate, phosphomonoester, phosphodiester and other P compounds) determined by solution ^{31}P NMR spectroscopy in mg P/kg_{soil} measured in the fraction (<5 kDa) of 0.25 M NaOH + 0.05 M EDTA soil extracts removed by the desalting process (Sephadex™ G-25 gel media). The percentage of the P class to the P_{tot} concentration is listed in parentheses.

	P_{tot} ICP-OES (mg P/kg_{soil})	**P_{tot} NMR (mg P/kg_{soil})**	**Ortho-P (mg P/kg_{soil})**	**P-monoester (mg P/kg_{soil})**	**P-diester (mg P/kg_{soil})**	**other P (mg P/kg_{soil})**
Cambisol (P)	255.8	254.4	231.2 (91)	17.9 (7)	0.5 (0)	4.8 (2)
Cambisol (F)	296.7	279.8	248.0 (89)	28.3 (10)	0.0 (0)	3.6 (1)
Gleysol	418.6	391.6	331.8 (85)	49.7 (13)	0.0 (0)	10.1 (3)
Cambisol (A)	115.6	101.9	90.3 (89)	9.2 (9)	0.0 (0)	2.5 (2)

Table SI-2. Iron, aluminium, manganese, potassium, and calcium concentrations in size fractions of soil extracts as determined by SEC and ICP-OES. Values below lower limit of detection are not listed. Value above upper limit of detection is marked with an asterisk.

	Fe (mg/kg_{soil})	Al (mg/kg_{soil})	Mn (mg/kg_{soil})	K (mg/kg_{soil})	Ca (mg/kg_{soil})
Cambisol (P)					
<5 kDa	13.3	37.2	1.9	124.9	157.1
5-10 kDa	3.7	11.4	0.7	35.9	58.3
10-20 kDa	3.3	-	-	29.8	4.6
20-50 kDa	5.1	-	-	38.1	-
50->70 kDa[4]	21.5	-	-	67.2	-
Cambisol (F)					
<5 kDa	973.0	462.0	27.0	63.7	268.1
5-10 kDa	203.2	151.7	9.8	31.5	92.4
10-20 kDa	69.8	23.2	0.4	31.6	5.1
20-50 kDa	238.5	49.7	1.5	39.6	1.3
50-70 kDa	287.0	58.8	3.0	15.1	0.8
>70 kDa	*1855.3	667.3	20.5	69.7	8.5
Gleysol					
<5 kDa	289.2	160.1	-	8.7	1245.9
5-10 kDa	6.6	24.3	-	-	165.6
10-20 kDa	15.0	-	-	-	-
20-50 kDa	70.5	14.8	-	-	-
50-70 kDa	144.6	21.0	-	-	-
>70 kDa	222.4	50.2	-	-	26.2
Cambisol (A)					
<5 kDa	22.6	-	-	-	2398.1
5-10 kDa	-	-	-	-	282.4
10-20 kDa	-	-	-	-	-
20-50 kDa	-	-	-	-	-
50-70 kDa	4.2	-	-	-	-
>70 kDa	49.4	10.4	-	-	-

[4] The 50-70 kDa and >70 kDa fractions of Cambisol (P) have been pooled together. The given values represent the concentrations measured in the pooled fraction.

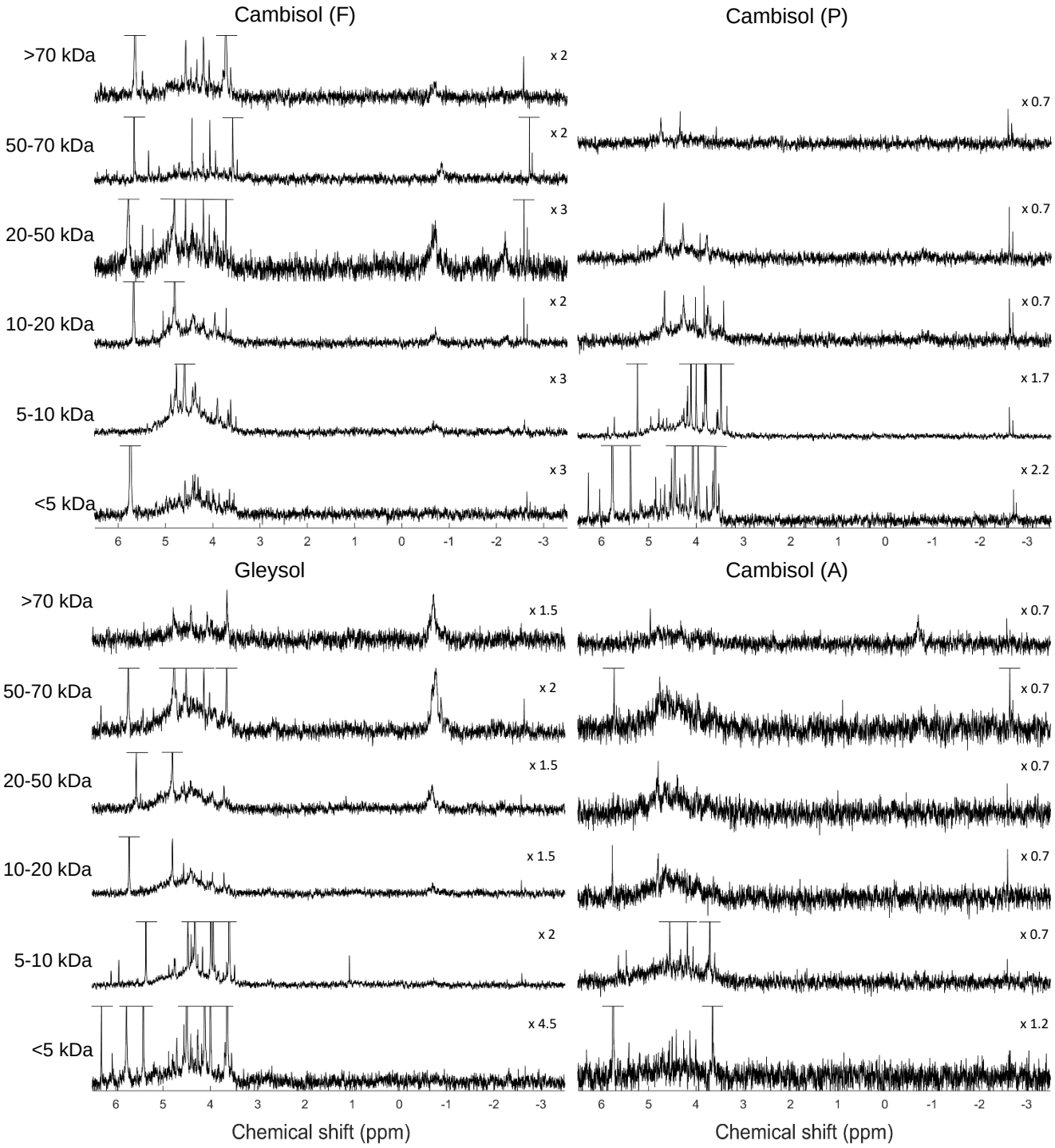


Figure SI-9. Solution ^{31}P nuclear magnetic resonance (NMR) spectra (500 MHz) of the phosphomono- and phosphodiester region on the molecular size fractions (Superdex™ 75 prep grade column) of 0.25 M NaOH + 0.05 M EDTA soil extracts (Cambisol (F), Cambisol (P), Gleysol and Cambisol (A)). Signal intensities were normalised to the MDP peak intensity. The vertical axes were scaled for improved visibility of spectral features, as indicated by a factor. The molecular size fractions >70 kDa and 50-70 kDa of the Cambisol (P) were pooled.

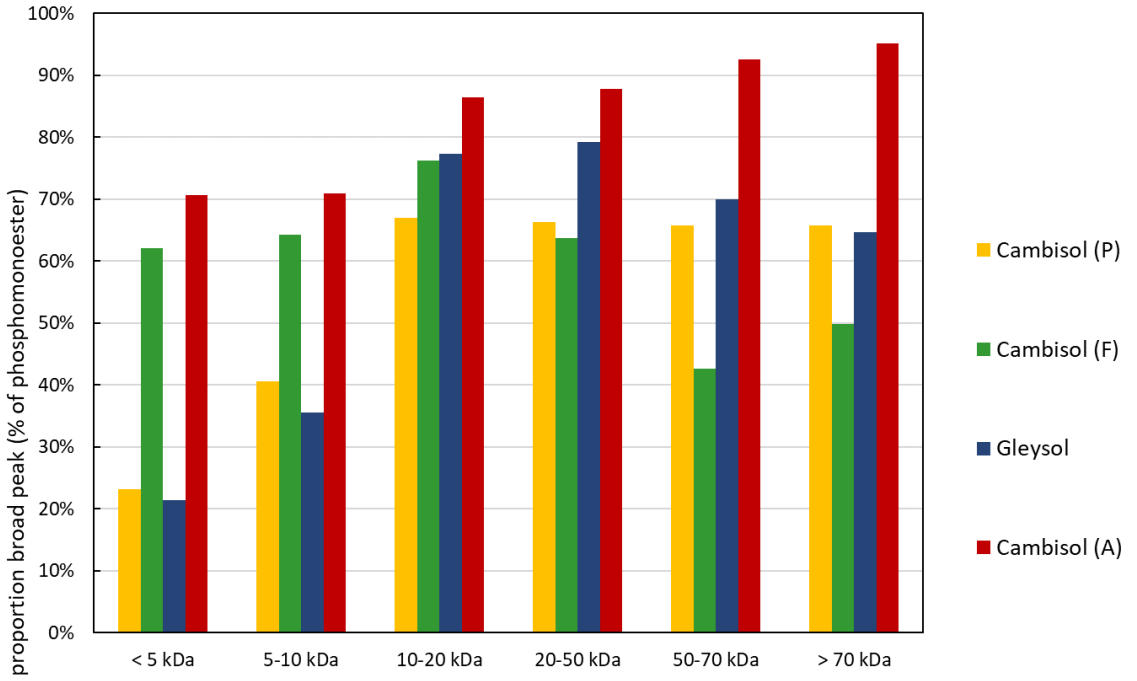


Figure SI-10. Relative concentrations of compounds causing the broad signals in the phosphomonoester region to the total concentrations of phosphomonoesters measured in the indicated molecular size fractions of solution ^{31}P NMR soil spectra.

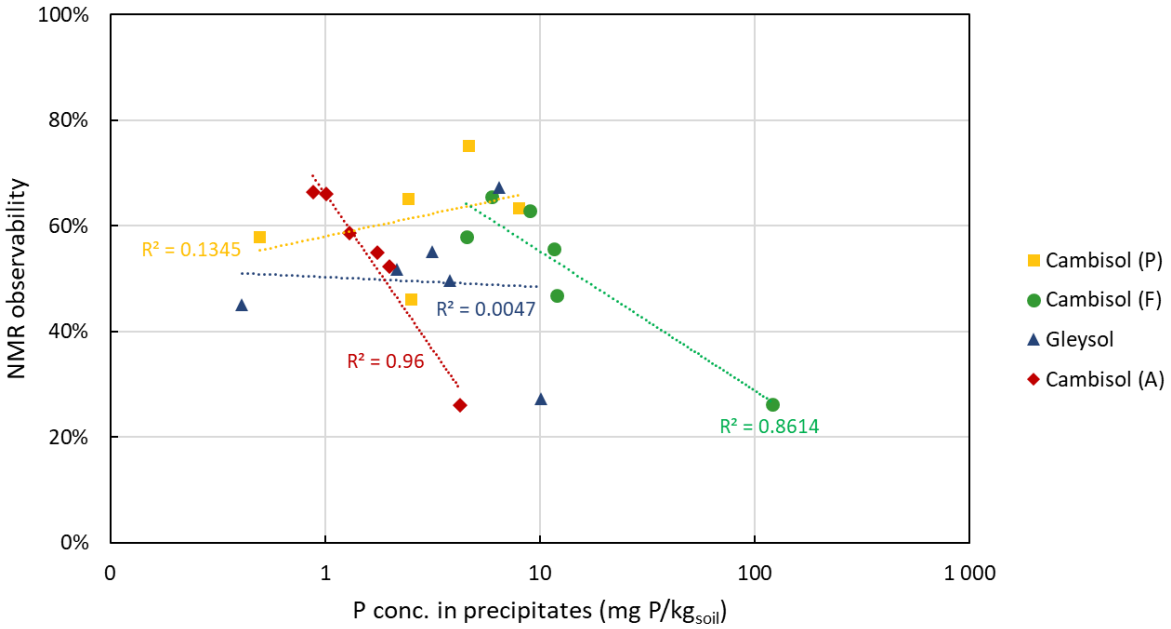


Figure SI-11. Correlation of P concentration measured in precipitates of molecular size fractions (mg P/kg$_{soil}$) and NMR observability of respective supernatants (% of P$_{tot}$ measured by ICP-OES before centrifugation). The x-axis and the fitted correlation curves are in logarithmic scale.

Chapter 4

Complex phosphomonoesters and inositol phosphates are associated to soil organic matter via multiple bonding types

Abstract

The majority of organic P (P_{org}) in soil is considered to be closely associated with the soil organic matter (SOM). The chemical nature of this pool of P_{org} is largely 'unresolved' and of apparent high molecular weight. In this study, we investigated the bonding types of unresolved P_{org} within the SOM structure using the Humeomics sequential chemical fractionation (SCF) procedure and nuclear magnetic resonance (NMR) spectroscopy. In summary, the SCF procedure extracted 118 % more P from a Gleysol soil compared to a single-step NaOH-EDTA extraction. Approx. 38 % of the extracted P_{org} was present in the form of the unresolved P_{org} pool represented by one to two underlying broad signals in the phosphomonoester region of solution ^{31}P NMR spectra. The Humeomics SCF revealed that 47 % of this unresolved P_{org} pool was associated with the SOM, mainly through ester bonds (40 %) and ether bonds (7 %). Another part of the unresolved P_{org} pool (30 %) appeared in the SOM fraction closely associated with the soil mineral phase. These findings provide further evidence that the unresolved P_{org} pool in soil is mostly comprised of a range of smaller P containing molecules associated with the SOM through ester linkages and not of a single macropolymeric structure. Furthermore, our results revealed that especially the pool of inositol phosphates (IP) associated with the soil mineral phase was highly underestimated using a single-step NaOH-EDTA extraction, which only extracted 58 % of total IP compared to the Humeomics sequential extraction. These findings on the associations and bonding types of the majority of P_{org} with the SOM will have major implications on our understanding of P_{org} stabilisation, transformation and cycling in soil systems. By knowing the bonding types, fertilisation and soil management strategies can be developed targeting these associations in order to enhance the liberation and mineralisation of the unresolved P_{org} pool for plant nutrition.

Introduction

Phosphorus (P) is an essential constituent of the energy storing compounds in cells and the genome of living organisms. Organic molecules containing P in soil, i.e. soil organic P (P_{org}), generally comprise between 20 % and 80 % of the total P in soil (Anderson, 1980). Soil P_{org} exists as P covalently bound to an organic moiety R (R–PO_3, phosphonates), via an ester linkage (R–O–PO_3, phosphomonoester), or via two ester linkages (R^1–O–PO_2–O–R^2, phosphodiesters) (Condron et al., 2005). In addition, soil P_{org} can include phosphate associated with soil organic matter (SOM), or condensed forms of P (e.g. polyphosphates) present within microorganisms (George et al., 2018; McLaren et al., 2020).

The chemical nature of the majority of P_{org} in soil remains unresolved (McLaren et al., 2020). This 'unresolved' pool of P_{org} can be accounted for as an underlying broad signal in the phosphomonoester region of ^{31}P NMR spectra on soil extracts (McLaren et al., 2015b; Reusser et al., 2020a). The molecules comprising the unresolved pool of P_{org} are of apparent large molecular weight, complex in structure, and resistant to enzymatic hydrolysis (He et al., 2006; Jarosch et al., 2015; McLaren et al., 2015b; McLaren et al., 2020). Furthermore, there is increasing evidence that this pool contains several components distributed along a continuum of molecular weights rather than a single polymeric structure (McLaren et al., 2019).

The occurrence of unresolved pool of P_{org} pool in high molecular size material of alkaline (NaOH) soil extracts suggests a close association with the SOM, because by definition, humic and fulvic acids are extracted with alkali from soil (Stevenson, 1994). For example, Ogner (1983) extracted humic acids from four Norwegian raw humus types using 0.5 M NaOH. The subsequent solution ^{31}P NMR spectroscopic analyses revealed a great diversity and abundance of phosphomonoesters in high molecular weight material of these humic acids. The close association of the unresolved P_{org} pool with the organic matter matrix was also hypothesised by Dougherty et al. (2007). The authors reported that an underlying broad signal in the phosphomonoester region of ^{31}P NMR spectra on soil extracts partially resisted an HF treatment whereas phytate was completely removed. Gerke (2010) reported that humic-metal-P associations can account for more than 50 % of P in soil solutions. The author described the association of P with the SOM not as orthophosphate being directly bound to the organic molecules, as it is the case for P_{org} in its closer definition, but as orthophosphate bound to metal cations, which are complexed with humic substances. In contrast, McLaren et al. (2015b) suggested that the unresolved P_{org} compounds are associated with the SOM by ester linkages due to the appearance of

underlying broad signals in the phosphomonoester region of ^{31}P NMR soil spectra. However, knowledge on the associations and bonding types of the unresolved P_{org} pool with the SOM and mineral phase is very limited.

Inositol phosphates (IP) are another major pool of P_{org} in soils. They are the predominant pool of P in seeds, and can accumulate in soil due to their high sorption affinity to soil constituents (Celi and Barberis, 2007). Concentrations of IP vary widely in soil ranging from less than 1 % up to more than 80 % of the total pool of P_{org} in soil (Cosgrove and Irving, 1980; Harrison, 1987; McDowell and Stewart, 2006; Smernik and Dougherty, 2007). The most abundant IP form in soil is *myo*-IP_6, commonly known as phytate, followed by the *scyllo* isomer of IP_6 (Cosgrove and Irving, 1980; Turner et al., 2002). The other two isomers of IP (*neo*- and *chiro*) are less abundant in soil. New findings of Reusser et al. (2020b) on solution ^{31}P NMR spectra of hypobromite oxidised soil extracts suggest that the diversity of IP in soils could be much higher than previously thought, including a variety of lower-order IP forms.

Pools of IP are largely considered to be associated with the mineral phase of soils, e.g. amorphous Al and Fe, clay minerals, or calcite (Ognalaga et al., 1994; Celi et al., 1999; Turner et al., 2002; Celi and Barberis, 2007). However, there is supporting evidence that IP can be associated to the organic phase of soils (Moyer and Thomas, 1970; Hong and Yamane, 1980; Borie et al., 1989). Furthermore, Jørgensen et al. (2015) indicated that the association of IP_6 with SOM can be important even in soils rich in amorphous Al and Fe. The authors reported that more than half of IP_6 was associated with the SOM in strongly weathered Oxisols and one fourth in organic rich grassland soils. Indeed, despite the apparent high abundance of these IP-SOM associations and the possible role of IP in the stabilisation of SOM (Tipping et al., 2016), the types of these associations and the formation mechanisms are largely unknown.

Soil organic matter consists of the total organic compounds in the soil except undecayed plant and animal residues as well as the living microbial biomass (Stevenson, 1994). Piccolo (2001) described SOM as an association of smaller molecules forming ‘suprastructures’ of apparent large molecular size. These smaller molecules are stabilised within the suprastructure by various bonding types, including van der Waals and hydrogen bonds (Piccolo, 2001). Nebbioso and Piccolo (2011) developed the ‘Humeomics’ sequential chemical fractionation (SCF) procedure to fractionate the SOM using different extractants without breaking any C–C bonds. In the Humeomics SCF procedure, extraction with a dichloromethane/methanol solution yields the ‘unbound’ SOM fraction, extraction

with a 12 % BF_3 in methanol solution the ‘weakly ester-bound’, extraction with a 0.5 M KOH in methanol solution the ‘strongly ester-bound’, and extraction with 47 % HI the ‘ether-bound’ fraction (Nebbioso and Piccolo, 2011). The residual fraction is assumed to be closely associated with the mineral phase. In summary, the unbound fraction consists of free molecules and molecules liberated from associations held together by weak dispersive forces (Piccolo et al., 2019). Subsequently, readily accessible ester linkages are transesterified, giving rise to the operationally defined ‘weakly’ ester-bound fraction. In a next step, ester linkages are cleaved which were not physically accessible during the first transesterification process, resulting in the ‘strongly’ ester-bound fraction. In the last step, ether linkages are cleaved, resulting in the ether-bound and residual fractions. Drosos et al. (2017) applied the SCF procedure to an agricultural sandy loam soil and reported the extraction of 235 % more soil organic C compared to a single-step alkaline (0.5 M NaOH + 0.1 M $Na_4P_2O_7$) extraction. The authors found that most C was extracted from the SOM suprastructure by breaking the weak (32 %) and strong (16 %) ester linkages. The ether-bound fraction was minor, comprising 5 % to total C. The residue after the Humeomics SCF contained 39 % of total C. This final residue is assumed to consist mostly of SOM bound to the soil mineral matrix via two processes (Drosos et al., 2017): 1) adsorption of hydrophobic components on aluminosilicate surfaces and 2) complex formation between mineral iron and oxygen containing hydrophilic groups.

George et al. (2018) declared that the interactions between P speciation, availability and SOM are of prime importance in order to evaluate the nutrient status in terrestrial ecosystems and improve land management practices. In this study, we combine the SOM sequential fractionation method of Humeomics (Nebbioso and Piccolo, 2011) with solution ^{31}P NMR spectroscopy, the most commonly used spectroscopic technique for P_{org} speciation (Cade-Menun, 2005; McLaren et al., 2020), in order to unravel the chemical nature and bonding type of soil P_{org} associated with the SOM. The aim of this study was to fractionate the complex structure of SOM into smaller components based on their bonding type with subsequent speciation of P_{org} compounds within the different fractions of Humeomics and residues.

We hypothesise that: 1) the Humeomics SCF procedure will extract more soil P compared to a single-step NaOH-EDTA extraction due to the stepwise ‘decoating’ of the SOM suprastructure; 2) subsequent fractionation of the SOM suprastructure causes a decrease in the pool of unresolved P_{org}, i.e. broad signals in the phosphomonoester region of solution ^{31}P NMR spectra; 3) the predominant bonding type for unresolved pools of P_{org} associated with large structures (suprastructures) of SOM is via ester linkages, and; 4) the residual fraction following SCF will be dominated by IP.

Materials and methods

Soil collection and preparation

A sample from a Gleysol (WRB, 2014) soil under pasture was collected in 2017 from the 0-10 cm layer in Switzerland. Detailed information on the location and soil properties can be found in Reusser et al. (2020a). The soil sample was sieved (<5 mm), dried for 5 days at 60°C, and then ground to powder using a mortar and pestle to pass through a 0.5 mm sieve.

Humeomics sequential chemical fractionation

The Humeomics SCF procedure was carried out on three analytical replicates of the Gleysol soil at the Institute of Agricultural Chemistry, University of Naples Federico II in Naples. The procedure is based on a slightly modified version of the method of Nebbioso and Piccolo (2011), which is summarised in Figure 1. If not otherwise stated, extraction of soils occurred at room temperature (21°C), centrifugation of extracts was carried out at 7500 rpm (8678 *g*) at 10°C and the supernatant passed through a Whatman® no. 41 filter paper.

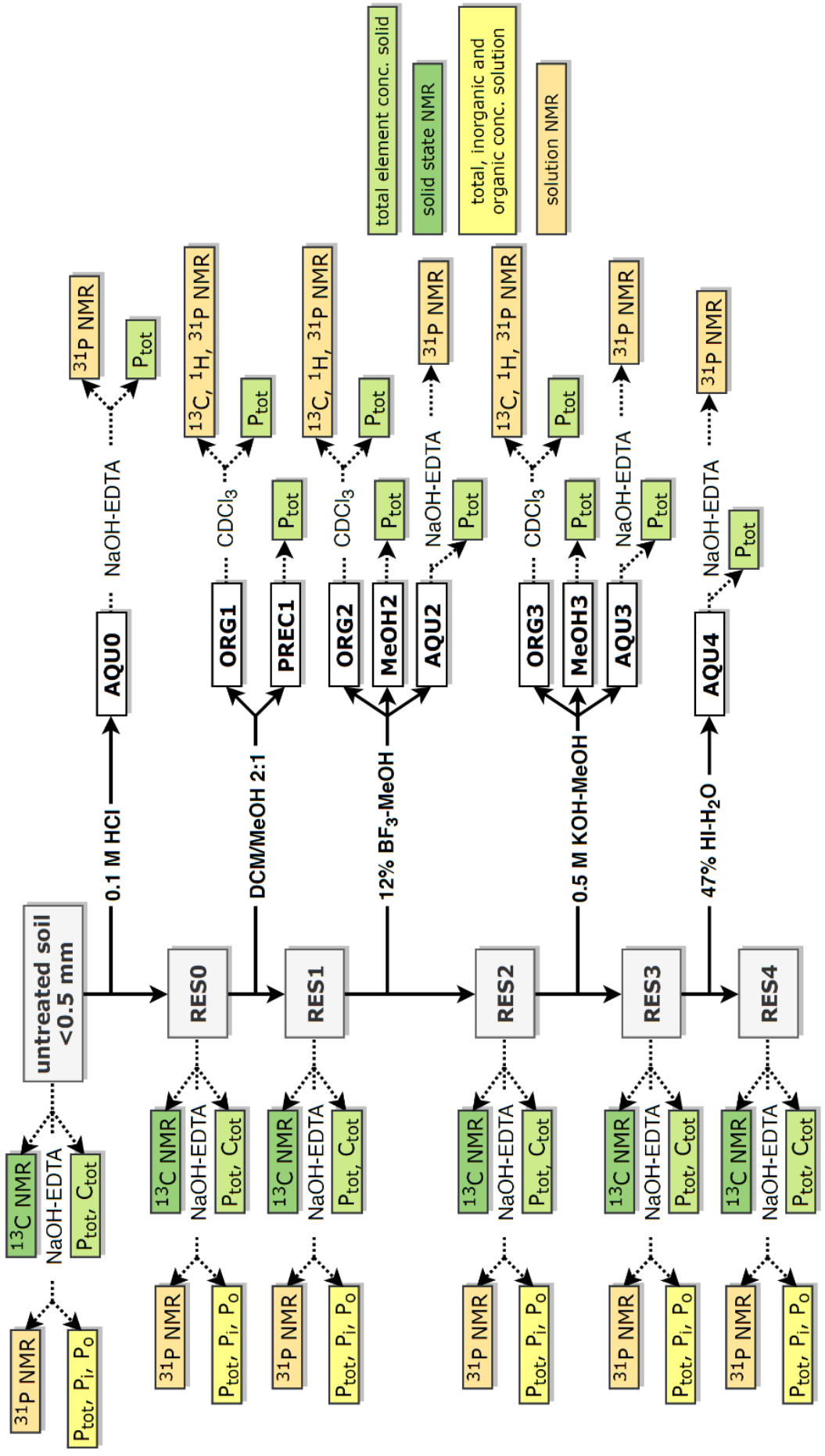


Figure 1. A summary of the Humeomics sequential chemical fractionation procedure, as modified from Nebbioso and Piccolo (2011), and the different measurements taken on all samples. The term 'RES' refers to the dried soil residue after the respective extraction step, 'AQU' to the aqueous phase, 'ORG' to the organic solvent phase, 'MeOH' to the methanol phase and 'PREC' to the precipitate after centrifugation of the respective extract.

Extraction with a 0.1 M HCl in water solution

Carbonates were removed by extracting 300 g of soil with 900 mL of a 0.1 M HCl solution for 3 h at 130 rpm. The soil extracts were then centrifuged for 30 min and the supernatant filtered. A 20 mL aliquot of the filtrate was collected and freeze-dried for further analyses (AQU0). The soil residue was suspended in 600 mL of H_2O, shaken manually, and the supernatant discarded following centrifugation for 15 min. The soil residue (RES0) was then dried at 40°C for 48 h and freeze-dried.

Extraction with a dichloromethane-methanol solution

The ORG1 fraction was obtained by extracting 40 g of RES0 with 120 mL of a dichloromethane 2:1 methanol (DCM/MeOH) solution (v/v) for 12 h at 130 rpm. The extracts were then centrifuged for 20 min and the supernatant filtered into a round-bottom flask. The extraction procedure was repeated but with 60 mL of solution for 1 h. The volume of the combined extracts was reduced to about 10 mL by rotary evaporation, and subsequently centrifuged at 12000 rpm (13870 *g*) for 20 min in order to separate the clay minerals (PREC1). The remaining supernatant (ORG1) was further rotary evaporated and dried under a N_2 gas flow. The soil residue (RES1) was then dried at 40°C for 72 h.

Extraction with a 12 % BF_3 in methanol solution

In the next step, transesterification was carried out using a 12 % (1.5 M) BF_3 in methanol solution. In this transesterification reaction, the BF_3 served as a lewis acid catalyst. It catalyses the exchange of an organic group of an ester with the organic group of an alcohol, i.e. methanol (McCarthy et al., 2020). The transesterification reaction mechanism on the example of triglyceride is illustrated in Figure SI-1 in the supporting information.

Briefly, 30 g of RES1 was suspended in 40 mL solution of 12 % BF_3 in methanol using Teflon tubes, and a constant temperature of 77°C for 16 h using an oven. The extraction procedure was repeated twice but with 90 mL and 40 mL of solution for 12 h and 7 h, respectively. The combined extracts were then centrifuged for 20 min and the supernatant filtered into a round-bottom flask. The soil residue was suspended sequentially five times in 30 mL of methanol and shaken manually. Each extract was centrifuged for 20 min and the supernatant filtered into the round-bottom flask containing the BF_3 extract. Furthermore, the soil residue was resuspended again twice but in 20 mL of H_2O and shaken manually. Each extract was centrifuged for 20 min and the supernatant filtered into

the round-bottom flask containing the BF_3 extract. This extract was then rotary-evaporated and transferred to a separation funnel, in order to obtain an organic phase (ORG2), a methanol phase (MeOH2) and an aqueous phase (AQU2). This procedure involved adding a total of 110 mL of H_2O and 250 mL of chloroform to the separation funnel. The ORG2 fraction was dried with anhydrous sodium sulphate, filtered, rotary evaporated, and then dried under a N_2 gas flow. Similarly, the MeOH2 fraction was dried with anhydrous sodium sulphate following centrifugation for 20 min, filtered, rotary evaporated, and then dried under a N_2 gas flow. The precipitate after centrifugation was combined with the soil residue. The AQU2 fraction was filtered and then rotary evaporated. Replicates 1 and 2 of AQU2 were dialysed prior to freeze-drying by ultrafiltration with a 500 Da cut-off membrane against H_2O in order to remove ions, which could otherwise interfere with the analysis of C using GC-MS (not part of this paper). The water was changed twice a day until the electrical conductivity was below 1 µS/cm. Replicate 3 of AQU2 was directly freeze-dried for subsequent analysis of total P (P_{tot}). The soil residue (RES2) was then dried at 40°C for 144 h.

Extraction with a 0.5 M KOH in methanol solution

The soil residue RES2 was further fractionated by extracting 20 g with 30 mL of 0.5 M KOH in methanol, which was stirred under a N_2 atmosphere and reflux at a constant temperature of 80°C for 2 h. The soil residue was suspended sequentially three times in acidified methanol (pH=2, HCl) and shaken manually. Each extract was centrifuged for 20 min and the supernatant filtered and added to the KOH extract. Furthermore, the soil residue was washed again but with a non-acidified methanol solution, and processed as previously described. The combined extracts were then centrifuged, filtered, rotary evaporated, and transferred to a separation funnel, in order to obtain an organic phase (ORG3), a methanol phase (MeOH3) and an aqueous phase (AQU3). This involved adding a total of 250 mL of dichloromethane, 20 mL of methanol, and 80 mL of H_2O to the separation funnel. The ORG3, MeOH3, and AQU3 fractions were processed the same as that previously described. The soil residue (RES3) was then dried at 40°C for 48 h.

Extraction with a 47 % HI in water solution

The last Humeomics SCF fraction was obtained by suspending 15 g of RES3 in a 20 mL solution of 47 % HI in H_2O using Teflon tubes, and a constant temperature of 77°C for 48 h using an oven. After cooling, the extract was centrifuged for 20 min and filtered into a plastic beaker. The soil residue was washed sequentially eleven times with 20 mL of H_2O

for 30 min at 130 rpm. Each extract was centrifuged for 20 min and the supernatant filtered and added to the HI extract. However, approximately 200 mg of sodium carbonate was added to the eleventh H_2O extraction in order to increase the pH of the soil to above 5. Sodium carbonate was added stepwise under stirring to the soil extract in the plastic beaker until a pH of 7 was reached. The I_2 was neutralised with a stepwise addition of sodium thiosulfate (pentahydrate) until a colour change from reddish brown to grey brown was observed. Replicates 1 and 2 of AQU4 were dialysed prior to freeze-drying as previously described. Replicate 3 of AQU4 was directly freeze-dried for subsequent analysis of P_{tot}. The soil residue (RES4) was then dried at 40°C for at least 72 h.

Concentrations of total phosphorus in soil

Concentrations of P_{tot} in soil residues following each extraction step (UN, RES0, RES1, RES2, RES3, and RES4) were determined using microwave acid digestion. Briefly, 200 mg of ground soil was digested in 3 mL of a 65 % nitric acid solution and 1 mL of a 37 % hydrochloric acid using a turboWAVE® MRT microwave digestion system (MILESTONE Srl, Sorisole, Italy) at 250°C for 35 min (Fioroto et al., 2017). The digest was then made up to volume with H_2O and subsequently analysed for P using the malachite green method of Ohno and Zibilske (1991).

Concentrations of P_{tot} in the rotary evaporated fractions of the SCF were measured using the aforementioned microwave acid digestion method, except about 100 mg of dried material was digested in 2 mL of a 65 % nitric acid solution.

The proportion of P_{tot} extracted by the Humeomics SCF procedure ($P_{tot,H}$) to P_{tot} in the untreated soil was calculated using Equation 1:

$$P_{tot,H}{:}\,P_{tot,UT,MD}\ (\%) = \frac{P_{ORG1}+P_{ORG2}+P_{ORG3}+P_{AQU0}+P_{AQU1}+P_{AQU2}+P_{AQU3}+P_{AQU4}+P_{PREC1}+P_{MeOH2}+P_{MeOH3}}{P_{tot,UT,MD}} * 100 \quad (1)$$

, where $P_{tot,UT,MD}$ is the total P concentration (mg P/kg_{soil}) in the untreated soil, and P_{ORG1}-P_{MeOH3} are the P concentrations (mg P/kg_{soil}) in the respective dried Humeomics SCF extracts. The concentrations were determined using microwave acid digestion with subsequent malachite green measurements of the diluted digests.

NaOH-EDTA extractable phosphorus

Concentrations of NaOH-EDTA extractable P in the untreated soil (UT) and soil residues following each extraction step (RES0, RES1, RES2, RES3, and RES4) were measured using the method of Cade-Menun et al. (2002). The soil was extracted with 0.25 M NaOH + 0.05 M EDTA at a 1:10 soil-to-solution ratio for 16 h at 150 rpm. Soil extracts were centrifuged for 15 min and filtered prior to analysis of P_{tot} by ICP-OES, and molybdate reactive P (MRP) by the malachite green method of Ohno and Zibilske (1991). The difference between P_{tot} and MRP is MUP, and is considered to be largely organic P. An aliquot of all NaOH-EDTA filtrate for Replicate 3 was frozen at -80°C and then freeze-dried for NMR analysis.

The proportion of P_{tot} extracted by the Humeomics SCF procedure to P_{tot} measured in the NaOH-EDTA extract of the untreated soil was calculated based on Equation 1, but by replacing $P_{tot,UT,MD}$ with the measured P_{tot} concentration in the NaOH-EDTA extract ($P_{tot,UT,NE}$, mg P/kg_{soil}).

The recovery (*R*) of P_{tot} extracted by the Humeomics SCF procedure with subsequent NaOH-EDTA extraction of the final residue compared to P_{tot} measured in the NaOH-EDTA extract of the untreated soil was calculated using Equation 2:

$$R\ (\%) = \frac{P_{ORG1}+P_{ORG2}+P_{ORG3}+P_{AQU0}+P_{AQU1}+P_{AQU2}+P_{AQU3}+P_{AQU4}+P_{PREC1}+P_{MeOH2}+P_{MeOH3}+P_{tot,RES4,NE}}{P_{tot,UT,NE}} * 100 \qquad (2)$$

, where $P_{tot,UT,NE}$ is the total P concentration (mg P/kg_{soil}) in the NaOH-EDTA extract of the untreated soil, P_{ORG1}-P_{MeOH3} are the P concentrations (mg P/kg_{soil}) in the respective dried Humeomics SCF extracts determined by microwave digestion, and $P_{tot,RES4,NE}$ is the total P concentration (mg P/kg_{soil}) in the NaOH-EDTA extract of the final residue RES4.

Sample preparation for solution ^{31}P NMR spectroscopy

Sample preparation of freeze-dried material for solution ^{31}P NMR spectroscopy is similar to that reported in Reusser et al. (2020a). Briefly, 40 mg (UT, RES1, and RES2) or 80 mg (RES0, RES3, and RES4) of freeze-dried material was dissolved in 600 µL of NaOH-EDTA solution. The solution was then shaken and allowed to rest for at least 2 h, prior to centrifugation at 10000 rpm (10621 *g*) for 15 min. A 500 µL aliquot of the supernatant was transferred to a 1.5 mL microcentrifuge tube, which was then spiked with a 25 µL aliquot

of 0.03 M methylenediphosphonic acid (MDP) standard in D_2O, and a 25 µL aliquot of sodium deuteroxide (NaOD) at 40 % (w/w) in D_2O. The solution was then transferred to a 5 mm NMR tube for analysis.

Sample preparation of the AQU0, AQU1, AQU3, and AQU4 freeze-dried materials was similar to that described above, except at a different material to solution ratio. The freeze-dried material of AQU0 was dissolved in 2 mL of NaOH-EDTA, and after centrifugation at 10000 rpm (10621 *g*) for 15 min, a 600 µL aliquot of the supernatant was transferred to a 1.5 mL microcentrifuge tube and spiked with 25 µL of MDP in D_2O and 30 µL of NaOD. The freeze-dried material of AQU2 still contained moisture due to its high ion content. Briefly, a subsample of 40 mL was freeze-dried and dissolved in NaOH-EDTA to make a final volume of 12 mL. Several different dilution factors and additions of NaOD were tested but resulted in low-quality NMR spectra. Consequently, the concentrated ion solution was removed using ultrafiltration with a 1 kDa cut-off at 5000 rpm for 20 min. This resulted in an aliquot of 200 µL for the 1 kDa retentate, which was subsequently diluted with 550 µL of NaOD and centrifuged. A 500 µL aliquot was transferred to a 1.5 mL microcentrifuge tube and spiked with 25 µL of MDP, which then produced a high-quality NMR spectrum. The freeze-dried material of AQU3 (80 mg) and AQU4 (120 mg) was dissolved in 600 µL of NaOH-EDTA, and after centrifugation, a 500 µL aliquot of the supernatant was transferred to a 1.5 mL microcentrifuge tube and spiked with 25 µL of MDP and 30 µL of NaOD. All solutions were then transferred to a 5 mm NMR tube for analysis.

Sample preparation of the ORG1, ORG2, and ORG3 freeze-dried materials was similar to that described above, except 750 µL of deuterated chloroform ($CDCl_3$) was used as the dissolving agent. After centrifugation, a 500 µL aliquot was transferred to a 1.5 mL microcentrifuge tube and spiked with 25 µL of a 0.03 M triphenylphosphate (TPP) in $CDCl_3$. All solutions were then transferred to a 5 mm NMR tube for analysis.

Solution ^{31}P NMR spectroscopy and processing of spectra

All solution ^{31}P NMR analyses were carried out at the NMR facility of the Laboratory of Inorganic Chemistry (Hönggerberg, ETH Zürich). All samples were analysed using a Bruker Avance III HD 500 MHz NMR spectrometer equipped with a 5 mm liquid-state Prodigy™ CryoProbe (Bruker Corporation; Billerica, MA). The ^{31}P frequency was set to 202.5 MHz with inverse gated broadband proton decoupling and 90° pulses (duration of 12 µs). The longitudinal relaxation time (T_1) of each sample was assessed in advance by an inversion recovery experiment. Detailed description of this experiment can be found in

Reusser et al. (2020a). Recycle delays were then calculated by multiplying the longest T_1 value of the inversion recovery experiment times by 5 in order to achieve a full recovery (99.33 %) of point magnetisation (Claridge, 2016b). This resulted in recycle delays ranging from 5 to 35 s for the NaOH-EDTA extracts and 5 to 14 s for the $CDCl_3$ extracts. An average of 4096 scans for each NMR spectrum was acquired for the NaOH-EDTA and $CDCl_3$ extracts.

Solution ^{31}P NMR spectra were Fourier transformed, phase corrected, and baseline adjusted using the TopSpin® software of Bruker (Version 4.1.0, Bruker Corporation; Billerica, MA). Line broadening was set to 0.6 Hz for all spectra. Quantification of NMR signal involved comparing the integral region of the added P standard (MDP or TPP) of known P concentration, with all other NMR signals which are directly relative to each other. Integral regions included: 1) phosphonates (δ 19.9 to 13.3 ppm), 2) the combined orthophosphate and phosphomonoester region (δ 6.2 to 2.9 ppm), 3) phosphodiester (δ 2.8 to -2.0 ppm), 4) pyrophosphates (δ -4.8 to -5.5 ppm) and 4) polyphosphates (δ -17.1 to -20.6 ppm). Spectral deconvolution fitting (SDF) was applied to the orthophosphate and phosphomonoester region due to overlapping NMR signals in this region, as described in Reusser et al. (2020a). Peak assignments were based on spiking the extracts with standard solutions of the respective P_{org} compound and comparing the chemical shift with reported values of known compounds (Doolette et al., 2009; Cade-Menun, 2015; Reusser et al., 2020b).

The proportion of a specific P compound removed by the Humeomics SCF from the soil residue ($P_{removed}$) was calculated using Equation 3:

$$P_{removed}\ (\%) = \frac{P\ conc.\ RES_x - P\ conc.RES_{x+1}}{P\ conc.RES_x} * 100 \quad (3)$$

, where 'P conc. RES_x' is the concentration of the respective P compound (mg P/kg_{soil}) determined by NMR in the NaOH-EDTA extract of the soil residue before the Humeomics SCF extraction step (RES_x), and 'P conc. RES_{x+1}' is the concentration of the respective P compound (mg P/kg_{soil}) in the soil residue after the Humeomics SCF extraction step (RES_{x+1}).

Concentrations of total carbon and total nitrogen

Concentrations of total C and total N in the soil residues following each extraction step (RES0, RES1, RES2, RES3, and RES4) were determined using combustion (vario PYRO

cube®, Elementar Analysesysteme GmbH). About 5 mg of soil was weighed into tin foil capsules, which were then placed into the C and N analyser.

Solid-state CPMAS ^{13}C NMR spectroscopy

All solid-state cross polarisation magic angle spinning (CPMAS) ^{13}C NMR analyses were carried out at the Interdepartmental Research Centre on NMR for the Environment, Agro-food and New Materials (CERMANU) of the University of Naples Federico II in Portici, Italy. All samples were analysed using a Bruker Avance 300 MHz wide-bore NMR spectrometer equipped with a CPMAS probe. Soil was packed into a 4 mm Zirconia rotor fitted with Kel-F end-caps, which was spun at 13000 Hz. The ^{13}C frequency was set to 75.47 MHz and spectra were obtained using a ramped-amplitude CP pulse sequence, in which the ^{1}H spin lock power was varied linearly during the contact time. Each NMR spectrum was acquired using a 1 ms contact time, a recycle delay of 2 s, and an average of 32000 scans.

Solid-state ^{13}C NMR spectra were Fourier transformed, phase corrected, and baseline adjusted using the TopSpin® software of Bruker (Version 4.1.0, Bruker Corporation; Billerica, MA). Line broadening was set to 200 Hz for all spectra. Integral regions were measured and peak assignments were carried out according to Stevenson (1994) and Spaccini et al. (2006). This included: 1) 0-50 ppm, alkyl-C (e.g. alkanes and fatty acids); 2) 50-85 ppm, C–O and C–N in ethers, esters, carbohydrates and amines (e.g. amino acid-peptide- and protein-C); 3) 85-105 ppm, aliphatic C–O (carbohydrates); 4) 105-160 ppm, aromatic C with phenolic C between 150-160 ppm, and; 5) 160-200 ppm carbonyl-C. Integral regions 1) and 4) are considered to represent mostly hydrophobic compounds, whereas those of 2), 3), and 5) are considered mostly to comprise of hydrophilic compounds (Spaccini et al., 2006). Furthermore, the alkyl-C region of 1) can also be referred to the aliphatic region, whereas regions 2) and 3) are also known as the *O*-alkyl region (Simpson and Simpson, 2017).

Statistical analyses and graphics

Calculations of replicate averages, standard deviations and correlation coefficients were carried out using Microsoft® Excel 2016. One-way ANOVA with subsequent multi-comparison of mean values were processed in MATLAB R2017a (©The MathWorks, Inc.). The multi-comparison of means calculation was based on the Tukey's honestly significant difference procedure using the Studentised range distribution (Hochberg and Tamhane, 1987; Milliken and Johnson, 2009). All graphics of solution ^{31}P NMR spectra were created

in MATLAB R2017a. All solution ^{31}P NMR spectra were normalised to the peak maximum of the MDP peak. All other graphics, including solid-state ^{13}C NMR spectra, were created in Microsoft® Excel 2016.

Results

Carbon

Concentrations of total carbon in soil Humeomics residues

The removal of carbonates with 0.1 M HCl caused a significant decrease in total soil C of 24 g C/kg_{soil} (15 %) (Table 1, Figure SI-2). Following this, the SCF procedure extracted on average 48 % of the total C (C_{tot}) when compared to the UT soil. The majority (44 %) of C was removed by the BF_3 in methanol extraction. No significant differences in C_{tot} were found between RES0 and RES1 (DCM/MeOH extraction) and between RES2 and RES3 (0.5 M KOH in methanol extraction). There was a significant increase in C_{tot} of 10 g C/kg_{soil} from RES3 to RES4, following the 47 % HI in H_2O extraction step and subsequent neutralisation of soil acidity using sodium carbonate.

Chemical nature of carbon in soil Humeomics residues

The chemical nature of C in the UT soil and soil residues following each step of the SCF procedure was assessed by solid-state CPMAS ^{13}C NMR spectroscopy (Figure 2). There was no difference in the chemical nature of C in the UT soil and the soil residue following 0.1 M HCl treatment (RES0). However, NMR signals in the alkyl-C region slightly decreased following the DCM/MeOH treatment (RES1) compared to RES0. Most notably, there was a large decrease of NMR signals in the alkyl-C region and the region representing C–O and C–N bonds in ethers, esters, carbohydrates and amines following the 12 % BF_3 in methanol treatment (RES2) compared to RES1. The 12 % BF_3 in methanol treatment also revealed the presence of a peak (δ 50-60 ppm), which was assigned to N-alkyl and methoxyl-C (Stevenson, 1994). Furthermore, the 12 % BF_3 in methanol treatment also removed some carbonyl-C (region 5) and aliphatic C–O (region 3), and N-alkyl, methoxyl C and aliphatic C–O (region 2). The 0.5 M KOH in methanol treatment resulted in a large decrease of NMR signals in the aliphatic C–O and N-alkyl, methoxyl C and aliphatic C–O regions of RES3 compared to RES2. The chemical nature of C in RES4 was largely hydrophobic alkyl-C (region 1) and aromatic C (region 4) following 47 % HI in H_2O.

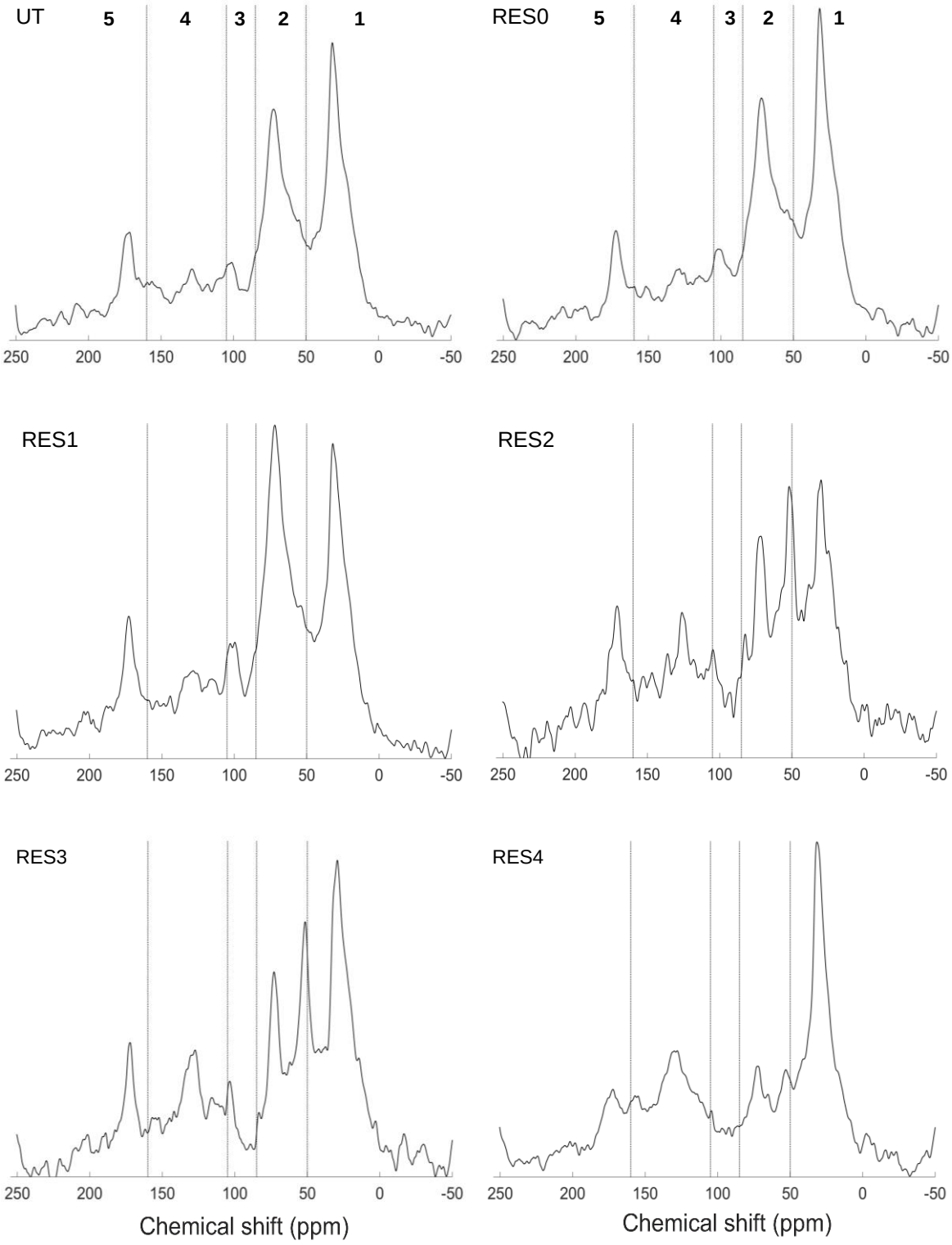


Figure 2. Solid-state ^{13}C cross polarisation magic angle spinning (CPMAS) NMR spectra on the untreated soil (UT) and the soil residues (RES0, RES1, RES2, RES3, and RES4) after each step of the Humeomics sequential chemical fractionation procedure. Chemical shift ranges of different C species are separated by dotted lines and labelled with numbers: alkyl C (1), C–O and C–N in ethers, esters, carbohydrates and amines (2), aliphatic C–O as in carbohydrates (3), aromatic C (4) and carbonyl C (5).

Phosphorus

Concentrations of total phosphorus in soil Humeomics extracts

The Humeomics SCF procedure extracted a total of 1158 mg P/kg_{soil} (Table SI-1 in the supporting information), which accounted for 64 % of P_{tot} in the Gleysol soil, as determined by microwave acid digestion (Table 1). Furthermore, the SCF procedure extracted 69 % of the total P pool determined in a single-step NaOH-EDTA extraction. Concentrations of P_{tot} in each of the Humeomics extracts were low, except for the 47 % HI in H_2O extract (AQU4), which was 1040 mg P/kg_{soil} and comprised 90 % of extractable P by Humeomics (Table SI-1). Negligible concentrations of P were present in the organic phases of Humeomics extracts.

Table 1. Concentrations of total P (P_{tot}) as measured by HNO_3-HCl microwave digestion; P_{tot}, organic P (P_{org}) and inorganic P (P_{inorg}) in 0.25 M NaOH + 0.05 M EDTA extracts; total C (C_{tot}) and total N (N_{tot}) as measured by combustion. Standard deviations of n=3 replicates are shown in brackets.

Measure	UT	RES0	RES1	RES2	RES3	RES4
P_{tot} digestion	1796	1579	1561	995	934	437
(mg P/kg_{soil})	(43)	(28)	(33)	(105)	(40)	(60)
P_{tot} NaOH-EDTA extract	1682	1518	1555	648	661	621
(mg P/kg_{soil})	-	-	(44)	(35)	(80)	(14)
P_{org} NaOH-EDTA extract	972	829	865	450	450	462
(mg P/kg_{soil})	-	-	(50)	(24)	(53)	(5)
P_{inorg} NaOH-EDTA extract	709	689	690	199	211	160
(mg P/kg_{soil})	-	-	(10)	(19)	(28)	(9)
C_{tot} (g C/kg_{soil})	159	134	131	73	70	80
	(0.4)	(0.9)	(0.6)	(3.9)	(1.6)	(0.9)
N_{tot} (g N/kg_{soil})	11	9	10	5	5	3
	(0.1)	(0.1)	(0.1)	(0.4)	(0.1)	(0.2)
C_{tot} : P_{org} (-)	163	162	149	172	151	175

Chemical nature of phosphorus in soil Humeomics extracts

Solution ^{31}P NMR spectroscopy revealed that the majority of P in soil Humeomics extracts occurred as orthophosphate in the aqueous phases of the 12 % BF_3 in methanol (AQU2) and 47 % HI in H_2O (AQU4) extracts (Table SI-1 and Figure SI-3), which comprised more than 90 % of extractable P by Humeomics. In general, phosphomonoesters were the dominant pool of organic P in Humeomics extracts, particularly in the aqueous phases of 0.5 M HCl (AQU0) and 47 % HI in H_2O (AQU4) extracts, albeit at low concentrations (<5 mg P/kg_{soil}). Several sharp (and broad) unidentified signals were present in the phosphomonoester region across all aqueous phases.

Concentrations of phosphorus pools in soil Humeomics residues

In total, 1359 mg P/kg_{soil} of P_{tot} in soil was removed by the Humeomics sequential chemical fractionation procedure, resulting in 76 % of P_{tot} in the untreated soil (Table 1). Most P was extracted by the 12 % BF_3 solution, leading to a decrease of the P_{tot} concentrations in the soil residues of 36 % (Table 1). Another significant pool of P (on average 497 mg P/kg_{soil}) was removed by the 47 % HI extraction. In contrast, the other Humeomics extraction steps did not decrease the P_{tot} contents in the soil residues in a significant amount except for the HCl removal of carbonates (Figure SI-5). The remaining soil residue RES4 contained 24 % of P_{tot} in the untreated soil. In summary, 118 % of P_{tot} could be extracted with the Humeomics sequential fractionation method with subsequent NaOH-EDTA extraction of RES4 compared to a single-step NaOH-EDTA extraction.

Similar to the P_{tot} results, the 12 % BF_3 extraction led to a significant decrease of the NaOH-EDTA extractable P pool of 58 % (Table 1, Figure SI-6). This decrease was more pronounced for P_{inorg}, with a reduction of 71 % (on average 491 mg P/kg_{soil}), compared to P_{org}, with a reduction of 48 % (on average 415 mg P/kg_{soil}). Consequently, the proportion of P_{org} to total NaOH-EDTA extractable P increased from 58 % in the untreated soil to 74 % in RES4 after the Humeomics sequential extraction method. In addition, slightly more P_{org} was extracted in RES4 compared to RES3.

Chemical nature of phosphorus in soil Humeomics residues

The chemical nature of P in the UT soil and soil residues following each step of the Humeomics SCF procedure was assessed by solution ^{31}P NMR spectroscopy on NaOH-EDTA extracts (Figure 3). The whole spectra of UT and RES4 as well as a detailed peak

assignment in the combined phosphomonoester and orthophosphate region can be found in Figures SI-7. SI-8 and SI-9.

In general, most P was detected in the combined ortho-P + phosphomonoester region, followed by phosphodiesters and pyrophosphates (Table 2). Phosphonates and polyphosphates were only measured in trace amounts. The phosphomonoester region comprised of the dominant ortho-P peak, of on average 1.5 underlying broad signals and 31 sharp peaks. The 0.5 M KOH extraction resulted in an increase of detectable sharp peaks in the phosphomonoester region of the soil residues, i.e. from 24 sharp peaks in RES2 to 39 resp. 40 in RES3 resp. RES4. The distribution of specific P compounds along the Humeomics fractions are presented in the following sections in more detail.

Untreated soil

The majority of P in the untreated soil was present in the form of orthophosphate (46 % of P_{tot}), inositol phosphates (22 % of P_{tot}) and one underlying broad signal in the phosphomonoester region (19 % of P_{tot}). The remaining 13 % of P_{tot} as measured by NMR included forms of phosphodiesters, other phosphomonoesters, pyrophosphates, polyphosphates and phosphonates.

The most abundant form of IP was the *myo*-isomer of IP_6 (51 % of total IP). In addition, the two inositol pentakisphosphates *myo*-(1,2,4,5,6)-IP_5 and *scyllo*-IP_5 were detected.

0.1 M HCl extraction

The extraction with 0.1 M HCl did not cause a considerable change of the P composition in the NMR spectrum on the soil residue RES0. However, regarding concentrations, 3 % of the unresolved P_{org} pool and 27 % of the total IP pool were apparently removed by the 0.1 M HCl extraction.

Dichloromethane-methanol extraction

Similar to the previous extraction step, the organic solvent extraction did not cause a considerable change of the P composition, except that 20 % of the unresolved P_{org} pool was removed. The remaining unresolved P_{org} pool was represented by two instead of one underlying broad signal in the phosphomonoester region of RES 1. In contrast to this pool, the organic solvent extraction resulted in a higher NaOH-EDTA extractability of IP and phosphodiesters in RES1.

12 % BF_3 extraction

The 12 % BF_3 extraction did not only remove most of P from the soil, but also markedly changed the molecular composition of the remaining P pool in the residue RES2. Compared to RES1, 71 % of orthophosphate, 35 % of the unresolved P_{org} pool, 87 % of IP, and 80 % of phosphodiesters were removed by the 12 % BF_3 extractant. Furthermore, pyrophosphates were completely removed.

This extraction step gave rise to a sharp peak at δ 4.79 ppm, which was attributed to glucose-6-phosphate based on spiking the AQU3 extract with a standard solution. Contrasting to all other P compounds, the concentration of glucose-6-phosphate increased markedly (4-times) in the NaOH-EDTA extract of RES2 compared to RES1.

0.5 M KOH extraction

The subsequent 0.5 M KOH extraction caused again a 26 % decrease of the unresolved P_{org} pool represented by one underlying broad signal in the phosphomonoester region of RES3. However, the extraction did not much affect *myo*-IP_6 but decreased the *scyllo*-IP_6 concentration by 40 % compared to RES2. This decrease was in line with the reappearance of *myo*-(1,2,4,5,6)-IP_5 and a major increase of *scyllo*-IP_5 by 250 % in the NaOH-EDTA extract of RES3.

47 % HI extraction and residual fraction

Only 31 % of the unresolved P_{org} pool was left in the residual soil RES4 of the Humeomics SCF procedure compared to the untreated soil. This unresolved P_{org} pool was represented by two underlying broad signals in the phosphomonoester region of the NMR extract on the residual fraction.

The last Humeomics extraction step caused a large increase of the NaOH-EDTA extractable IP pool. All four isomers of IP_6 were again detected in the last residue. In summary, an additional pool of 172 mg P/kg_{soil} IP could be extracted compared to RES3, with the total IP pool in RES4 comprising 45 % of P_{tot} analysed by NMR. The dominance of IP in the residual soil after the Humeomics SCF can also be seen in the NMR spectra of RES4 in Figure 3. A plethora of sharp peaks in the phosphomonoester region were detected (40), half of them were assigned to different IP isomers and conformations. It was possible to distinguish the peaks of the 2-equatorial-4-axial and 4-equatorial-2-axial conformation of D-*chiro*-IP_6 as well as the ones of the 2-equatorial-4-axial and 4-equatorial-

2-axial conformations of *neo*-IP_6 (Figure SI-9). All increases in the NaOH-EDTA extractable IP add up to 223 mg P/kg_{soil}. Therefore, at least 71 % more IP could be extracted by Humeomics compared to a single-step NaOH-EDTA extraction (UT).

Phosphodiesters were almost completely removed following the Humeomics SCF. Two pools of phosphodiesters can be distinguished: one extracted with 12 % BF_3 (52 mg P/kg_{soil}) and one extracted with 47 % HI (12 mg P/kg_{soil}).

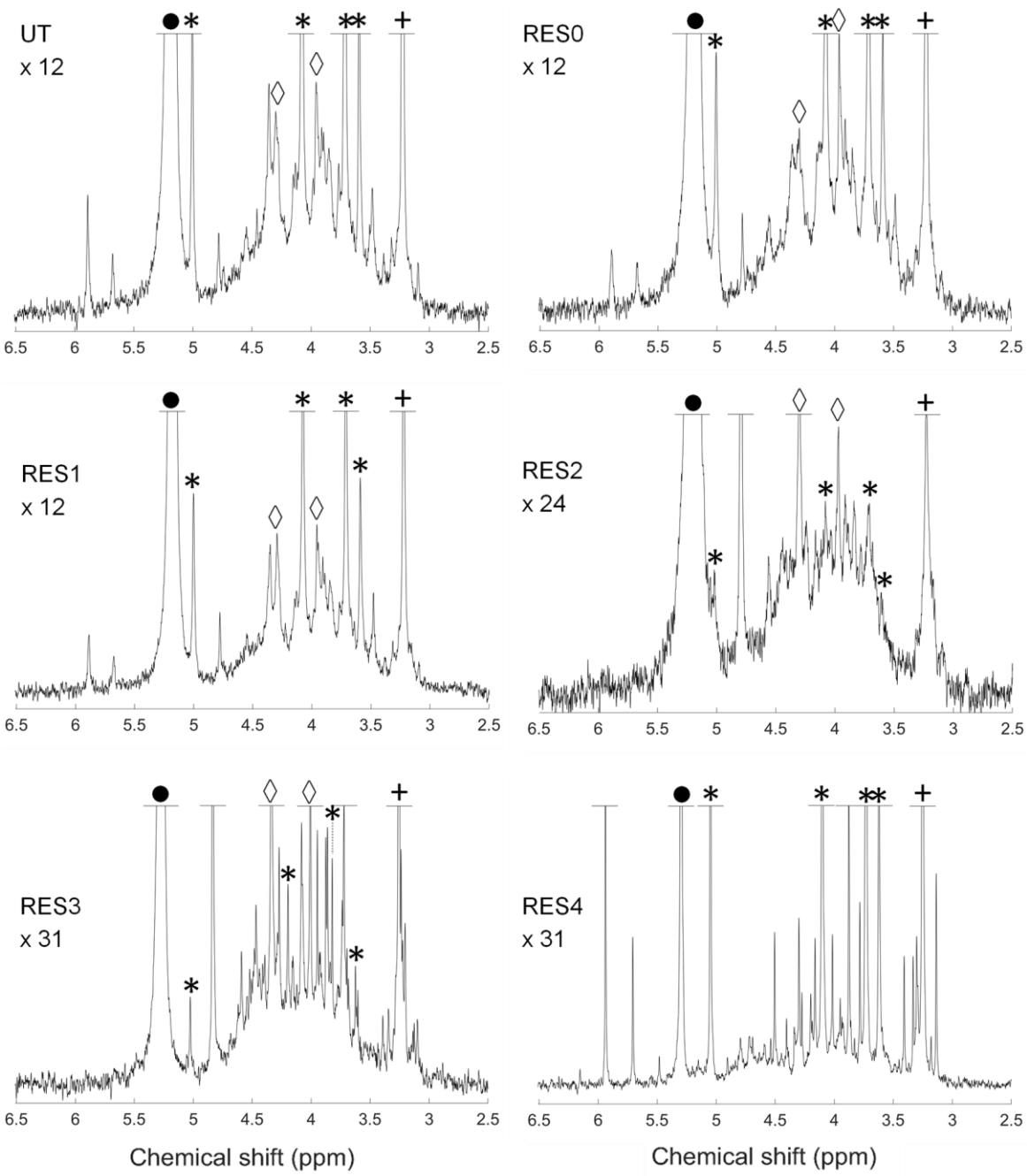


Figure 3. Solution ^{31}P NMR spectra (500 MHz) of the orthophosphate and phosphomonoester region on 0.25 M NaOH + 0.05 M EDTA Humeomics soil residues extracts. Signal intensities were normalised to the MDP peak intensity. The vertical axes were increased for improved visibility of spectral features, as indicated by a factor. The asterisks mark the four individual peaks of *myo*-IP_6, the '+' the peak of *scyllo*-IP_6, the '●' the orthophosphate peak and the symbol '◊' the two peaks of α- and β-glycerophosphate.

Table 2. Concentrations (mg P/kg_{soil}) of phosphonates, orthophosphate (ortho-P), the sum of all inositol phosphates (inositol-P) with a detailed list of all identified IP isomers below, α- and β-glycerophosphate (glycerol-P), broad peaks, other phosphomonoester (P-monoester), phosphodiester (P-diester), pyrophosphates (pyro-P), polyphosphates (poly-P) and total P (P_{tot}) measured in solution ^{31}P NMR spectra of 0.25 M NaOH + 0.05 M EDTA soil residues of the Humeomics fractionation procedure. Quantification was based on spectral integration, deconvolution fitting and the relative proportion to the peak of the added MDP standard of known concentration.

P-class	UT	RES0	RES1	RES2	RES3	RES4
Phosphonates	2	3	2	2	2	1
Ortho-P	642	630	597	173	164	100
Total Inositol-P	312	229	275	35	40	212
myo-IP_6	159	119	139	11	10	93
scyllo-IP_6	70	58	57	20	12	54
neo-IP_6	38	31	27	0	0	13
D-*chiro*-IP_6	26	10	37	0	0	15
myo-(1,2,4,5,6)-IP_5	12	8	8	0	9	22
scyllo-IP_5	7	5	7	4	10	14
Glycero-P	33	35	20	19	21	0
Broad peaks	272	263	211	138	102	83
Glucose-6-P	4	4	5	20	21	0
Other P-monoester	39	32	48	26	38	76
P-diester	49	51	65	13	13	1
Pyro-P	43	37	38	0	1	0
Poly-P	4	0	0	0	0	0
P_{tot} NMR	1400	1283	1262	427	402	473

Discussion

Carbon content and speciation in different SOM fractions

About half of the C present in the soil was removed by the Humeomics SCF procedure, suggesting a close association of this C pool with the SOM suprastructure. Similar to our findings, Drosos et al. (2017) found that the organo- and hydrosoluble fractions of the 12 % BF_3 transesterification process contained most of the extracted C of an agricultural sandy loam soil classified as Typic Ustifluvent (Piccolo, 2012). The authors reported intense *O*-alkyl signals and high abundance of fatty acids and sugars with minor amounts of phenolic acids, imines, esters, dicarboxylic acids and hydroxyacids in the ORG2 fraction. Equally to ORG2, the AQU2 fraction of their study showed intense signals in the *O*-alkyl region indicating an abundance of hydroxyl- and methoxyl-rich molecules and minor amounts of alkyl-C and aromatic-C (Drosos et al., 2017). The ^{13}C CPMAS NMR results on the soil residues of this study indicate that mostly *O*-alkyl rich compounds as carbohydrates but also alkyl-C rich compounds were extracted by cleavage of weak ester bonds. Furthermore, methoxyl-C compounds appear to be less abundant in the ORG2 + AQU2 fraction compared to the results of Drosos et al. (2017) because of their enrichment in RES2.

Interestingly, the Gleysol did not contain significant amounts of unbound C. In contrast, Nebbioso and Piccolo (2011) detected most C of a humic acid from a volcanic soil in the unbound fraction ORG1, containing the largest pool of measured hydrocarbons, fatty acids and hydroxyacids. In contrast to the volcanic soil, our results indicate that most organic molecules present in the peaty soil of this study are either part of a larger association/bound to the mineral phase or are rapidly biodegraded/incorporated when unbound. Furthermore, strong ester bonds and ether bonds appear to be rather irrelevant for the assembling of organic molecules in the SOM suprastructure of our soil. These results are in agreement with Nebbioso and Piccolo (2011), who found that only 0.9 % of the original humic acid weight was extracted with the strongly-bound fraction. The organosoluble and hydrosoluble molecules after methanolic hydrolysis (strongly-bound ester fraction) accounted for 2.4 % of total SOM solubilised by Humeomics (Drosos et al., 2017). An explanation for the negligible effect of the KOH in methanol extraction could be that the released OH-ions reacted with the soil buffer system, i.e. exchangeable Al^{3+} (Blume et al., 2009), before reacting with the SOM. Hence, the catalysis capacity for the methanolic hydrolysis could have been reduced.

The soil residue after the Humeomics procedure still contained 8 % C. However, there was a slight increase in C_{tot} from RES3 to RES4, possibly due to the addition of sodium carbonate to the soil residue in order to increase the pH or due to residues of methanol. Drosos et al. (2017) reported that 37.4 % of total soil organic C was left in the residue after the Humeomics fractionation procedure. The SOM in this unextracted pool is assumed to be closely associated to the soil mineral matrix and consist of alkyl structures (Spaccini et al., 2006; Drosos et al., 2017). These hydrophobic components are thought to be of great importance in the stabilisation of SOM due to entropy driven separation from the soil solution (Spaccini et al., 2006) and adsorption to aluminosilicate surfaces (Kleber et al., 2007; Drosos et al., 2017). The general increase of the relative amount of alkyl-C during biodegradation as part of the 'persistent' SOM is well known (Kögel-Knabner, 1997; Simpson and Simpson, 2017). Kögel-Knabner et al. (1992) found an accumulation of rigid alkyl C moieties in forest SOM but attributed it to an increase in cross-linking during the humification process. Hence, our findings on the dominance of hydrophobic alkyl-C and aromatic-C in the non-extracted SOM pool agree well with previous studies on the 'persistent' SOM pool.

In summary, approximately half of C in our soil appears to be associated with the SOM suprastructure and the other half with the soil mineral phase.

Extractability of P using the Humeomics SCF procedure of SOM

Overall, the Humeomics sequential fractionation procedure with subsequent NaOH-EDTA extraction of the final soil residue recovered about 118 % of P_{tot} when compared to a single-step NaOH-EDTA extraction of the UT soil. This could be due to the sequential 'decoating' of the SOM suprastructure by Humeomics and the resulting enhanced accessibility of P compounds for the NaOH-EDTA extractant. Especially hydrophobic domains separating the SOM suprastructure from the soil solution were removed by the DCM/MeOH extraction (Spaccini et al., 2006; Nebbioso et al., 2015), presumably leading to better extractability of incorporated P by the aqueous NaOH-EDTA solution.

Analysis of the different extracts and residues revealed that 66 % of P_{tot} was extracted using the 12 %BF_3 in methanol solution, which is attributed to the fraction associated with the SOM suprastructure through weak ester bonds (Nebbioso and Piccolo, 2011). However, not all of the extracted P was recovered in the extracts, mostly due to losses in the preparation process of the extracts for total microwave digestion. This was mostly the case for the 12 % BF_3 extraction, where the NaOH-EDTA extractable P_{tot} in the soil

residues decreased by 907 mg P/kg_{soil}, but only 99 mg P/kg_{soil} were measured by microwave digestion in the respective extracts combined. Especially AQU2 was difficult to freeze-dry and quantitatively measure due to the high ion content. Only removal of the <1 kDa fraction allowed for solution ^{31}P NMR spectroscopic analysis, which revealed that 279 mg P/kg_{soil} was present in the AQU2 extract. Therefore, we assume that concentrations of P_{tot} especially in the AQU2 extract are underestimated and quantitative evaluation of the extractability of P compounds using Humeomics should rather be based on the differences in the soil residues in this study.

Associations of phosphorus species with SOM fractions

The findings of this study suggest that not only the total content of P associated with SOM varies between SOM pools but also the composition of P. So far, most studies on ratios of soil C to P focus on total P_{org} concentrations and do not investigate the P_{org} composition (John et al., 1965; Neptune et al., 1975; Goh and Williams, 1982; Stevenson, 1986). The average C:P_{org} ratio in this study was 162 and therefore in range of SOM ratios reported by various studies (Goh and Williams, 1982; Stevenson, 1986; Tipping et al., 2016).

The findings of our study are summarised in Figure 4. The summary is based on the solution ^{31}P NMR spectroscopy results of the Humeomics SCF soil residues and shows the distribution of different P_{org} compounds within the fractions as well as their relative abundance. The first fraction ‘cation-complexed water soluble’ represents P_{org} removed by the HCl washing step, which is not bound to the SOM suprastructure but presumably free in soil solution, complexed with cations i.e. Ca and/or adsorbed to surfaces of calcite (Celi et al., 2000). The second fraction ‘unbound organo soluble’ consists of P_{org} compounds that are soluble in organic solvents (i. e. DCM/MeOH), such as phospholipids. These compounds are either free or associated by weak dispersive forces (Piccolo et al., 2019). According to Nebbioso and Piccolo (2011), the SOM suprastructure is differentiated by the different bonding types based on the applied extraction methods: weakly ester-bound (12 % BF_3), strongly ester-bound (0.5 M KOH) and ether-bound (47 % HI). Furthermore, the summary includes P_{org} in the final residue, which is assumed to represent SOM associated with the mineral phase (Nebbioso and Piccolo, 2011).

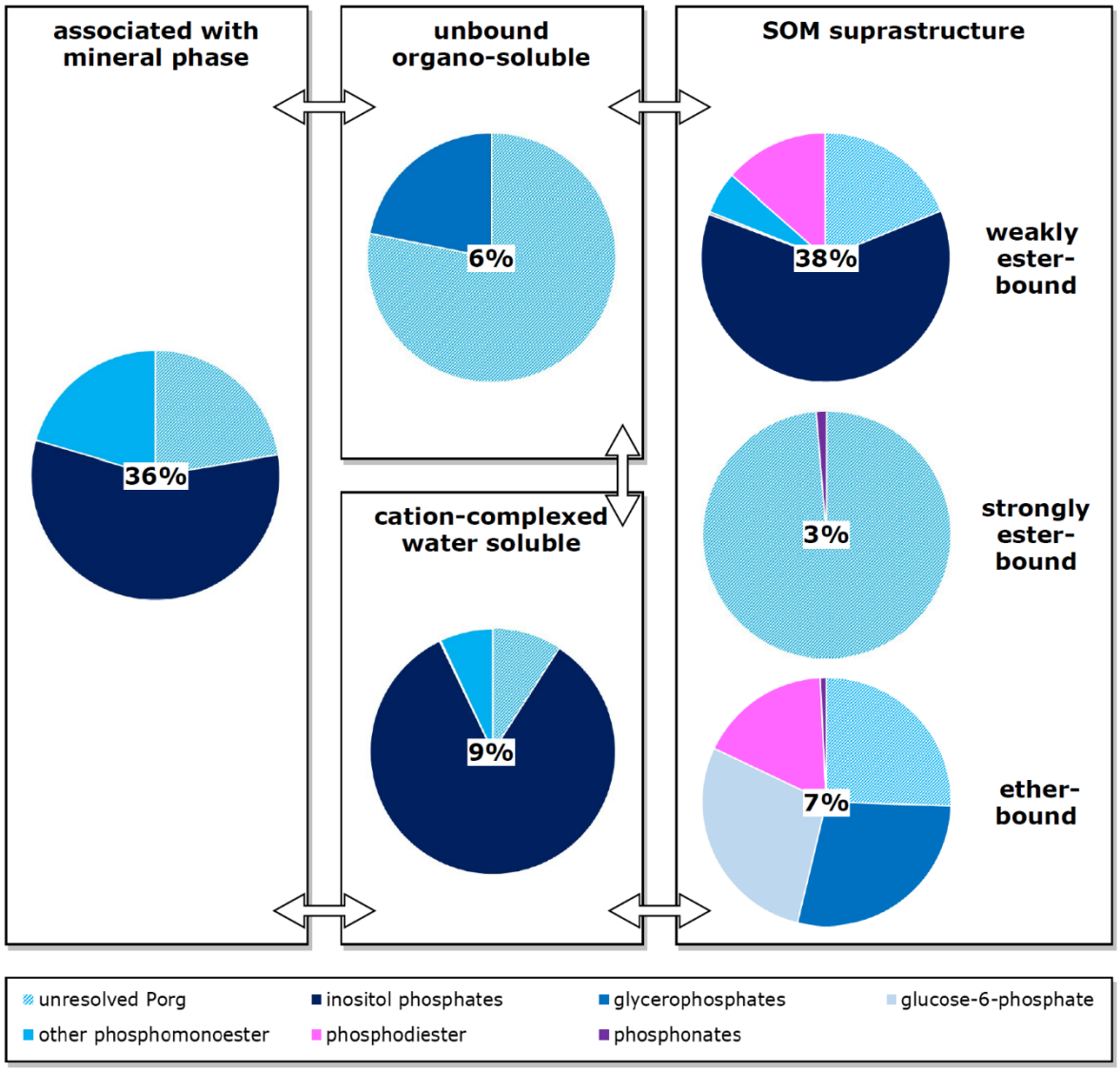


Figure 4. Summary of soil organic P extracted with the Humeomics sequential extraction procedure of the SOM suprastructure and in the residual fraction (associated with mineral phase). The fractions include: 0.1 M HCl extraction (water soluble and/or cation-complexed P_{org}), DCM/MeOH extraction (unbound organo-soluble P_{org}), 12 % BF_3 extraction (weakly ester-bound P_{org}), 0.5 M KOH extraction (strongly ester-bound P_{org}) and 47 % HI extraction (ether-bound P_{org}). The quantitative pie charts are based on ^{31}P NMR results of the NaOH-EDTA extracts of soil residues after each Humeomics sequential extraction step (Table 2). Percentages indicate the proportion of P_{org} removed with the respective fraction compared to total removed P_{org} values as measured by ^{31}P NMR spectroscopy. The arrows represent possible exchange/transformation mechanisms between the different pools.

Unresolved P_{org} pool

Our results provide further evidence that the unresolved P_{org} pool is not comprised of a single, large polymeric structure with C–C bonds and phosphate as functional groups but mainly as part of the SOM associations of smaller, recognisable molecules. The associations of these unresolved P_{org} molecules with other organic molecules of the SOM would result in the apparent high molecular size reported by Jarosch et al. (2015); McLaren et al. (2015b) and McLaren et al. (2019). The phosphate containing molecules causing the broad signals are mostly associated with the SOM through weak and strong ester linkages (40 %) as defined by Nebbioso and Piccolo (2011). However, the extraction with KOH in methanol could also have caused a ligand exchange with phosphate groups bound to the mineral phase, hence additionally extracting unresolved P_{org} molecules associated with the soil mineral phase. This effect needs to be further investigated on model compounds containing P_{org} adsorbed to mineral surfaces.

The close association with the SOM structure could result in the described resistance of the unresolved P_{org} pool to enzymatic hydrolysis (Jarosch et al., 2015) and chemical stability (Reusser et al., 2020b), because the accessibility is limited due to the hydrophobic nature and complex structure of these associations, possibly hindering microbial attack (Schmidt et al., 2011).

A smaller proportion of the unresolved P_{org} pool (unbound fraction) appears to be free or part of associations held together only by weak dispersive forces (22 %). Another part (30 %) is still present in the SOM pool associated with the soil mineral phase. It is not known whether this remaining unresolved P_{org} pool exhibits a polymeric structure or not. Nebbioso et al. (2015) found an enrichment of aromatic compounds in this residual fraction. Similarly, Drosos et al. (2017) reported that the SOM fraction associated with the mineral phase consisted of mostly oxygen and nitrogen heterocyclic compounds, benzoic acids, and esters. The authors proposed that in this fraction, alkyl apolar components assemble by hydrophobic dispersive forces into larger size aggregates. These findings would rather not support the model of a macropolymeric structure of the unresolved P_{org} pool in the final residue.

McLaren et al. (2019) and Reusser et al. (2020b) found that the underlying broad peaks in the phosphomonoester region are not comprised of a range of sharper signals arising from IP. This study provides further evidence that the broad peaks represent one or several separate, unresolved P_{org} pools, as the continuous decrease of these compounds in the

NaOH-EDTA extracts by subsequent decoating of the SOM suprastructure is decoupled from the extractability trend of the IP and other phosphomonoesters.

Inositol phosphates

Two major pools of IP have been identified: the first pool being closely associated with the SOM and the second pool with the soil mineral phase. The association of IP with the SOM suprastructure would explain the occurrence of IP high molecular weight material (Moyer and Thomas, 1970; Veinot and Thomas, 1972; Borie et al., 1989).

The extractability of IP could be markedly enhanced using the Humeomics SCF with results indicating that only 58 % of the total IP pool in soil can be extracted using a single-step NaOH-EDTA extraction. Furthermore, we assume that many sharp peaks in the phosphomonoester region of RES4 could be attributed to unknown IP based on a hypobromite oxidation study of the same soil (Reusser et al., 2020b), resulting in an even higher total IP pool in the investigated soil. Previous studies on quantitative analysis of IP using NMR were mostly based on a single alkaline extraction step (Turner et al., 2003c; Turner and Richardson, 2004; McLaren et al., 2019; Reusser et al., 2020a). The findings of our study imply that especially the IP pool associated with the mineral phase of some soils has been underestimated using a single-step NaOH-EDTA extraction.

Drosos et al. (2017) suggested that stabilisation of the SOM pool associated with the mineral phase should also be attributed to complex formation between oxygen-containing hydrophilic groups and mineral iron. The phosphate groups of IP would fulfil these requirements and serve as important anchor molecules between the soil mineral phase and SOM. Furthermore, IP are well known for their adsorption affinity to minerals and organic matter (Cosgrove and Irving, 1980; Ognalaga et al., 1994). Hence, the high abundance of IP in the final residue could be due to IP released from the soil mineral phase through acidic hydrolysis by HI.

Tipping et al. (2016) reported that P contributes 3 % to the total mass of ‘nutrient rich’ SOM (according to their mixing-model), which is stabilised by sorption to the mineral phase. The authors suggested that specific molecules as *myo*-IP_6 are possibly selectively incorporated into this SOM pool because of their strong sorption capacity. Our findings are in agreement with the suggestion of the authors that *myo*-IP_6 is the most prevalent form in the SOM fraction associated with the mineral phase. Moreover, our study provides evidence that not

only *myo*-IP_6 but also other IP forms and phosphomonoesters could play a role in stabilising the SOM pool.

The significant reduction of IP in the transesterification process with BF_3 is assumed to result mainly from breaking weak ester bonds between the SOM and organic molecules associated with the phosphate groups of IP and not from breaking the ester bond between the phosphate groups and the inositol ring itself. The reason for this assumption is that the phosphate ester linkage of IP is considered to be quite stable with slow chemical hydrolysis in acidic media (Turner et al., 2002). However, no IP has been recovered in the AQU2 fraction. Therefore, we suggest to test the effect of the BF_3 catalysed transesterification on pure IP standards as well as on soils containing a known concentration of IP.

Other P species

Glycerophosphates are hydrolysis products of phospholipids e.g. phosphatidylcholine (Doolette et al., 2009) as part of cell membranes (Strickland, 1973). Therefore, we can assume that the original phospholipids in the investigated soil were either unbound, part of the microbial biomass or completely liberated by breaking the ether bonds of the SOM structure through acidic hydrolysis. The removal of phospholipids in the unbound fraction appears to be coupled with the removal of alkyl-C as in fatty acids (Stevenson, 1994). However, the observed removal of phospholipids in the last residue was not coupled with a decrease in alkyl-C. Hence, phospholipids appear to be less important in the stabilisation of the SOM structure compared to IP, despite their hydrophobic character.

Most of the phosphodiesters in the NaOH-EDTA residues extracts are present in form of DNA (Makarov et al., 2002). The removal of the hydrophobic unbound fraction resulted in an increase of 133 % of the NaOH-EDTA extractable phosphodiester pool compared to the untreated soil. The reason for this could be the better accessibility of DNA by the alkaline extractant. In contrast to glucose-6-phosphate and phospholipids, the subsequent transesterification process resulted in a significant decrease of the DNA pool in the soil residue, suggesting an association of DNA with the SOM structure. However, similar to phospholipids and glucose-6-phosphate, almost no DNA was extracted from the SOM pool associated with the mineral phase. These results are consistent with findings of Nebbioso et al. (2015), who reported that plant-derived acids remained in the SOM associated with the mineral phase rather than microbial metabolites. Furthermore, Drosos et al. (2017) showed that the majority of lipids in the stable fraction was derived from plant polyesters rather than from microbial activity. Our results indicate that metabolites and degradation

products of the microbial biomass could be more abundant in the SOM structure compared to the SOM pool associated with the mineral phase. These findings partially contradict Kirkby et al. (2011), who proposed that the stable SOM pool is comprised of microbial detritus due to similar stoichiometries. The authors identified microbial biomass as an immediate precursor of stable SOM. However, the effect of the 12 % BF_3 extraction on phosphodiesters needs further investigation in order to calculate the proportion of phosphodiesters destroyed by the extractant to the proportion liberated from the SOM structure.

In summary, we identified weak ester bonds as the main bonding type in P-SOM associations. Further research is needed to investigate the formation and transformation processes of these P-SOM associations in soils of diverse SOM and P_{org} composition. Furthermore, the effects of the Humeomics SCF steps on standard P_{org} compounds should be tested in order to rule out any extraction artefacts. Especially the transesterification process catalysed by BF_3 needs further investigation. The large amount of orthophosphate with the concomitant absence of P_{org} compounds in AQU2 could indicate that the P_{org} compounds themselves were transesterified, i.e. the bond between the phosphate group and the organic molecule being broken. The proportion of P_{org} compounds destroyed by the BF_3 extraction needs to be evaluated in order to assess the proportion of P_{org} actually liberated from the SOM suprastructure.

Implications on SOM models

The SOM pool associated with the mineral phase of this study is comparable to SOM adsorbed to the mineral surfaces in the soil continuum model of Lehmann and Kleber (2015). In contrast to the proposed soil continuum model of Lehmann and Kleber (2015), our mineral associated SOM fraction includes small organic molecules containing P, i.e. IP, which directly originate from plant and animal residues and are of high abundance in the P_{org} pool. Therefore, not only large biopolymers from plant and animal residues source the SOM pool and more attention should be drawn on specific molecules, e.g. IP, which could function as anchor in the SOM stabilisation process (Tipping et al., 2016).

Conclusions

The combination of the Humeomics SOM fractionation procedure with solution ^{31}P NMR spectroscopy revealed that the most abundant form of P_{org} in soil, the unresolved pool, is associated with the SOM structure, mainly through ester and ether bonds. The subsequent decrease of this unresolved P_{org} pool with the SOM fractionation advances our understanding of its chemical nature away from the macropolymeric theory and more towards an association point of view. Furthermore, our findings indicate that especially the pool of IP in soil is underestimated using a single-step NaOH-EDTA extraction for soil P_{org} determination. These relatively stable P_{org} compounds dominated the residual fraction of the Humeomics SOM fractionation procedure, which represents SOM associated with the mineral phase. However, the effects of the Humeomics extractants on individual P_{org} compounds need further investigation.

The improved knowledge on the chemical nature of the unresolved P_{org} pool as well as the revelation of the bonding types of diverse P compounds with the SOM will impact our understanding of P cycling, transformation and stabilisation processes in soils.

Supporting information

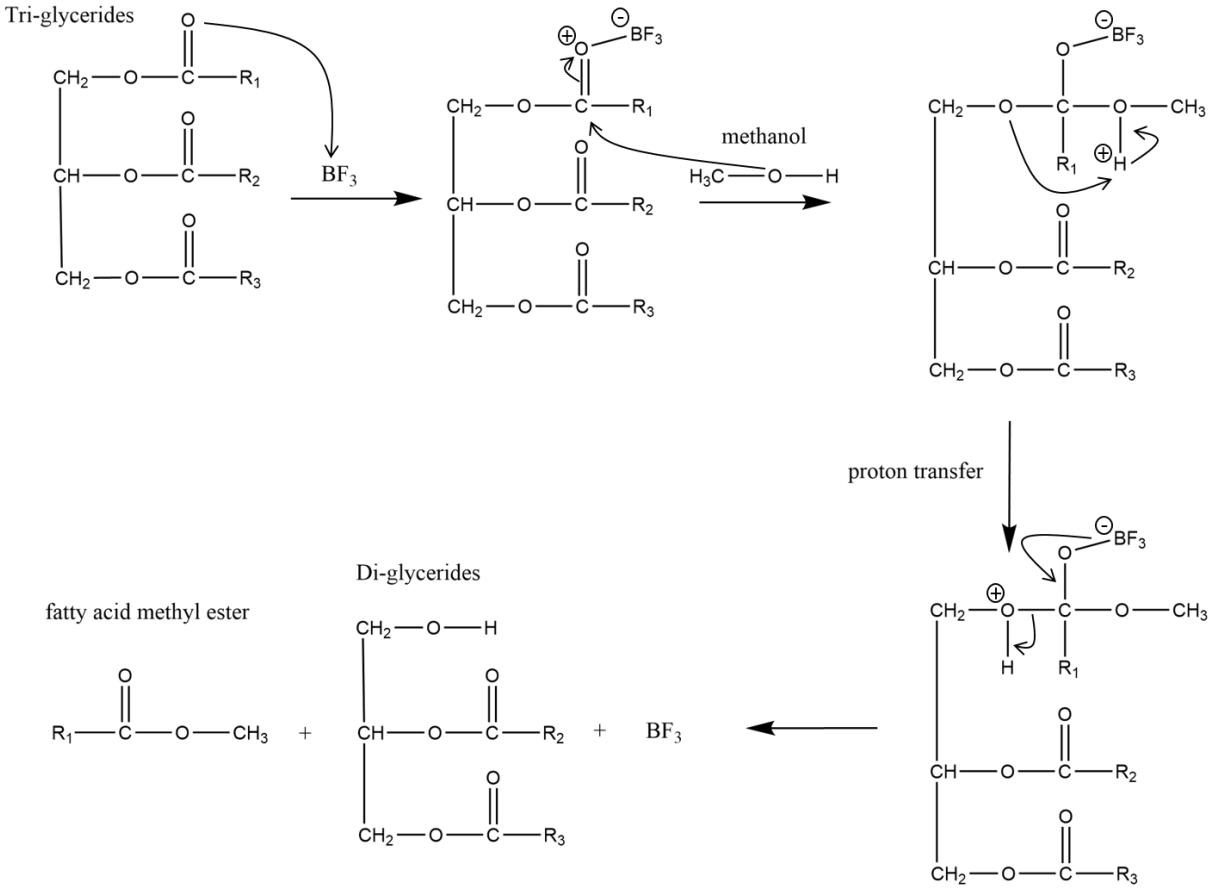


Figure SI-1. Transesterification reaction mechanism modified from Thangarasu and Anand (2019). This example illustrates the BF_3 catalysed transesterification of tri-glycerides to di-glycerides using methanol. Illustrated using ChemDraw Professional 17.0 PerkinElmer Inc.

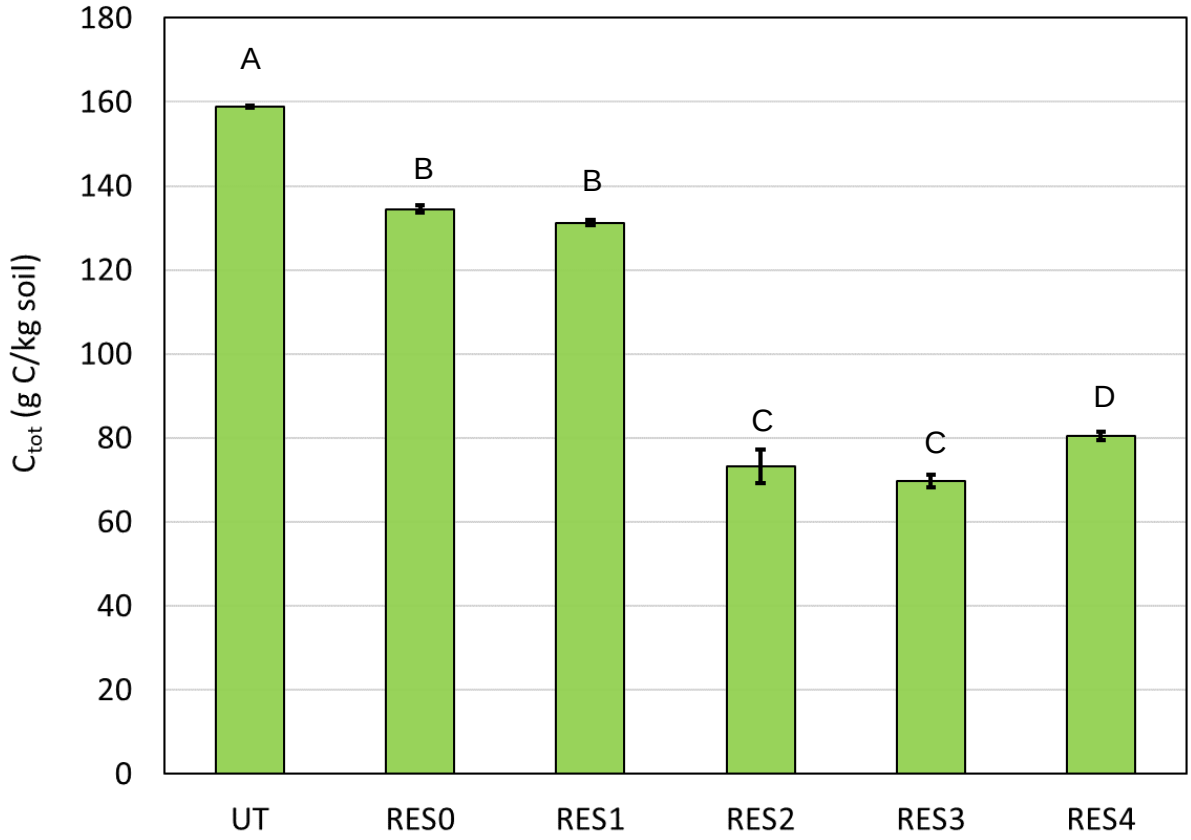


Figure SI-2. Total C concentrations in the soil residues of the Humeomics fractionation procedure in g C per kg soil. Error bars denote standard deviation of n=3 replicates, letters indicate significant differences of means determined with a one-way ANOVA and Tukey's honestly significant difference procedure ($p<0.0001$). For RES4, sodium carbonate has been added to the soil residue in order to increase the pH.

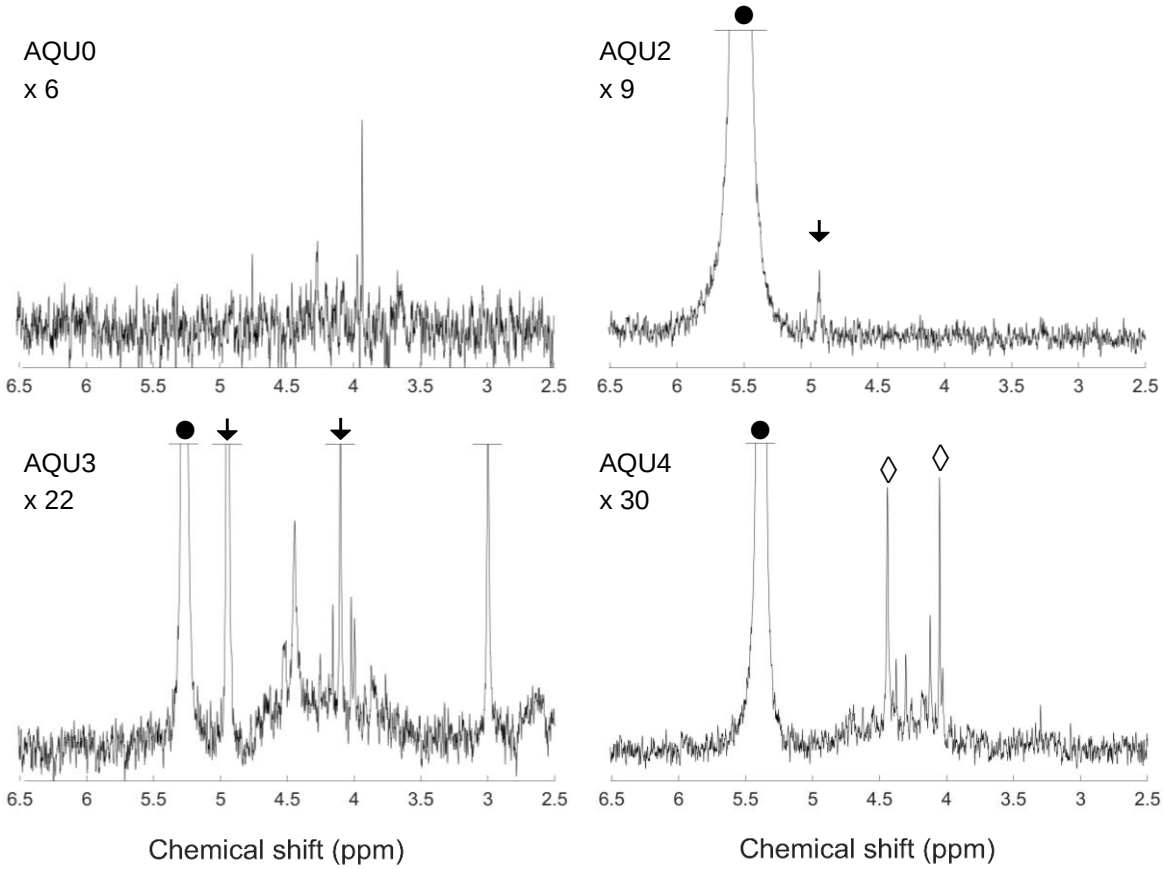


Figure SI-3. Solution ^{31}P NMR spectra (500 MHz) of the orthophosphate and phosphomonoester region on 0.25 M NaOH + 0.05 M EDTA Humeomics aqueous extracts. Signal intensities were normalised to the MDP peak intensity. The vertical axes were increased for improved visibility of spectral features, as indicated by a factor. The '●' marks the orthophosphate peak, the arrows the peak of glucose-6-phosphate and the symbol '◊' the two peaks of α- and β-glycerophosphate.

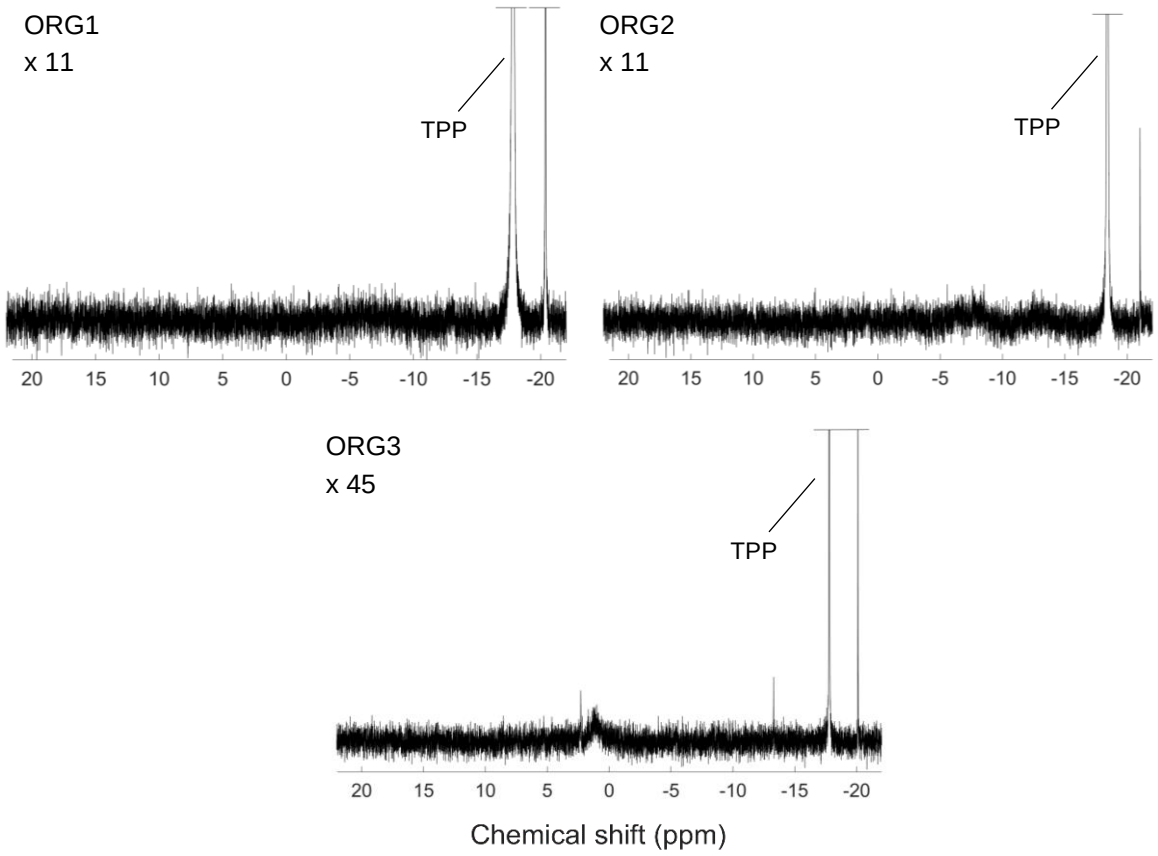


Figure SI-4. Solution ^{31}P NMR spectra (500 MHz) of the $CDCl_3$ Humeomics ORG extracts. The added triphenylphosphate standard (TPP) is marked. Signal intensities were normalised to the TPP peak intensity. The vertical axes were increased for improved visibility of spectral features, as indicated by a factor.

Table SI-1. Concentrations (mg P/kg_{soil}) of phosphonates, orthophosphate (ortho-P), inositol phosphates (inositol-P), α- and β-glycerophosphate (glycerol-P), broad peaks, other phosphomonoester (other P-monoester), phosphodiester (P-diester), pyrophosphates (pyro-P), polyphosphates (poly-P) and total P (P_{tot}) measured in solution ^{31}P NMR spectra of redissolved Humeomics extract, either redissolved in 0.25 M NaOH + 0.05M EDTA (aqueous extracts AQU0-4) or in $CDCl_3$ (organic solvents extracts ORG1-3). Quantification was based on spectral integration of regions with NMR signals, and spectral deconvolution fitting of the phosphomonoester region. Concentrations of total P (P_{tot}) as measured by microwave acid digestion.

P-class	AQU0	AQU2	AQU3	AQU4	ORG1	ORG2	ORG3
Phosphonates	0.0	0.0	1.2	4.7	0.0	0.0	0.0
Ortho-P	0.0	273.5	2.0	533.6	0.0	0.0	0.0
Inositol-P	0.0	0.0	0.0	0.7	0.0	0.0	0.0
Glycero-P	0.0	0.0	0.0	11.7	0.0	0.0	0.0
Broad peaks	0.0	0.0	0.7	29.6	0.0	0.0	0.0
Glucose-6-P	0.0	1.1	0.5	0.0	0.0	0.0	0.0
Other P-monoester	21.1	0.0	0.6	15.1	0.0	0.0	0.0
P-diester	0.0	1.1	1.6	0.0	0.0	0.0	0.3
P-triester	0.0	3.1	0.0	0.0	0.0	0.0	0.0
Pyro-P	0.0	0.0	0.0	0.0	0.0	0.0	0.02
Poly-P	0.0	0.0	0.0	0.0	0.3	0.1	0.2
P_{tot} NMR	21.1	278.8	6.5	595.4	0.3	0.1	0.5
P_{tot} digestion	11.7	71.5	6.3	1039.9	0.8	1.6	0.0

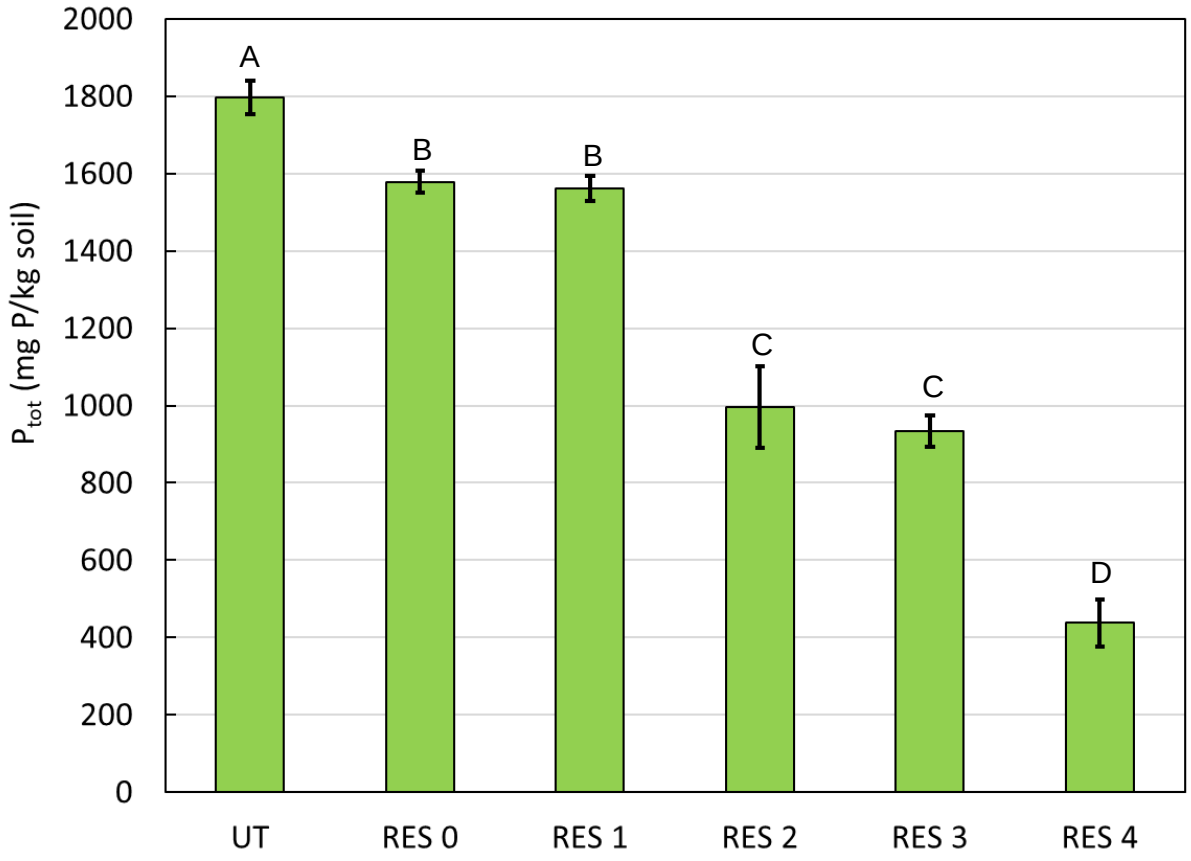


Figure SI-5. Total P concentrations in the soil residues of the Humeomics fractionation procedure in mg P per kg soil (microwave digestion). Error bars denote standard deviation of n=3 replicates, letters indicate significant differences of means determined with a one-way ANOVA and Tukey's honestly significant difference procedure ($p<0.0001$).

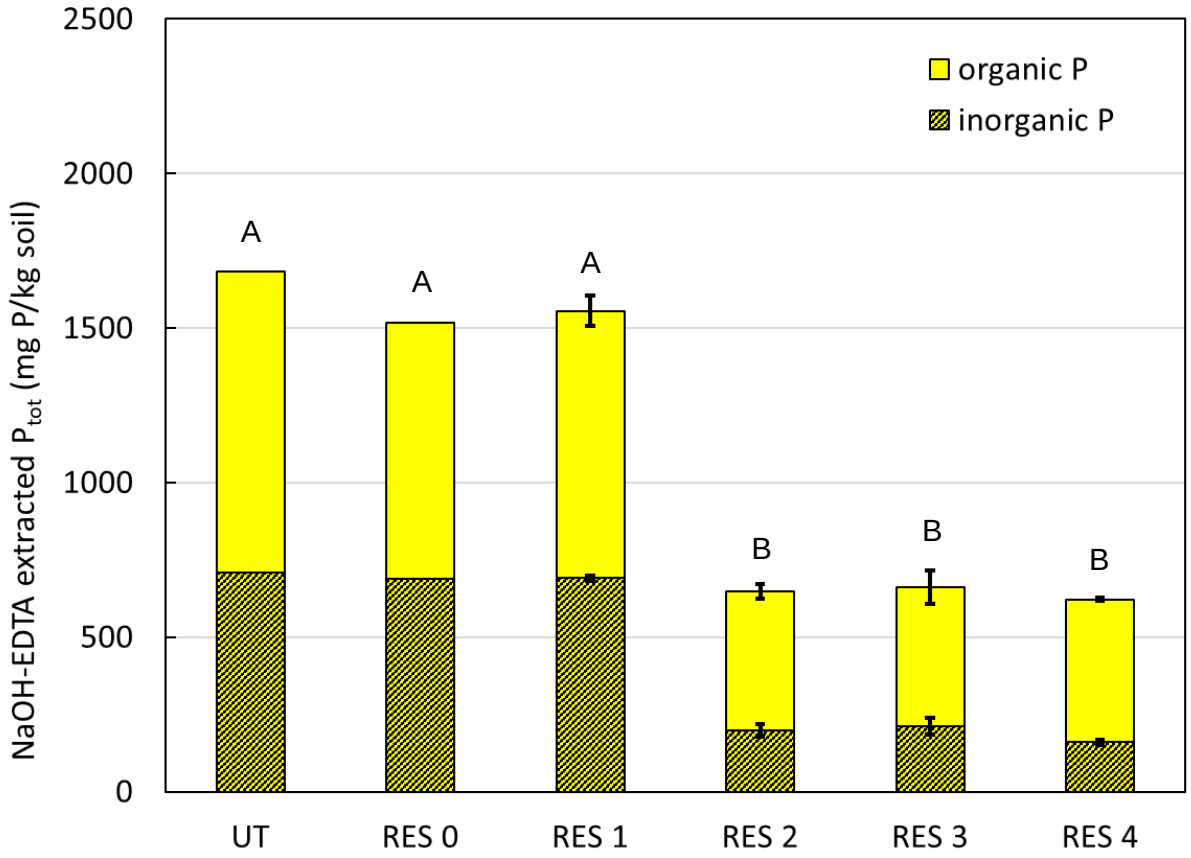


Figure SI-6. Organic and inorganic (shaded) P concentrations in 0.25 M NaOH + 0.05 M EDTA extracts of the Humeomics soil residues. Error bars denote standard deviation of n=3 replicates (RES4 n=2). Letters indicate significant differences of means determined with a one-way ANOVA and Tukey's honestly significant difference procedure ($p<0.0001$).

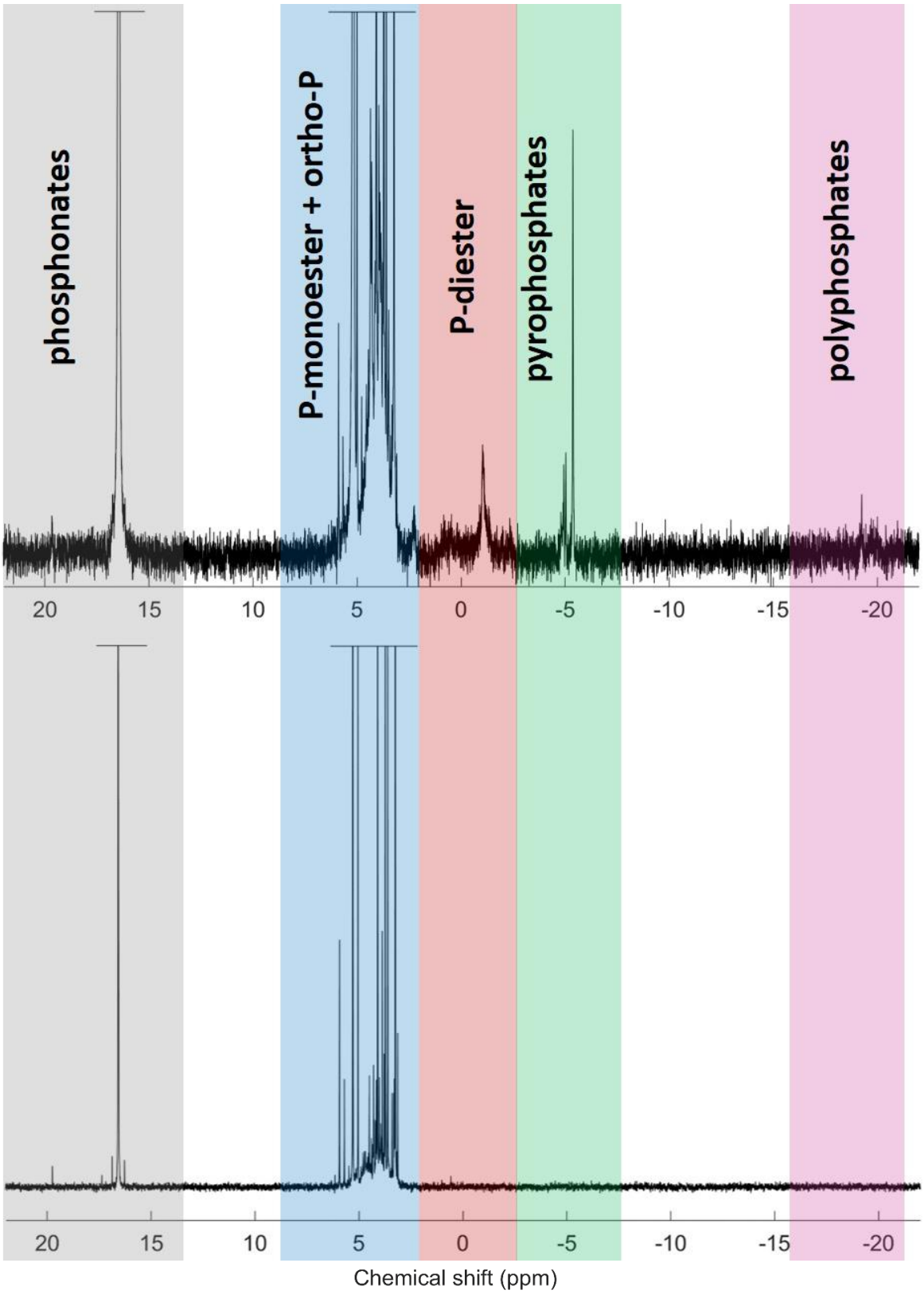


Figure SI-7. Solution ^{31}P NMR spectra (500 MHz) of 0.25 M NaOH + 0.05 M EDTA extracts of the untreated soil (above) and the residual soil RES4 after the Humeomics sequential chemical fractionation procedure (below). Signal intensities were normalised to 0.085-times the MDP peak intensity. The following chemical shift regions are highlighted: phosphonates (δ 19.9 to 13.3 ppm), the combined orthophosphate and phosphomonoester region (δ 6.2 to 2.9 ppm), phosphodiester (δ 2.8 to -2.0 ppm), pyrophosphates (δ -4.8 to -5.5 ppm) and polyphosphates (δ -17.1 to -20.6 ppm).

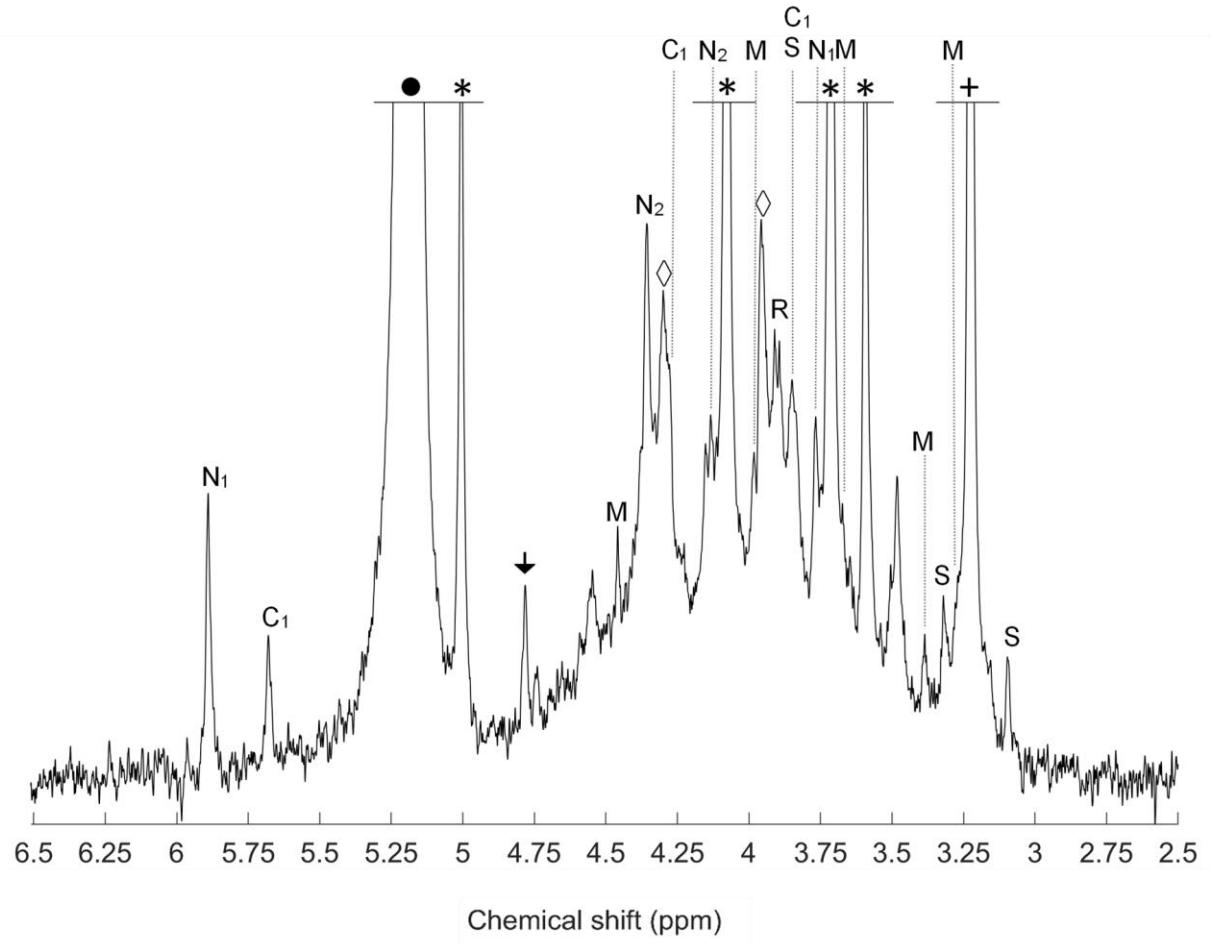


Figure SI-8. Solution ^{31}P NMR spectra (500 MHz) of the orthophosphate and phosphomonoester region on the 0.25 M NaOH + 0.05 M EDTA extract of the untreated soil. Signal intensities were normalised to 0.085-times the MDP peak intensity. The '*' mark the four individual peaks of *myo*-IP$_6$, the '+' the peak of *scyllo*-IP$_6$, the '●' the orthophosphate peak, the symbol '◊' the two peaks of α- and β-glycerophosphate and the arrow the peak of glucose-6-phosphate. Furthermore, the following peaks are marked: 4-eq-2-ax *neo*-IP$_6$ (N$_1$), 2-eq-4-ax *neo*-IP$_6$ (N$_2$), 2-eq-4-ax *chiro*-IP$_6$ (C$_1$), *myo*-(1,2,4,5,6)-IP$_5$ (M), *scyllo*-IP$_5$ (S) and RNA mononucleotides (R).

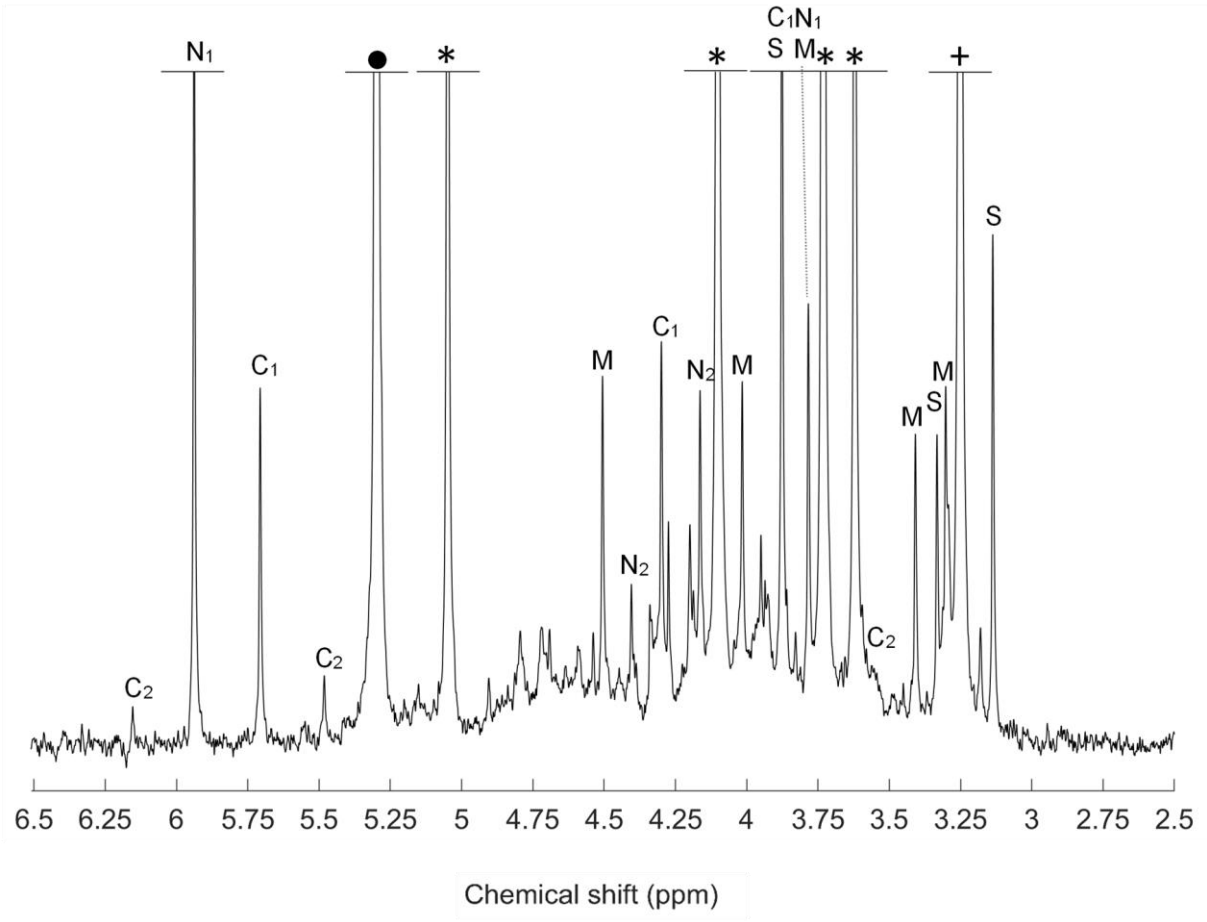


Figure SI-9. Solution ^{31}P NMR spectra (500 MHz) of the orthophosphate and phosphomonoester region on the 0.25 M NaOH + 0.05 M EDTA extract of the final soil residue RES4 after the Humeomics sequential chemical fractionation procedure. Signal intensities were normalised to 0.032-times the MDP peak intensity. The '*' mark the four individual peaks of *myo*-IP_6, the '+' the peak of *scyllo*-IP_6 and the '●' the orthophosphate peak. Furthermore, the following peaks are marked: 4-eq-2-ax *neo*-IP_6 (N_1), 2-eq-4-ax *neo*-IP_6 (N_2), 2-eq-4-ax *chiro*-IP_6 (C_1), 4-eq-2-ax *chiro*-IP_6 (C_2), *myo*-(1,2,4,5,6)-IP_5 (M), *scyllo*-IP_5 (S).

General discussion

Main findings

The objective of this PhD project was to identify the chemical nature of soil P_{org} with increasing molecular weight by combining several chemical extraction approaches, size separation procedures, and solution ^{31}P NMR spectroscopy techniques. The P_{org} pool of the six soils investigated in this PhD project ranged from 70 to 1135 mg P/kg_{soil} and comprised 12 % to 63 % of P_{tot} measured by total digestion.

The main findings of this PhD project are:

- There is a much wider diversity and greater abundance of IP in soil than previously thought, which are present as sharp NMR signals in the phosphomonoester region.

- Quantification of IP requires an underlying broad signal in the phosphomonoester region when carrying out SDF.

- Several lower-order IP were detected in soil extracts for the first time using solution ^{31}P NMR spectroscopy, including four stereoisomers of IP_5 (*myo*-, *scyllo*-, *neo*-, and *chiro*-IP_5), *scyllo*-IP_4 and two isomers of *myo*-IP_5 (*myo*-(1,2,4,5,6)-IP_5 and *myo*-(1,2,4,5,6)-IP_5). Moreover, there is supporting evidence that up to 35 unidentified sharp signals in the phosphomonoester region are also due to lower-order IP, based on experiments involving hypobromite oxidation of soil extracts.

- The IP pool in soil can be further divided into IP closely associated with the SOM and forming apparent large molecular size compounds, or IP closely associated with the soil mineral phase.

- The pool of unresolved P_{org} in the phosphomonoester region, which can be accounted for as a broad underlying signal in solution ^{31}P NMR soil spectra, is chemically and structurally different compared to the sharp signals within this region (e.g. *myo*-IP_6).

- The pool of unresolved P_{org} is comprised of a continuum of apparent large molecular size compounds rather than a single macropolymeric structure at distinct molecular size.

- The unresolved P_{org} compounds are closely associated with the SOM structure, mainly through ester linkages.

The findings of this PhD project on the chemical nature of soil P_{org} with increasing molecular weight, their advancement of our understanding of soil P_{org} as well as their implications in a broader context are discussed in more detail in the following sections.

Inositol phosphates

The proportion of P_{org} in soil comprised of IP is considered to vary considerable, ranging from 0 % to 90 % (Cosgrove and Irving, 1980; Harrison, 1987; McDowell and Stewart, 2006; Smernik and Dougherty, 2007). Studies reporting concentrations of IP often focused on the *myo* isomer of IP_6 or its salts (i.e. phytate) (McDowell and Stewart, 2006; Negassa et al., 2010; Colocho Hurtarte et al., 2020) due to the assumed highest abundance of this isomer in soils (Halstead and Anderson, 1970; Turner, 2007) and/or analytical limitations, e.g. the spectral resolution and sensitivity of some solution ^{31}P NMR spectrometers. Furthermore, Smernik and Dougherty (2007) suggested that *myo*-IP_6 concentrations were overestimated in past studies using solution ^{31}P NMR due to possible misassignment of peaks and exclusion of the unresolved P_{org} pool in the spectral processing and quantification procedure.

In this thesis, concentrations of *myo*-IP_6 across the six soils ranged from ~1 (Vertisol) to 90 mg P/kg_{soil} (Gleysol), comprising 1 % to 23 % of total P_{org} as measured by solution ^{31}P NMR on NaOH-EDTA extracts (chapter 1, single-step extraction). Furthermore, we detected all four stereoisomers of IP_6 as well as the two different conformations of *neo*- and *chiro*-IP_6 (chapters 2-4). We proved the existence of lower-order IP in soil for the first time using direct spectroscopic techniques (NMR analysis) on NaOH-EDTA extracts following hypobromite oxidation (chapter 2). One third of detected IP was in the form of IP_5 and IP_4 across four soil samples.

Sequential chemical fractionation of the Gleysol using the Humeomics method of Nebbioso and Piccolo (2011) revealed that the extent of the IP pool in soil increased to 535 mg P/kg_{soil}, compared to 312 mg P/kg_{soil} in the initial NaOH-EDTA soil extract (chapter 4). However, these extracts also contained a plethora of unidentified sharp signals, which are most likely due to IP based on findings from chapter 2. Consequently, the single-step NaOH-EDTA extraction of soil may significantly underestimate the pool of IP in soil. The findings of this thesis provide evidence that the abundance and variety of the IP pool in soil can be underestimated using a single-step extraction with NaOH-EDTA. This is in line with findings of Celi et al. (2020), who reported that the NaOH-EDTA extraction of IP_6 performed differently depending on the type of Fe oxides involved in the adsorption process of IP_6. The authors suggested an increase of the NaOH concentration in the extractant or the use of sequential extractants, e.g. NaOH and oxalic acid, for a more efficient IP_6 extraction from soil.

Inositol hexakisphosphates

The accurate quantification of *myo*-IP_6 in alkaline soil extracts using solution ^{31}P NMR spectroscopy requires the inclusion of an underlying broad signal in the SDF process of the phosphomonoester region (chapter 1). Otherwise, concentrations of *myo*-IP_6 will be overestimated in the soil extract. Prior to this PhD project, the influence of the underlying broad signal on the quantification process of *myo*-IP_6 was only tested on one sample, a calcareous soil originating from a wheat field (Doolette et al., 2010; Doolette et al., 2011a). In contrast, the six soil samples examined in our study covered a wide diversity of P_{org} concentrations and composition of the phosphomonoester region. Furthermore, the improved resolution and sensitivity of our NMR spectra compared to the previously cited studies allowed for detection of considerably more sharp signals in the phosphomonoester region. Consequently, we were able to calculate the spike recoveries of the four individual peaks of *myo*-IP_6. The resulting difference of spike recoveries between individual peaks provided more evidence of an imperfect Gaussian/Lorentzian distribution of the underlying broad signal, caused by its composition of more than one component. This partitioning of the underlying broad signal was already suggested by Doolette and Smernik (2015) and further supported by McLaren et al. (2019) and in the course of this PhD project (chapters 2, 3 and 4). The four peaks of *myo*-IP_6 are distributed along the phosphomonoester region from the assumed beginning almost to the end of the underlying broad signals. Therefore, it can be assumed that also the net peak area and hence the concentration of all sharp peaks arising from other phosphomonoesters located between the C2 and C5 peak of *myo*-IP_6 (e.g. *chiro*- and *neo*-IP_6, lower-order IP, α- and β-glycerophosphates and RNA mononucleotides) are overestimated when fitting the sharp peaks directly to the baseline without an underlying broad signal (chapter 1).

The SDF approach without an underlying broad signal was introduced by Turner et al. (2003c) and applied in various studies, e.g. Hill and Cade-Menun (2009) and Deiss et al. (2016). Turner et al. (2003c) tested the accuracy of the applied SDF approach by analysing solutions that contained a mixture of P compound standards along with a *myo*-IP_6 standard. The calculated mean recovery of the *myo*-IP_6 standard in the model P compound mixture was approx. 100 %, leading to the assumption that the SDF approach resulted in an accurate quantification of *myo*-IP_6. Nevertheless, the authors did not test the quantification accuracy in a real soil mixture but in a model mixture, limiting the significance of this method. The reason for this limitation is that model mixtures and model soils as well as extracts of plants anyway lack the P_{org} compounds which cause the underlying broad

signal in the phosphomonoester region (Bünemann et al., 2008b; Noack et al., 2012; Doolette and Smernik, 2016). Another approach for assessing the quantification accuracy of *myo*-IP_6 by the applied SDF procedure was to compare the peak ratio of the four *myo*-IP_6 peaks (Vincent et al., 2012), which should be close to the theoretical ratio of 1:2:2:1 (Costello et al., 1976). However, as reported in chapter 1, this ratio was similar between the two SDF procedures with and without an underlying broad signal, questioning the usefulness of this method to assess the accuracy of the quantification.

The majority of IP detected in the six soil samples of this PhD project were present in the highest phosphorylation state of the inositol ring (IP_6). The abundance of the four stereoisomers of IP_6 was in the order of *myo* (61 %) > *scyllo* (29 %) > *chiro* (6 %) > *neo* (4 %) (chapter 2). These findings are in agreement with previous studies on the abundance of the four IP_6 stereoisomers detected in soil (McKercher and Anderson, 1968a; Cosgrove and Irving, 1980; Turner, 2007).

In addition to previous literature (Turner et al., 2012), we were able to detect for the first time the two different conformations of *neo*-IP_6 and *chiro*-IP_6 (4-equatorial-2-axial and 2-equatorial-4-axial) in alkaline soil extracts. In contrast to *myo*-IP_6, which undergoes a conformational change at pH 9.4 caused by ring inversion (Isbrandt and Oertel, 1980) resulting in the detection of only one conformer in NaOH-EDTA soil extracts, both conformers of *neo*- and *chiro*-IP_6 can be measured in alkaline extracts (Turner et al., 2012). Usually, analysis of *chiro*- and *neo*-IP_6 in soil extracts relies on the detection of the 4-equatorial-2-axial (C1,3,4,6) peak of *neo*-IP_6 and the 2-equatorial-4-axial (C1,6) peak of *chiro*-IP_6 downfield of the orthophosphate peak, because the other peaks of these two compounds overlap with other peaks in the phosphomonoester region. However, by using different fractionation and purification techniques, i.e. hypobromite oxidation, size exclusion chromatography and the Humeomics method, we were able to resolve these peaks in the phosphomonoester region of solution ^{31}P NMR soil extracts. The origin of IP isomers other than *myo* in soil is still not fully explained due to the fact that plant residues only contain the *myo*- and, to a lesser extent, the *chiro*-isomer (Turner et al., 2002; L'Annunziata, 2007). The process of microbial epimerisation of the *myo*- to the *chiro*-isomer was reported by L'Annunziata et al. (1977), but it is not known yet if this process in the soil system also yields the *scyllo*- and *neo*-isomers (L'Annunziata, 2007).

The relatively high abundance of IP in soil is due to their strong complexation with soil constituents, with factors such as soil texture, pH, SOM content, climate and soil management/land use influencing their stability (Turner et al., 2002). Furthermore, the

spatial arrangement of the phosphate groups on the inositol ring can have major implications on the accumulation and turnover of IP. Cosgrove (1970a) reported that initial attack of phytase on the IP_6 molecule was apparently facilitated by the presence of an axial group. The author provided evidence that the reaction rate of the IP_6 hydrolysis by phytase was highest for the *myo* isomer, followed by *neo* and *chiro*, and was lowest for *scyllo*. Therefore, it is assumed that *myo*-IP_6 is less resistant to phytase hydrolysis compared to the other three isomers. Moreover, not only resistance to enzymatic attack but also sorption affinity on soil constituents could depend on the spatial arrangement of the phosphate groups (Anderson and Arlidge, 1962; Ognalaga et al., 1994; Celi and Barberis, 2007). Further research on the diversity, abundance and sorption behaviour of IP isomers and their conformers in soil is needed in order to understand their transformation and turnover in the plant-soil systems.

Lower order inositol phosphates

For the first time, we identified lower-order IP in soil extracts using solution ^{31}P NMR spectroscopy (chapter 2). We were able to detect *myo*-(1,2,4,5,6)-IP_5, *myo*-(1,3,4,5,6)-IP_5, *scyllo*-IP_5, *neo*-IP_5, *chiro*-IP_5 and *scyllo*-(1,2,3,4)-IP_4 by prior removal of peaks arising from compounds other than IP using hypobromite oxidation (chapter 2). The presence of lower-order IP in soil was already reported in the chromatographic era (McLaren et al., 2020) by various studies (Smith and Clark, 1951; Anderson, 1955; Cosgrove, 1963; McKercher and Anderson, 1968a, b; Halstead and Anderson, 1970; Anderson and Malcolm, 1974; Irving and Cosgrove, 1982) and more recently by Almeida et al. (2018) using chromatography.

Both isomers of *myo*-IP_5 (*myo*-(1,2,4,5,6)-IP_5 and *myo*-(1,3,4,5,6)-IP_5) result from enzymatic attack on IP_6 at positions other than C2 or C5 (Johnson and Tate, 1970; Cosgrove and Irving, 1980). The authors reported based on studies of Johnson and Tate (1970), Cosgrove (1970a, 1970b) and Irving and Cosgrove (1972) that D-*myo*-(1,2,4,5,6)-IP_5 was the major initial product of the IP_6 hydrolysis by microbial phytases. In contrast, *myo*-(1,3,4,5,6)-IP_5 was a minor hydrolysis product of IP_6 by wheat bran phytase (Lim and Tate, 1973). Further possible biotic and abiotic pathways in the generation of these two IP_5 compounds, which lead to a possible addition to the soil, are discussed in chapter 2. The other isomers of IP_5 detected in the course of this PhD project were reported to arise from acidic hydrolysis of the initial IP_6 compound as well as from enzymatic dephosphorylation of the initial IP_6 compound (Cosgrove, 1970a; Cosgrove and Irving, 1980). We propose that the latter process occurred in the soils sampled and led to the

detection of IP_5 compounds in our samples because we carried out an alkaline extraction. Furthermore, IP_6 and IP_5 are assumed to be resistant to hypobromite oxidation (Irving and Cosgrove, 1982).

Prior to this PhD project, lower-order IP had not yet been detected in solution ^{31}P NMR spectra on soil extracts. It was only hypothesised that unidentified sharp peaks in the phosphomonoester region could arise from lower-order IP (Turner et al., 2012). In contrast, Vestergren et al. (2012) and Vincent et al. (2013) concluded that the unidentified sharp signals in the phosphomonoester region were mostly due to degradation products of RNA and phospholipids, as well as other small cellular metabolites.

As discussed previously, we found that lower-order IP can comprise one third of the total IP pool measured by NMR and a plethora of unidentified sharp peaks (up to 35) in the phosphomonoester region presumably arise from other lower-order IP isomers and conformers, based on the assumption that other P_{org} compounds were oxidised using hypobromite oxidation (chapter 2). Turner et al. (2003a) carried out NaOH-EDTA extraction of 29 permanent pasture soils from England and Wales with subsequent P_{org} characterisation using solution ^{31}P NMR spectroscopy. The authors detected 16 peaks in the phosphomonoester region of the NMR soil spectra (ranging from δ 6.78 to 3.27 ppm). Similarly, Vestergren et al. (2012) reported 18 peaks in the phosphomonoester region of reference compounds added boreal forest humus soil extracts. In the same year, Turner et al. (2012) detected 15 sharp peaks arising from IP in hypobromite oxidised soil solutions. Later, Cade-Menun (2015) created a library containing solution ^{31}P NMR chemical shifts of standard compounds that could theoretically be present in soils. In this library, the authors reported the presence of 34 sharp peaks in the phosphomonoester region. In contrast to these studies, we were able to detect up to 70 sharp peaks in hypobromite oxidised soil extracts. In addition, several of the unidentified IP peaks occur in chemical shift regions of the peaks arising from RNA mononucleotides as well as α- and β-glycerophosphates, what may lead to a misassignment of these peaks in untreated soil extracts (chapter 2).

The information on the abundance and diversity of the IP pool acquired during this PhD project provides new insights into the IP_6 to lower-order IP ratio, which could serve in the future as an indicator for IP turnover and stabilisation in soil (chapter 2). It has been shown by Anderson and Arlidge (1962) that the degree of sorption decreases with decreasing number of phosphate groups bound to the IP ring. Furthermore, improved knowledge on the diversity of IP in soil will help to gain further insights into enzymatic processes. These

processes may influence the transformation of the IP pool either by dephosphorylation, by epimerisation or by phosphorylation (L'Annunziata and Gonzalez, 1977; L'Annunziata et al., 1977).

Our findings emphasise the importance of including lower-order IP into ^{31}P NMR soil research. We recommend that peak identification libraries such as the one of Cade-Menun (2015) should be updated by including lower-order IP. Spiking experiments using a diversity of lower-order IP standards should be carried out on untreated and, if possible, on hypobromite oxidised soil extracts to identify more of the plethora of sharp peaks in the phosphomonoester region. Furthermore, the effect of the hypobromite oxidation on these compounds has to be investigated in order to exclude a possible degradation of higher order IP to lower-order IP during the isolation and purification processes. The reason for this is that the steric hindrance and charge density of the phosphate groups bound to the inositol ring decrease with decreasing phosphorylation state. Therefore, the resistance to hypobromite oxidation of IP with a lower-order than IP_5 (Irving and Cosgrove, 1982) may decrease as well.

Cosgrove and Irving (1980) reported that the small amount of commercially available IP standards made it difficult to carry out accurate measurements of enzymatic degradation processes. This is similarly the case for characterisation of P_{org} using solution ^{31}P NMR spectroscopy. Therefore, the synthesis and availability of lower-order IP standards would greatly help advance the identification of IP and understand the mechanisms governing their flux in soil.

Inositol phosphates in different molecular size fractions

More than half (56 %) of total IP was detected in molecular size fractions above 5 kDa across four investigated soil samples (chapter 3). Pools of IP were most abundant in the 5-10 kDa fraction (33 % of total IP), followed by the above 70 kDa fraction (13 % of total IP). The other large molecular size fractions each contained less than 5 % of total IP.

Veinot and Thomas (1972), Hong and Yamane (1981) and other studies reported the presence of IP in large molecular size material of soil extracts using chromatography. In addition to these studies, we proved the existence of IP_6 as well as lower-order IP in various large molecular size fractions using a different analytical technique for speciation. High concentrations of IP in large molecular size fractions were most prominent in the two soil samples with high amounts of SOM and high concentrations of iron along the

molecular size distribution. These findings provide more evidence that IP in large molecular size material are in close association with SOM compounds (Moyer and Thomas, 1970; Omotoso and Wild, 1970; Borie et al., 1989). These associations could be held together via Fe bridges (Veinot and Thomas, 1972; Vincent et al., 2012), given the high abundance of the metal in the respective fractions (chapter 3). Another possible association mechanism of IP with the SOM could be via ester-linkages of the oxygen of the phosphate group with the carbon of an organic moiety, as it is the case for inositol-phospholipids (Hawthorne, 1982). This would also explain the large decrease in IP concentrations when ester bonds of the SOM structure are destroyed, as described in chapter 4. Another possibility of the occurrence of IP in large molecular size material could be IP complexed with nanoparticles of soil minerals that were not completely removed in the filtration and centrifugation process. As discussed in chapter 3, a spherical molecule with a molecular size of 70 kDa would exhibit a minimum diameter of approx. 5.5 nm, being in range of small Fe oxyhydroxides (Lead and Wilkinson, 2006).

We suggest three possible mechanisms responsible for the occurrence of IP_6 and lower-order IP in large molecular size material based on our findings and the work of Celi and Barberis (2007) and Jørgensen et al. (2015): 1) the association with SOM through Fe bridges or weak ester bonds; 2) the association with Fe minerals and/or amorphous Fe at nanoscale and 3) the association with clay minerals.

Pools of inositol phosphates in soil

In summary, two main pools of IP in soils can be distinguished: the first one being closely associated with the SOM structure and the second being closely associated with the soil mineral phase, possibly through Fe bridges and other adsorption mechanisms (chapters 3 and 4). These findings are in line with the study of Borie et al. (1989) and Jørgensen et al. (2015), who identified the association with the organic matrix, amorphous metal oxides or clay minerals as the main bindings sites for IP_6 in soils. Jørgensen et al. (2015) reported that more than 50 % of the *myo*- and *scyllo*-IP_6 were associated with organic matter in a tropical Oxisol, showing that association of IP with SOM is important even in soils with abundant amorphous Al and Fe oxides. In this PhD project, we provide additional evidence that a major pool of IP is incorporated in the SOM. Moreover, our results from chapter 3 and chapter 4 suggest that the second major pool of IP is associated via Fe bridges to the mineral phase and the SOM fraction resisting the Humeomics SCF extractions (Celi and Barberis, 2007; Jørgensen et al., 2015; Celi et al., 2020).

The bioavailability of P originating from IP relies on the mineralisation with phytases (George et al., 2007). However, the accessibility of IP for enzymatic attack is hindered by sorption and precipitation mechanisms. These mechanisms not only generally decrease the solubility of IP but also affect the charge density of the surface area of the IP-mineral complex and could change the conformation of the IP (Ognalaga et al., 1994; Celi et al., 1999; Giles et al., 2012). One goal of IP research is overcoming this reduced accessibility of the soil IP pool in order to increase its enzymatic hydrolysis. The resulting phosphate ions are potentially available for crops, what would decrease the need of additional mineral fertilisers. Furthermore, the engineered secretion of enzymes targeting highly abundant IP forms by transgenic plants could enhance the otherwise poor utilisation of the soil IP pool for plant nutrition (Richardson et al., 2001; Zimmermann et al., 2003). In the past, studies mostly focused on IP-mineral associations and precipitates rather than associations with SOM (Turner et al., 2002). In contrast, the findings of this PhD project provide evidence for including the IP pool associated with the SOM into P_{org} research, especially on highly organic soils, in order to better understand stabilising and incorporation mechanisms. Only when these mechanisms are better understood, strategies for more sustainable fertiliser practices can be developed that involve the enhanced accessibility of IP for biological hydrolysis.

Moreover, IP may play an important role in stabilising the SOM fraction closely associated with the soil mineral phase (chapter 4) (Tipping et al., 2016). However, this stabilisation function of the IP pool may be reduced by enhancing its accessibility for enzymatic hydrolysis and increase its mineralisation for P fertilisation purposes. Therefore, the role of IP in the soil C cycle and the stabilisation mechanisms involved have to be investigated in order to not adversely affect C sequestration efforts by increased IP mineralisation.

The unresolved organic P pool

The unresolved P_{org} pool represented by underlying broad signals in the phosphomonoester region of NMR soil spectra comprised 53 % to 79 % (on average 64 %) of total phosphomonoesters across all six soil samples (chapter 1). Similar or even higher proportions were reported in various studies: 56 % in a calcareous soil from Australia by Doolette et al. (2010), 61 % to 73 % across five diverse soil samples by McLaren et al. (2015b), 51 % to 77 % across 8 soil samples by Jarosch et al. (2015) and 65 % by McLaren et al. (2017). The chemical nature of soil P_{org} that remained unidentified was on average 45 % across a global compilation of NMR datasets by McLaren et al. (2020). Studies which

focus on the identification of small, identifiable molecules by fitting their peaks to the baseline of NMR soil spectra do not account for this unresolved P_{org} pool of apparent large molecular size (Turner et al., 2003c; Hill and Cade-Menun, 2009; Deiss et al., 2016).

This project confirms the findings of previous studies that the majority of P_{org} is represented by underlying broad signals in the phosphomonoester region of ^{31}P NMR spectra across soils of diverse origins, soil type and agricultural management practices. The advances in our knowledge on the chemical nature of this abundant P_{org} pool will be discussed in the following sections.

Structural composition of the underlying broad signals

The underlying broad signals in the phosphomonoester region are of homogeneous broadening and therefore not comprised of a range of sharp signals arising from small, identifiable molecules (chapter 2). These findings of McLaren et al. (2019) on NaOH-EDTA soil extracts and of Reusser et al. (2020b) on hypobromite oxidised soil extracts using T_2 NMR experiments provide further evidence that the unresolved P_{org} pool is chemically distinct from IP, which was already suggested by Smernik and Dougherty (2007) and Dougherty et al. (2007). Furthermore, the resistance of approx. half of the unresolved P_{org} to hypobromite oxidation showed that the underlying broad feature in solution ^{31}P NMR spectra is also not comprised of IP complexed with organic matter as proposed by Smernik and Dougherty (2007). This is due to the fact that hypobromite oxidation would destroy such complexes (Cosgrove, 1963; Irving and Cosgrove, 1981). Indeed, these complexes of IP with SOM appear to be an important pool of P_{org} (chapter 4), but this PhD project provides evidence that they are not represented by the underlying broad signals in the phosphomonoester region of NMR spectra.

The T_2 experiments further revealed that the underlying broad signal itself was comprised of more than one broad peak. These findings are based on theoretical calculations of the peak's line widths at half-height, which were considerably lower than the measured ones using SDF (McLaren et al., 2019). In chapter 2, only one underlying broad signal was fitted, resulting in a line width at half-height of on average 256.1 Hz. The calculated line width based on the T_2 times was 50-times narrower (5.2 Hz). In chapter 3, the fitting of on average three to four underlying broad signals in the SDF procedure resulted in an average line width at half-height of 88 Hz. Therefore, fitting several underlying broad signals better approximates the calculated line widths reported by McLaren et al. (2019) (11 Hz) and Reusser et al. (2020b). Further research is needed in order to fully understand the

structural composition of the broad signals by combining T_2 experiments with simulations of different amounts of underlying broad signals with varying peak distributions (Lorentzian vs. Gaussian).

The T_2 experiments did not only provide insights into the composition of the underlying broad signals but also into the molecular size of the compounds causing these signals. The T_2 times are inversely related to a molecule's tumbling (Claridge, 2016b). It is assumed that more rigid molecules tumble less compared to less rigid ones. Furthermore, the rigidity of a molecule can be related to its molecular size, meaning that a large molecule is considered to exhibit a higher rigidity compared to a molecule of small molecular size (Claridge, 2016b; McLaren et al., 2019). Therefore, it is considered that the shorter the T_2 time is, the larger is the molecular size of a compound. The significant shorter T_2 times of the underlying broad signals compared to IP confirm the results of McLaren et al. (2015b) and Jarosch et al. (2015), who reported the large molecular size and complex structural nature of the underlying broad signal by using a different analytical approach. The large molecular size nature of the compounds causing these signals was then further investigated in chapter 3 of this thesis using size-exclusion chromatography and discussed in the next section.

The molecular size distribution of the unresolved organic P pool

The unresolved P_{org} pool represented by underlying broad signals is comprised of a range of molecules in a continuum of molecular sizes rather than a macropolymer at distinct molecular size (chapter 3). The theory of the macropolymeric nature of the unresolved P_{org} pool was based on the macropolymeric model of humic substances (Kononova et al., 1961) and findings of Swift and Posner (1972) as well as Moyer and Thomas (1970) that the P containing organic macromolecules were derived from humification processes (McLaren et al., 2020). The macropolymeric nature of the unresolved P_{org} pool was further proposed by Doolette et al. (2010), McLaren et al. (2014) and McLaren et al. (2015b). The results of this thesis advance these findings by revealing that the unresolved P_{org} is comprised of an association of smaller molecules with SOM (chapter 4) and that these associations dominate the P_{org} size fractions of 10-20 and 20-50 kDa. These associations are also present in smaller (<10 kDa) and higher (>50 kDa) fractions but along with other, identifiable P_{org} compounds, e.g. IP, other phosphomonoesters and phosphodiesters. Furthermore, large molecular size material of soil extracts does not only consist of the

unresolved P_{org} pool but also large biopolymers containing P, e.g. DNA and associations of IP with SOM and the soil mineral phase (chapter 3 and 4).

Nevertheless, the unresolved P_{org} pool represented by broad signals comprised from 44 % (<5 kDa) up to 77 % (20-50 kDa) of total phosphomonoesters and on average 54 % of total P over all size fractions. Therefore, these unresolved P associations exhibit a major pool of P_{org} across all molecular size fractions and need to be further investigated in order to evaluate their importance in the soil P cycle. Chapter 4 provides evidence that the associations are strongly coupled with the SOM, as already proposed by various studies (Smernik and Dougherty, 2007; Bünemann et al., 2008b; Doolette et al., 2011a; McLaren et al., 2015b; McLaren et al., 2020).

Associations of the unresolved organic P pool with SOM

More than two thirds of the unresolved P_{org} pool was associated with the SOM structure (chapter 4). The close association of the unresolved P_{org} pool with the SOM has been suggested in various studies using ^{31}P NMR spectroscopy based on its apparent large molecular size and complex structure in humic extracts (He et al., 2006; Smernik and Dougherty, 2007; Jarosch et al., 2015; McLaren et al., 2015b). McLaren et al. (2015b) hypothesised that the association occurs via ester linkages because of the appearance of the underlying broad signal in the phosphomonoester region of ^{31}P NMR soil spectra. In this thesis, we provide evidence that most (40 %) of the unresolved P_{org} pool is associated with the SOM through weak ester bonds. Another proportion (22 %) appears to be unbound and the remaining proportion in the residue (30 %) appears to be closely associated with the mineral phase, possibly through Fe bridges or sorption processes (chapter 3). The association with the SOM superstructure through weak ester bonds explains the appearance of this P_{org} pool in large molecular size material and its representation by underlying broad signals in the phosphomonoester region.

In this PhD project, the associations of the unresolved P_{org} pool with the SOM were investigated by combining the Humeomics chemical fractionation procedure with P analyses (chapter 4). It is the first time that P concentrations were measured in the different fractions and residues of the Humeomics procedure. However, the study is based on only one soil sample with high organic matter and P contents due to the time-consuming procedure. Future studies involving Humeomics on soil samples of diverse SOM and P contents and compositions are necessary to confirm our results on the association of the unresolved P_{org} pool with the SOM. Furthermore, the effects of each fractionation step on

pure P_{org} compounds and their recovery should be studied, especially whether the innate ester bond between the phosphates and the inositol ring are stable to the BF_3 extraction or not. In our study, we used microwave digestion to assess the total P contents in the fractions. The organic solvent fractions could not be completely dried and therefore complicated the sample preparation for the analyses. Hence, better suited analytical techniques for total P measurements in organic extracts should be tested, e.g. solution ^{31}P NMR spectroscopy on organic solvent extracts. We carried out solution ^{31}P NMR spectroscopy on the organic fractions of the Humeomics procedure (chapter 4) dissolved in deuterated chloroform and spiked with triphenylphosphate. First results are promising, but the NMR experiment and sample preparation have to be optimised. Furthermore, known chemical shift ranges of specific P_{org} compounds analysed in NaOH-EDTA cannot be directly transferred to the organic solvent extract due to the differing matrix. Hence, new chemical shift libraries of P_{org} standards soluble in organic solvents need to be elaborated and the NMR observability of these P_{org} standards tested.

The findings on the association of the unresolved P_{org} pool with the SOM and the mineral phase will help to improve our understanding on possible processes involved in the formation of this abundant P pool. McLaren et al. (2020) hypothesised that the polymerisation of an array of P_{org} degradation products could result in the formation of the unresolved P_{org} pool. Our findings on the molecular size distribution and the associations of P_{org} compounds with the SOM indicate that this process, if at all, only plays a minor role in the formation of the unresolved P pool. Another proposed formation mechanism of phosphomonoester-like compounds is the phosphorylation of organic moieties by abiotic processes (Auten, 1923; Brannon and Sommers, 1985; Kamerlin et al., 2013; McLaren et al., 2020). Indeed, findings of chapters 3 and 4 on the observed association of the unresolved P_{org} pool with the SOM structure via ester linkages could indicate that these abiotic phosphorylation processes of SOM compounds could result in the formation of the unresolved P_{org} pool. This would also be supported by the close coupling of this P_{org} pool with the soil C fractions (chapter 4). Another explanation on the formation of the unresolved P_{org} pool could be that P containing compounds of known chemical nature and originating from plant residues and/or microbial turnover products are incorporated in the SOM mainly via ester bonds.

Being part of the complex SOM structure and the association with the mineral phase can limit the accessibility for biological and chemical degradation, similarly to the IP pool. This could explain the observed resistance of the unresolved P_{org} pool to enzymatic attack

(Jarosch et al., 2015), its chemical stability (chapter 3) and stabilisation resp. accumulation in soil (Annaheim et al., 2015; McLaren et al., 2017).

The improved knowledge on formation, association and stabilisation processes of the unresolved P_{org} pool acquired during this PhD project helps to better target future research projects on this highly abundant P_{org} pool under field conditions in order to better understand its turnover in terrestrial as well as possibly aquatic ecosystems. In addition, when these processes on the formation and stabilisation are known, the reversal processes which lead to destabilisation and potential release of P can be evaluated. This reversal processes could possibly be enhanced by specific soil management practices, the amendment of enzymes or other compounds, or the engineering of transgenic plants releasing compounds targeting this P_{org} pool. Therefore, understanding the formation, association and stabilisation processes provides the base to evaluate the contribution of this P_{org} pool towards more sustainable fertilisation practices.

Other organic P compounds

In general, the abundance of other P_{org} compounds was low compared to the unresolved P_{org} and IP pool. Other identified phosphomonoesters, such as α- and β-glycero-phosphates as well as glucose-6-phosphate, comprised 5 % of total P_{org} (excluding pyrophosphates) in the Gleysol (chapter 4). Similarly, phosphodiesters and phosphonates combined contributed on average 6 % to the total P_{org} pool across all six soil samples (chapter 1). Despite their low abundance, these compounds should not be neglected in speciation studies of soil P_{org}. Phosphomonoesters, -diesters and phosphonates are considered to be biologically relevant and are more labile in soil compared to the IP and unresolved P_{org} pool (McLaren et al., 2017; George et al., 2018; McLaren et al., 2020).

Phosphonates

Phosphonates can either be present as free molecules or bound to lipids or other macromolecular structures in living systems of animals, plants, bacteria and fungi (Hilderbrand, 1983), as component of herbicides i.e. glyphosate or result as a recalcitrant by-product of the sarin manufacture (Baldwin et al., 2005). In soils, phosphonates are assumed to originate from microbes and accumulate in cold, moist and acidic soils with limited microbial activity (Tate and Newman, 1982). In a global compilation of 204 samples, McLaren et al. (2020) reported that phosphonates comprised on average 2 % of total P_{org}

in soil, hence being the third most abundant form of P_{org} in soil after phosphomonoesters (80 %) and -diesters (18 %). Phosphonates were generally not only detected in solution ^{31}P NMR spectra of whole soil extracts (Newman and Tate, 1980; Cade-Menun et al., 2002; Turner et al., 2003b) but also in the soil humic acid fractions (Ogner, 1983; Bedrock et al., 1994; He et al., 2006). Besides that, little is known about this soil P_{org} compound.

In this PhD project, phosphonates were present in all soils and partially associated with the SOM (chapters 1 and 4). However, only one peak in the phosphonate region could be assigned to a naturally occurring compound of known chemical composition: 2-aminoethylphosphonic acid (chapter 1). Relative concentrations of phosphonates to total measured P (chapter 1) were highest in the Cambisol under forest (1.5 %), followed by the Ferralsol and Vertisol (1 % each). This distribution is dissimilar to other P_{org} compounds, such as phosphomonoesters, whose relative concentrations are lowest in the Ferralsol and Vertisol. These findings suggest that other factors and mechanisms influence the cycling and turnover of phosphonates in the soil-plant system compared to other P_{org} compounds.

Other phosphomonoester

Other phosphomonoesters include glucose-6-phosphate, α- and β-glycerophosphate as well as RNA mononucleotides, although these compounds originate from the alkaline hydrolysis of phosphodiesters (Tate and Newman, 1982; Doolette et al., 2009; Vestergren et al., 2012). α- and β-glycerophosphate originate from phospholipids, whose thicknesses can exceed 2 nm resp. 45 Å of the bilayer in the cell membrane (Lis et al., 1982). The large size of the phospholipids could explain why the two hydrolysis products were detected in large molecular size material in chapter 3. However, another explanation could be the close association of phospholipids with the SOM described in chapter 4. The hydrophobic tails of the phospholipids could contribute to the stabilisation of the large sized SOM associations by separating them from the aqueous solution (Piccolo et al., 2019). Nevertheless, the Humeomics fractionation procedure does not distinguish between SOM and microbial biomass, but the investigated soil in chapter 4 does contain a significant amount of P in form of microbial P (213 mg P/kg_{soil}). Hence, the removal of DNA, phospholipids and glucose-6-phosphate with the Humeomics fractionation procedure could possibly also result from the co-occurring destruction and extraction of microbial biomass. More studies investigating the effect of Humeomics on the soil microbial biomass are necessary.

The partially hydrophobic nature of phospholipids raised the questions whether these compounds could be fully extracted from soil by the aqueous NaOH-EDTA extraction procedure for solution ^{31}P NMR spectroscopy or not. Within the frame of this PhD project, Steck (2018) carried out Soxhlet extractions of a soil sample based on the method of Makarov et al. (2002) in order to assess the abundance of phospholipids in soil and compare the extraction efficacy of this method with the traditional NaOH-EDTA extraction. The author used a mixture of methanol and chloroform for the Soxhlet extraction with subsequent NaOH-EDTA extraction and solution ^{31}P NMR spectroscopy of the untreated and residual soil samples after extraction as well as the redissolved extract of the Soxhlet extraction. The author reported that the Soxhlet extract redissolved in NaOH-EDTA contained α- and β-glycerophosphate. However, there was no evident decrease of these two compounds between the NaOH-EDTA extracts of the untreated soil sample and the soil residue after Soxhlet extraction, suggesting that a different pool of phospholipids was extracted with the organic solvents compared to the NaOH-EDTA extraction. Therefore, we assume that phospholipids in soils are not fully recovered in NaOH-EDTA extracts. One disadvantage of the study was that the dried Soxhlet extracts were redissolved in aqueous NaOH-EDTA in order to compare the NMR spectra with the NaOH-EDTA soil extracts. This procedure resulted in hydrolysis of phospholipids and possibly incomplete dissolution of the hydrophobic components. Further research is needed using organic solvent extractions of soils with subsequent solution ^{31}P NMR spectroscopy directly on the organic solvent extracts in order to test if the extent and diversity of the phospholipid pool in soil is underestimated using ^{31}P NMR spectroscopy of NaOH-EDTA soil extracts.

Phosphodiesters

DNA, intermediate hydrolysis products of RNA and phospholipids (Makarov et al., 2002) with chemical shift ranges δ 2.5 to -2.4 ppm were summarised and reported as phosphodiesters in this PhD project. The NaOH-EDTA extracts of the Vertisol and the calcareous Cambisol did not contain any phosphodiesters (chapter 1). In the other four soils, concentrations of phosphodiesters ranged from 3 mg P/kg$_{soil}$ (Ferralsol) to 28 mg P/kg$_{soil}$ (Cambisol under forest), comprising on average less than 4 % of total P_{org}.

As reported in chapter 3, phosphodiesters were most abundant in molecular size fractions above 10 kDa. This is not surprising, for example the genome of *E. coli* can reach a length of 1.58 mm when unfolded (Madigan et al., 2012). The HPLC chromatogram of a DNA standard revealed that P was present along the whole molecular size range of the column,

indicating the presence of a diversity of DNA fragments of varying molecular sizes (chapter 3). The removal of most phosphodiesters by the Humeomics sequential extraction further suggest a strong association of DNA and other phosphodiesters with the SOM (chapter 4). Moreover, the association with the less 'persistent' SOM pool and the chemical instability to hypobromite oxidation underline the more labile nature and hence higher potential bioavailability of phosphodiesters compared to IP and the unresolved P_{org} pool in soil (Darch et al., 2014).

Phosphotriesters

In phosphotriesters, three oxygen atoms of the phosphate group are attached to an organic moiety. Examples for phosphotriesters are triisopropylphosphate, which is used as a warfare agent and pesticide (Ihdene et al., 2011), and *tris-p*-nitrophenyl phosphate, a substrate used for detection of phosphatase activity in soil (Eivazi and Tabatabai, 1977). By studying the hydrolysis of the latter compound in soil, Eivazi and Tabatabai (1977) determined the activity of phosphotriesterase in incubated soil. The activity of these enzymes were lower compared to phosphomono- and phosphodiesterase but measurable, suggesting that phosphotriesters could be present in soil. In contrast to the other enzymes, the phosphotriesterase activity increased by the addition of toluene, because the dissolution of the substrate *tris-p*-nitrophenyl phosphate was increased, which is insoluble in water. This insolubility in water could also be the reason why phosphotriesters have not yet been detected in NaOH-EDTA soil extracts using ^{31}P NMR spectroscopy (Vestergren et al., 2012; Cade-Menun, 2015; McLaren et al., 2020). In a pilot study, we carried out extraction of phospholipids from the Gleysol sample using a mixture of methanol, chloroform and a citrate buffer according to the methods of Bligh and Dyer (1959) and Frostegård et al. (1991). The dried lipid fraction was then redissolved in deuterated chloroform and analysed by ^{31}P NMR spectroscopy (Appendix 3). The spectrum revealed two dominant peaks at chemical shift regions δ -12 to -14 ppm, which could arise from a phosphotriester (Pretsch et al., 2010). However, this has to be confirmed using spiking experiments and HSQC NMR experiments in order to elucidate the chemical structure of these compounds. Furthermore, their origin from the soil sample needs to be tested, since phosphotriesters are also used as additives for lubricants in laboratory equipment.

Perspectives

This PhD project provides new insights into the abundance, chemical nature and associations of the two major P_{org} pools and other P_{org} compounds found in soil. However, further research is needed linking the chemical nature of P_{org} with biogeochemical processes in order to assess the transformation, cycling and turnover of these compounds in soil systems as well as their transfer to aquatic ecosystems. In addition, the findings of this PhD project reveal gaps in the following research topics, which should be addressed in future projects:

- The soil IP pool is much more diverse and abundant in soils as previously thought. There is likely an underestimated pool of IP, which appears to be mostly associated with the mineral phase. The findings of this PhD project about the detection of various isomers of IP and lower-order IP in soil should be verified in future studies, involving spiking of a wide range of different IP standards. Furthermore, differences in the sorption and association behaviour of different IP compounds in soil should be assessed in order to gain a more differentiated picture of the accumulation and stabilisation of IP in soil.

- Detection of the unresolved P_{org} pool in soil relies on wet chemistry methods with previous extraction of the P_{org} pool. However, as reported by George et al. (2018), it is critical to develop non-destructive methods without the need of prior extraction, such as solid state NMR, visible near-infrared reflectance spectroscopy (VNIRS) or X-ray absorption near edge structure spectroscopy (XANES). These non-destructive methods should include the detection of the unresolved P_{org} pool as well as the diversity of the IP pool and not only phytate (Negassa et al., 2010; Colocho Hurtarte et al., 2020). The combination of solution ^{31}NMR spectroscopy of soil extracts with these non-destructive methods would allow for a more comprehensive evaluation of the structural composition and abundance of P_{org} in soil.

- A majority of P_{org} is associated with the SOM. However, we only investigated one soil sample using one fractionation method of the SOM. Future research on the association of P_{org} with SOM should be carried out on soils of varying organic matter and P composition. Furthermore, different fractionation techniques of the SOM should be compared (e.g. Humeomics and the biomarker method of Simpson and

Simpson (2012)) with regard to assessing P contents, speciation and recoveries of P standard compounds in the different SOM fractions.

- The utilisation of the highly abundant and diverse IP pool as well as the unresolved P_{org} pool in soil for plant nutrition needs further attention. Studies on soil amendments and different soil management practices that could influence the accessibility and subsequent chemical and/or biological degradation of these P_{org} pools are recommended, as well as research on transgenic plants releasing compounds that target these pools. However, at the same time, the role of especially IP in stabilising the SOM needs to be investigated. An enhanced degradation of the IP pool for plant nutrition could lead to a destabilisation of the SOM, which in turn negatively affects soil quality and efforts for atmospheric C sequestration. In contrast, IP addition could help to improve soil quality and slow down degradation by stabilising the SOM.

- Phospholipids may not be completely recovered using NaOH-EDTA extraction prior to solution ^{31}NMR spectroscopy. A promising method appears to be the combination of organic solvent extraction, e.g. Soxhlet extraction with a methanol-chloroform mixture described by Makarov et al. (2002), with subsequent analysis of the organic solvent extract dissolved in $CDCl_3$ using solution ^{31}P NMR spectroscopy. Due to the absence of peaks arising from water-soluble P_{org} compounds, better resolution and possible application of 2D- and 3D-NMR spectroscopy for speciation of the phospholipids could be applicable.

- The longitudinal (or spin-lattice) relaxation times T_1 do not only vary markedly between different samples but also depend on the matrix of a sample and the chemical environment of each P nuclei within a molecule. Hence, the T_1 time differs for each P compound and is usually significantly longer for orthophosphate compared to IP (Appendix 4). If the T_1 time of an NMR experiment is not chosen long enough for complete relaxation of each P compound, the net peak area will not be completely ‘evolved’, resulting in an inaccurate quantification of this compound (Cade-Menun et al., 2002). In this thesis, we carried out inversion recovery experiments for each sample prior to the main NMR experiment in order to assess the compound with the longest T_1 time (McDowell et al., 2006; McLaren et al., 2019). Another approach was introduced by McDowell et al. (2006), who

reported that the ratio of total P to the total Mn and Fe concentrations in an extract can be used for quick estimation of the T_1 time of an extract. However, this relationship between T_1 and the element concentrations is mostly based on the regression relationship of phosphomonoesters (r^2=0.7), whose slope is significantly less compared to the generated orthophosphate relationship (McDowell et al., 2006). Hence, further research on the relationship of the T_1 time with P, Mn and Fe concentrations in a sample is needed to assess whether this simplified method results in accurate quantification of all P compounds in a sample or not. The studies should further compare these relationships with actual T_1 measurements of each P compound in a sample using inversion recovery experiments.

- The substitution of one or more oxygen atoms with the ^{18}O isotope in the phosphate molecule leads to an isotopic effect in NMR spectroscopy, resulting in a chemical shift change, which can be observed (Sorensen-Stowell and Hengge, 2005). This isotopic effect not only allows for investigating hydrolysis effects on P_{org} compounds induced by NaOH-EDTA extraction (Wang et al., 2021), but also for tracing experiments based on the incorporation of ^{18}O into specific P_{org} compounds, e.g. into the cell membrane represented by phospholipids. Cohn and Hu (1978) used this isotopic effect on labelled phosphate groups in order to study enzyme-catalysed phosphate-phosphate exchange reactions. The authors suggested that this labelling method combined with ^{31}P NMR spectroscopy could be used to investigate the phosphorylation rate of e.g. carboxylic acid or to label the phosphate groups of ATP. However, a high concentration of the ^{18}O containing P compound as well as sufficient peak resolution are prerequisites for the observation of the isotopic effect in solution ^{31}P NMR spectra. For example, the many overlapping peaks in the phosphomonoester region of alkaline soil extracts could hinder the detection of a small peak arising from a phosphomonoester compound labelled with ^{18}O. In order to overcome this problem, selective extraction of labelled P_{org} compounds from soil with subsequent solution ^{31}P NMR spectroscopy would be promising. One of these selective extractions could be the phospholipid extraction previously discussed and described in Appendix 3.
- The physicochemical properties of a compound are defined by its chemical nature. Hence, changes in soil properties, land management practices, climatic conditions etc. will affect P_{org} compounds differently depending on their chemical nature. However, there are still many uncertainties about the chemical composition of the ‘unresolved’ P_{org} pool, which represents a major fraction of total soil P_{org}. Further

research is needed to unravel the chemical composition of soil P_{org} in order to better understand the transformation and cycling behaviour of individual P_{org} compounds in terrestrial and aquatic ecosystems.

Conclusions

The molecular size of P_{org} present in soils ranges from small identifiable biomolecules, e.g. *scyllo*-IP_4, up to large associations with the SOM and mineral phase. The majority of soil P_{org} can be divided into two pools: the pool of IP and the 'unresolved' P_{org} pool. Our findings on the first pool provide evidence that it is much more diverse and abundant than previously thought. For the first time, we were able to identify six different lower-order IP in solution ^{31}P NMR spectra on soil extract. Additionally, up to 35 unidentified sharp signals in the phosphomonoester region of hypobromite oxidised soil extracts are presumably arising from other, so far unidentified IP forms. Furthermore, investigations on the P_{org} molecular size distribution and SOM fractionation revealed that a significant proportion of IP is associated with the SOM through ester linkages and present in large molecular size material. However, this large molecular size material was mostly dominated by the unresolved P_{org} pool, which comprised on average 77 % resp. 74 % of the total phosphomonoester concentration in the 10-20 and 20-50 kDa fractions. Our findings further indicate that the unresolved P_{org} pool does not consist of a single macropolymeric structure at distinct molecular size but rather a range of molecules of different molecular sizes and closely associated with the SOM through ester and ether linkages. These P_{org}-SOM associations generated on average three to four underlying broad signals in the phosphomonoester region of solution ^{31}P NMR soil spectra and comprised 53 % to 79 % of total phosphomonoesters.

The findings of this PhD project not only improved our understanding on the structural composition and association of the IP and unresolved P_{org} pool but also provide information on possible formation and stabilisation processes. Furthermore, the project reveals that the chemical nature of P_{org} in soil, especially the IP and unresolved P_{org} pool, is much more complex and diversified than previously thought. The key challenge of future research is to capture and acknowledge the complexity and diversity of P_{org} in soil, and to link these compounds with physical-, chemical-, and biogeochemical processes in the plant-soil system.

References

Almeida, D.S., Menezes-Blackburn, D., Turner, B.L., Wearing, C., Haygarth, P.M., Rosolem, C.A., 2018. *Urochloa ruziziensis* cover crop increases the cycling of soil inositol phosphates. Biology and Fertility of Soils 54, 935-947.

Anderson, G., 1955. Paper chromatography of inositol phosphates. Nature 175, 863-864.

Anderson, G., 1980. Assessing organic phosphorus in soils, In: Khasawneh, F.E., Sample, E.C., Kamprath, E.J. (Eds.), The role of phosphorus in agriculture. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI, pp. 411-431.

Anderson, G., Arlidge, E.Z., 1962. The adsorption of inositol phosphates and glycerophosphate by soil clays, clay minerals, and hydrated sesquioxides in acid media. Journal of Soil Science 13, 216-224.

Anderson, G., Malcolm, R.E., 1974. The nature of alkali-soluble soil organic phosphates. Journal of Soil Science 25, 282-297.

Anderson, G., Williams, E.G., Moir, J.O., 1974. A comparison of the sorption of inorganic orthophosphate and inositol hexaphosphate by six acid soils. Journal of Soil Science 25, 51-62.

Angyal, S.J., 1963. Chapter VIII - Cyclitols, In: Florkin, M., Stotz, E.H. (Eds.), Comprehensive Biochemistry. Elsevier, pp. 297-303.

Annaheim, K.E., Doolette, A.L., Smernik, R.J., Mayer, J., Oberson, A., Frossard, E., Bünemann, E.K., 2015. Long-term addition of organic fertilizers has little effect on soil organic phosphorus as characterized by ^{31}P NMR spectroscopy and enzyme additions. Geoderma 257-258, 67-77.

Auten, J.T., 1923. Organic phosphorus of soils. Soil Science 16, 281.

Baker, R.T., 1976. Changes in the chemical nature of soil organic phosphate during pedogenesis. Journal of Soil Science 27, 504-512.

Baldwin, D.S., Howitt, J.A., Beattie, J.K., 2005. Abiotic degradation of organic phosphorus compounds in the environment, In: Turner, B.L., Frossard, E., Baldwin, D.S. (Eds.), Organic phosphorus in the environment. CABI Publishing, Wallingford, UK, pp. 75-88.

Batjes, N.H., 1996. Total carbon and nitrogen in the soils of the world. European Journal of Soil Science 47, 151-163.

Bedrock, C.N., Cheshire, M.V., Chudek, J.A., Goodman, B.A., Shand, C.A., 1994. Use of ^{31}P-NMR to study the forms of phosphorus in peat soils. Science of The Total Environment 152, 1-8.

Bligh, E.G., Dyer, W.J., 1959. A rapid method of total lipid extraction and purification. Canadian Journal of Biochemistry and Physiology 37, 911-917.

Bloembergen, N., Purcell, E.M., Pound, R.V., 1948. Relaxation effects in nuclear magnetic resonance absorption. Physical Review 73, 679-712.

Blume, H.P., Brümmer, G.W., Scheffer, F., Horn, R., Kandeler, E., Schachtschabel, P., Kögel-Knabner, I., Welp, G., Kretzschmar, R., Thiele-Bruhn, S., 2009. Scheffer/Schachtschabel: Lehrbuch der Bodenkunde. Spektrum Akademischer Verlag.

Borie, F., Zunino, H., Martínez, L., 1989. Macromolecule-P associations and inositol phosphates in some chilean volcanic soils of temperate regions. Communications in Soil Science and Plant Analysis 20, 1881-1894.

Bowman, R.A., Moir, J.O., 1993. Basic EDTA as an extractant for soil organic phosphorus. Soil Science Society of America Journal 57, 1516-1518.

Brannon, C.A., Sommers, L.E., 1985. Preparation and characterization of model humic polymers containing organic phosphorus. Soil Biology and Biochemistry 17, 213-219.

Bühler, S., Oberson, A., Sinaj, S., Friesen, D.K., Frossard, E., 2003. Isotope methods for assessing plant available phosphorus in acid tropical soils. European Journal of Soil Science 54, 605-616.

Bünemann, E.K., Augstburger, S., Frossard, E., 2016. Dominance of either physicochemical or biological phosphorus cycling processes in temperate forest soils of contrasting phosphate availability. Soil Biology and Biochemistry 101, 85-95.

Bünemann, E.K., Oberson, A., Frossard, E., 2010. Phosphorus in action: biological processes in soil phosphorus cycling. Berlin, Heidelberg : Springer.

Bünemann, E.K., Smernik, R.J., Doolette, A.L., Marschner, P., Stonor, R., Wakelin, S.A., McNeill, A.M., 2008a. Forms of phosphorus in bacteria and fungi isolated from two Australian soils. Soil Biology and Biochemistry 40, 1908-1915.

Bünemann, E.K., Smernik, R.J., Marschner, P., McNeill, A.M., 2008b. Microbial synthesis of organic and condensed forms of phosphorus in acid and calcareous soils. Soil Biology and Biochemistry 40, 932-946.

Cade-Menun, B., Liu, C.W., 2014. 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 78, 19-37.

Cade-Menun, B.J., 2005. Characterizing phosphorus in environmental and agricultural samples by ^{31}P nuclear magnetic resonance spectroscopy. Talanta 66, 359-371.

Cade-Menun, B.J., 2015. Improved peak identification in ^{31}P-NMR spectra of environmental samples with a standardized method and peak library. Geoderma 257-258, 102-114.

Cade-Menun, B.J., Liu, C.W., Nunlist, R., McColl, J.G., 2002. Soil and litter phosphorus-31 nuclear magnetic resonance spectroscopy. Journal of Environmental Quality 31, 457-465.

Cade-Menun, B.J., Preston, C.M., 1996. A comparison of soil extraction procedures for ^{31}P NMR spectroscopy. Soil Science 161.

Celi, L., Barberis, E., 2005. Abiotic stabilisation of organic phosphorus in the environment, In: Turner, B.L., Frossard, E., Baldwin, D.S. (Eds.), Organic phosphorus in the environment. CABI Publishing, Wallingford, UK, pp. 113-132.

Celi, L., Barberis, E., 2007. Abiotic reactions of inositol phosphates in soil, In: Turner, B.L., Richardson, A.E., Mullaney, E.J. (Eds.), Inositol phosphates: linking agriculture and the environment. CABI, Wallingford, pp. 207-220.

Celi, L., Lamacchia, S., Barberis, E., 2000. Interaction of inositol phosphate with calcite. Nutrient Cycling in Agroecosystems 57, 271-277.

Celi, L., Lamacchia, S., Marsan, F.A., Barberis, E., 1999. Interaction of inositol hexaphosphate on clays: adsorption and charging phenomena. Soil Science 164.

Celi, L., Prati, M., Magnacca, G., Santoro, V., Martin, M., 2020. Role of crystalline iron oxides on stabilization of inositol phosphates in soil. Geoderma 374, 114442.

Chung, S.-K., Kwon, Y.-U., Chang, Y.-T., Sohn, K.-H., Shin, J.-H., Park, K.-H., Hong, B.-J., Chung, I.-H., 1999. Synthesis of all possible regioisomers of *scyllo*-inositol phosphate. Bioorganic & Medicinal Chemistry 7, 2577-2589.

Claridge, T.D.W., 2016a. Chapter 1 - Introduction, In: Claridge, T.D.W. (Ed.), High-resolution NMR techniques in organic chemistry (Third Edition). Elsevier, Boston, pp. 1-10.

Claridge, T.D.W., 2016b. Chapter 2 - Introducing High-Resolution NMR, In: Claridge, T.D.W. (Ed.), High-Resolution NMR techniques in organic chemistry, 3 ed. Elsevier, Boston, pp. 11-59.

Cohn, M., Hu, A., 1978. Isotopic (^{18}O) shift in ^{31}P nuclear magnetic resonance applied to a study of enzyme-catalyzed phosphate-phosphate exchange and phosphate (oxygen)-water exchange reactions. Proceedings of the National Academy of Sciences of the United States of America 75, 200-203.

Colocho Hurtarte, L.C., Santana Amorim, H.C., Kruse, J., Criginski Cezar, J., Klysubun, W., Prietzel, J., 2020. A novel approach for the quantification of different inorganic and organic phosphorus compounds in environmental samples by P L2,3-Edge X-ray absorption near-edge structure (XANES) spectroscopy. Environmental Science & Technology 54, 2812-2820.

Condron, L.M., Turner, B.L., Cade-Menun, B.J., 2005. Chemistry and dynamics of soil organic phosphorus, In: Sims, J.T., Sharpley, A.N. (Eds.), Phosphorus: agriculture and the environment. American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, Madison, WI, pp. 87-121.

Cosgrove, D., 1963. The chemical nature of soil organic phosphorus. I. Inositol phosphates. Soil Research 1, 203-214.

Cosgrove, D.J., 1969. The chemical nature of soil organic phosphorus. II. Characterization of the supposed dl-chiro-inositol hexaphosphate component of soil phytate as d-chiro-inositol hexaphosphate. Soil Biology and Biochemistry 1, 325-327.

Cosgrove, D.J., 1970a. Inositol phosphate phosphatases of microbiological origin. Inositol phosphate intermediates in the dephosphorylation of the hexaphosphates of *myo*-inositol, *scyllo*-inositol, and D-*chiro*-inositol by a bacterial (Pseudomonas sp.) phytase. Australian Journal of Biological Sciences 23, 1207-1220.

Cosgrove, D.J., 1970b. Ion-exchange chromatography of inositol polyphosphates. Annals of the New York Academy of Sciences 165, 677-686.

Cosgrove, D.J., 1977. Microbial transformations in the phosphorus cycle, In: Alexander, M. (Ed.), Advances in Microbial Ecology. Springer US, Boston, MA, pp. 95-134.

Cosgrove, D.J., Irving, G.C.J., 1980. Inositol phosphates: their chemistry, biochemistry and physiology. Amsterdam: Elsevier.

Costello, A.J.R., Glonek, T., Myers, T.C., 1976. ^{31}P nuclear magnetic resonance-pH titrations of *myo*-inositol hexaphosphate. Carbohydrate Research 46, 159-171.

Dalal, R.C., 1977. Soil organic phosphorus, In: Brady, N.C. (Ed.), Advances in Agronomy. Academic Press, pp. 83-117.

Darch, T., Blackwell, M.S.A., Hawkins, J.M.B., Haygarth, P.M., Chadwick, D., 2014. A meta-analysis of organic and inorganic phosphorus in organic fertilizers, soils, and water: implications for water quality. Critical Reviews in Environmental Science and Technology 44, 2172-2202.

Deiss, L., de Moraes, A., Dieckow, J., Franzluebbers, A.J., Gatiboni, L.C., Sassaki, G.I., Carvalho, P.C.F., 2016. Soil phosphorus compounds in integrated crop-livestock systems of subtropical Brazil. Geoderma 274, 88-96.

Dell'Aquila, C., Neal, A.L., Shewry, P.R., 2020. Development of a reproducible method of analysis of iron, zinc and phosphorus in vegetables digests by SEC-ICP-MS. Food Chemistry 308, 125652.

Doolette, A.L., Smernik, R.J., 2011. Soil organic phosphorus speciation using spectroscopic techniques, In: Bünemann, E., Oberson, A., Frossard, E. (Eds.), Phosphorus in action: biological processes in soil phosphorus cycling. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 3-36.

Doolette, A.L., Smernik, R.J., 2015. Quantitative analysis of ^{31}P NMR spectra of soil extracts – dealing with overlap of broad and sharp signals. Magnetic Resonance in Chemistry 53, 679-685.

Doolette, A.L., Smernik, R.J., 2016. Phosphorus speciation of dormant grapevine (Vitis vinifera L.) canes in the Barossa Valley, South Australia. Australian Journal of Grape and Wine Research 22, 462-468.

Doolette, A.L., Smernik, R.J., 2018. Facile decomposition of phytate in the solid-state: kinetics and decomposition pathways. Phosphorus, Sulfur, and Silicon and the Related Elements 193, 192-199.

Doolette, A.L., Smernik, R.J., Dougherty, W.J., 2009. Spiking improved solution phosphorus-31 nuclear magnetic resonance identification of soil phosphorus compounds. Soil Science Society of America Journal 73, 919-927.

Doolette, A.L., Smernik, R.J., Dougherty, W.J., 2010. Rapid decomposition of phytate applied to a calcareous soil demonstrated by a solution ^{31}P NMR study. European Journal of Soil Science 61, 563-575.

Doolette, A.L., Smernik, R.J., Dougherty, W.J., 2011a. Overestimation of the importance of phytate in NaOH–EDTA soil extracts as assessed by ^{31}P NMR analyses. Organic Geochemistry 42, 955-964.

Doolette, A.L., Smernik, R.J., Dougherty, W.J., 2011b. A quantitative assessment of phosphorus forms in some Australian soils. Soil Research 49, 152-165.

Dougherty, W.J., Smernik, R.J., Bünemann, E.K., Chittleborough, D.J., 2007. On the use of hydrofluoric acid pretreatment of soils for phosphorus-31 nuclear magnetic resonance analyses. Soil Science Society of America Journal 71, 1111-1118.

Dougherty, W.J., Smernik, R.J., Chittleborough, D.J., 2005. Application of spin counting to the solid-state ^{31}P NMR analysis of pasture soils with varying phosphorus content. Soil Science Society of America Journal 69, 2058-2070.

Drosos, M., Nebbioso, A., Mazzei, P., Vinci, G., Spaccini, R., Piccolo, A., 2017. A molecular zoom into soil Humeome by a direct sequential chemical fractionation of soil. Science of The Total Environment 586, 807-816.

Dyer, W.J., Wrenshall, C.L., 1941. Organic phosphorus in soils: III. The decomposition of some organic phosphorus compounds in soil cultures. Journal of Soil Science 51, 323.

Dyer, W.J., Wrenshall, C.L., Smith, G.R., 1940. Isolation of phytin from the soil. Science (Washington) 91, 319-320.

Eivazi, F., Tabatabai, M.A., 1977. Phosphatases in soils. Soil Biology and Biochemistry 9, 167-172.

Erickson, H.P., 2009. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. Biological Procedures Online 11, 32.

Fatiadi, A.J., 1968. Bromine oxidation of inositols for preparation of inosose phenylhydrazones and phenylosazones. Carbohydrate Research 8, 135-147.

Fioroto, A.M., Kelmer, G.A.R., Albuquerque, L.G.R., César Paixão, T.R.L., Oliveira, P.V., 2017. Microwave-assisted digestion with a single reaction chamber for mineral fertilizer analysis by inductively coupled plasma optical emission spectrometry. Spectroscopy Letters 50, 550-556.

Frostegård, Å., Tunlid, A., Bååth, E., 1991. Microbial biomass measured as total lipid phosphate in soils of different organic content. Journal of Microbiological Methods 14, 151-163.

George, T.S., Giles, C.D., Menezes-Blackburn, D., Condron, L.M., Gama-Rodrigues, A.C., Jaisi, D., Lang, F., Neal, A.L., Stutter, M.I., Almeida, D.S., Bol, R., Cabugao, K.G., Celi, L., Cotner, J.B.,

Feng, G., Goll, D.S., Hallama, M., Krueger, J., Plassard, C., Rosling, A., Darch, T., Fraser, T., Giesler, R., Richardson, A.E., Tamburini, F., Shand, C.A., Lumsdon, D.G., Zhang, H., Blackwell, M.S.A., Wearing, C., Mezeli, M.M., Almås, Å.R., Audette, Y., Bertrand, I., Beyhaut, E., Boitt, G., Bradshaw, N., Brearley, C.A., Bruulsema, T.W., Ciais, P., Cozzolino, V., Duran, P.C., Mora, M.L., de Menezes, A.B., Dodd, R.J., Dunfield, K., Engl, C., Frazão, J.J., Garland, G., González Jiménez, J.L., Graca, J., Granger, S.J., Harrison, A.F., Heuck, C., Hou, E.Q., Johnes, P.J., Kaiser, K., Kjær, H.A., Klumpp, E., Lamb, A.L., Macintosh, K.A., Mackay, E.B., McGrath, J., McIntyre, C., McLaren, T., Mészáros, E., Missong, A., Mooshammer, M., Negrón, C.P., Nelson, L.A., Pfahler, V., Poblete-Grant, P., Randall, M., Seguel, A., Seth, K., Smith, A.C., Smits, M.M., Sobarzo, J.A., Spohn, M., Tawaraya, K., Tibbett, M., Voroney, P., Wallander, H., Wang, L., Wasaki, J., Haygarth, P.M., 2018. Organic phosphorus in the terrestrial environment: a perspective on the state of the art and future priorities. Plant and Soil 427, 191-208.

George, T.S., Quiquampoix, H., Simpson, R.J., Richardson, A.E., 2007. Interactions between phytases and soil constituents: implications for the hydrolysis of inositol phosphates, In: Turner, B.L., Richardson, A.E., Mullaney, E.J. (Eds.), Inositol phosphates: linking agriculture and the environment. CABI International, Wallingford, pp. 221-241.

Gerke, J., 1992. Orthophosphate and organic phosphate in the soil solution of four sandy soils in relation to pH-evidence for humic-FE-(AL-) phosphate complexes. Communications in Soil Science and Plant Analysis 23, 601-612.

Gerke, J., 2010. Humic (organic matter)-Al(Fe)-phosphate complexes: an underestimated phosphate form in soils and source of plant-available phosphate. Soil Science 175, 417-425.

Giles, C.D., Richardson, A.E., Druschel, G.K., Hill, J.E., 2012. Organic anion–driven solubilization of precipitated and sorbed phytate improves hydrolysis by phytases and bioavailability to Nicotiana tabacum. Soil Science 177.

Goh, K.M., Williams, M.R., 1982. Distribution of carbon, nitrogen, phosphorus, sulphur, and acidity in two molecular weight fractions of organic matter in soil chronosequences. Journal of Soil Science 33, 73-87.

Golovan, S.P., Meidinger, R.G., Ajakaiye, A., Cottrill, M., Wiederkehr, M.Z., Barney, D.J., Plante, C., Pollard, J.W., Fan, M.Z., Hayes, M.A., Laursen, J., Hjorth, J.P., Hacker, R.R., Phillips, J.P., Forsberg, C.W., 2001. Pigs expressing salivary phytase produce low-phosphorus manure. Nature Biotechnology 19, 741-745.

Goring, C.A.I., Bartholomew, W.V., 1951. Microbial products and soil organic matter: III. Adsorption of carbohydrate phosphates by clays. Soil Science Society of America Journal 15, 189-194.

Gottselig, N., Nischwitz, V., Meyn, T., Amelung, W., Bol, R., Halle, C., Vereecken, H., Siemens, J., Klumpp, E., 2017. Phosphorus binding to nanoparticles and colloids in forest stream waters. Vadose Zone Journal 16, vzj2016.2007.0064.

Halstead, R.L., Anderson, G., 1970. Chromatographic fractionation of organic phosphates from alkali, acid, and aqueous acetylacetone extracts of soils. Canadian Journal of Soil Science 50, 111-119.

Harrison, A.F., 1982. ^{32}P-method to compare rates of mineralization of labile organic phosphorus in woodland soils. Soil Biology and Biochemistry 14, 337-341.

Harrison, H.F., 1987. Soil organic phosphorus: a review of world literature. CAB International, Wallingford.

Hawthorne, J.N., 1982. Chapter 7. Inositol phospholipids, In: Hawthorne, J.N., Ansell, G.B. (Eds.), New Comprehensive Biochemistry. Elsevier, pp. 263-278.

Hayes, M.H.B., Swift, R.S., 2020. Chapter One - Vindication of humic substances as a key component of organic matter in soil and water, In: Sparks, D.L. (Ed.), Advances in Agronomy. Academic Press, pp. 1-37.

He, Z., Ohno, T., Cade-Menun, B.J., Erich, M.S., Honeycutt, C.W., 2006. Spectral and chemical characterization of phosphates associated with humic substances Soil Science Society of America Journal 70, 1741-1751.

He, Z., Olk, D.C., Cade-Menun, B.J., 2011. Forms and lability of phosphorus in humic acid fractions of Hord silt loam soil. Soil Science Society of America Journal 75, 1712-1722.

Hens, M., Merckx, R., 2001. Functional characterization of colloidal phosphorus species in the soil solution of sandy soils. Environmental Science & Technology 35, 493-500.

Hilderbrand, R.L., 1983. The role of phosphonates in living systems. CRC Press, Boca Raton, Fla.

Hill, J.E., Cade-Menun, B.J., 2009. Phosphorus-31 nuclear magnetic resonance spectroscopy transect study of poultry operations on the Delmarva Peninsula. Journal of Environmental Quality 38, 130-138.

Hiradate, S., Yonezawa, T., Takesako, H., 2006. Isolation and purification of hydrophilic fulvic acids by precipitation. Geoderma 132, 196-205.

Hochberg, Y., Tamhane, A.C., 1987. Multiple comparison procedures. Wiley New York.

Hong, J.-K., Yamane, I., 1981. Distribution of inositol phosphate in the molecular size fractions of humic and fulvic acid fractions. Soil Science and Plant Nutrition 27, 295-303.

Hong, J.K., Yamane, I., 1980. Inositol phosphate and inositol in humic acid and fulvic acid fractions extracted by three methods. Soil Science and Plant Nutrition 26, 491-505.

Ihdene, Z., Dokhan, S., Krutikov, V., Hamada, B., 2011. Détection des organophosphorés par réaction enzymatique. Étude cinétique. Comptes Rendus Chimie 14, 1022-1028.

Irvine, R.F., Schell, M.J., 2001. Back in the water: the return of the inositol phosphates. Nature Reviews Molecular Cell Biology 2, 327.

Irving, G.C.J., Cosgrove, D.J., 1972. Inositol phosphate phosphatases of microbiological origin: the inositol pentaphosphate products of *Aspergillus ficuum* phytases. Journal of Bacteriology 112, 434-438.

Irving, G.C.J., Cosgrove, D.J., 1981. The use of hypobromite oxidation to evaluate two current methods for the estimation of inositol polyphosphates in alkaline extracts of soils. Communications in Soil Science and Plant Analysis 12, 495-509.

Irving, G.C.J., Cosgrove, D.J., 1982. The use of gas-liquid chromatography to determine the proportions of inositol isomers present as pentakis- and hexakisphosphates in alkaline extracts of soils. Communications in Soil Science and Plant Analysis 13, 957-967.

Isbrandt, L.R., Oertel, R.P., 1980. Conformational states of *myo*-inositol hexakis (phosphate) in aqueous solution. A carbon-13 NMR, phosphorus-31 NMR, and Raman spectroscopic investigation. Journal of the American Chemical Society 102, 3144-3148.

Jarosch, K.A., Doolette, A.L., Smernik, R.J., Tamburini, F., Frossard, E., Bünemann, E.K., 2015. Characterisation of soil organic phosphorus in NaOH-EDTA extracts: a comparison of ^{31}P NMR spectroscopy and enzyme addition assays. Soil Biology and Biochemistry 91, 298-309.

John, M.K., Sprout, P.N., Kelley, C.C., 1965. The distribution of organic phosphorus in British Columbia soils and its relationship to soil characteristics. Canadian Journal of Soil Science 45, 87-95.

Johnson, L.F., Tate, M.E., 1970. The structure of *myo*-inositol pentaphosphate. Annals of the New York Academy of Sciences 165, 526-532.

Jørgensen, C., Turner, B.L., Reitzel, K., 2015. Identification of inositol hexakisphosphate binding sites in soils by selective extraction and solution ^{31}P NMR spectroscopy. Geoderma 257-258, 22-28.

Kamerlin, S.C.L., Sharma, P.K., Prasad, R.B., Warshel, A., 2013. Why nature really chose phosphate. Quarterly reviews of biophysics 46, 1-132.

Keeler, J., 2010. Understanding NMR spectroscopy, 2nd ed. ed. Wiley-Blackwell, Chichester.

Kirkby, C.A., Kirkegaard, J.A., Richardson, A.E., Wade, L.J., Blanchard, C., Batten, G., 2011. Stable soil organic matter: a comparison of C:N:P:S ratios in Australian and other world soils. Geoderma 163, 197-208.

Kleber, M., 2010. What is recalcitrant soil organic matter. Environmental Chemistry 7, 320-332.

Kleber, M., Lehmann, J., 2019. Humic substances extracted by alkali are invalid proxies for the dynamics and functions of organic matter in terrestrial and aquatic ecosystems. Journal of Environmental Quality 48, 207-216.

Kleber, M., Sollins, P., Sutton, R., 2007. A conceptual model of organo-mineral interactions in soils: self-assembly of organic molecular fragments into zonal structures on mineral surfaces. Biogeochemistry 85, 9-24.

Kögel-Knabner, I., 1997. ^{13}C and ^{15}N NMR spectroscopy as a tool in soil organic matter studies. Geoderma 80, 243-270.

Kögel-Knabner, I., 2002. The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. Soil Biology and Biochemistry 34, 139-162.

Kögel-Knabner, I., 2017. The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter: fourteen years on. Soil Biology and Biochemistry 105, A3-A8.

Kögel-Knabner, I., de Leeuw, J.W., Hatcher, P.G., 1992. Nature and distribution of alkyl carbon in forest soil profiles: implications for the origin and humification of aliphatic biomacromolecules. Science of The Total Environment 117-118, 175-185.

Kögel-Knabner, I., Guggenberger, G., Kleber, M., Kandeler, E., Kalbitz, K., Scheu, S., Eusterhues, K., Leinweber, P., 2008. Organo-mineral associations in temperate soils: integrating biology, mineralogy, and organic matter chemistry. Journal of Plant Nutrition and Soil Science 171, 61-82.

Kögel-Knabner, I., Rumpel, C., 2018. Chapter One - Advances in molecular approaches for understanding soil organic matter composition, origin, and turnover: a historical overview, In: Sparks, D.L. (Ed.), Advances in Agronomy. Academic Press, pp. 1-48.

Kononova, M.M., Nowakowski, T.Z., Počvennyj Institut Imeni, V.V.D., 1961. Soil organic matter: its nature, its role in soil formation and in soil fertility. New York [etc.] : Pergamon Press.

Kouno, K., Tuchiya, Y., Ando, T., 1995. Measurement of soil microbial biomass phosphorus by an anion exchange membrane method. Soil Biology and Biochemistry 27, 1353-1357.

Kuo, S., 1996. Phosphorus, In: Sparks, D.L., Page, A.L., Helmke, P.A., Loeppert, R.H. (Eds.), Methods of soil analysis part 3—Chemical methods. Soil Science Society of America, American Society of Agronomy, Madison, WI, pp. 869-919.

L'Annunziata, M.F., 1975. The origin and transformations of the soil inositol phosphate isomers. Soil Science Society of America Journal 39, 377-379.

L'Annunziata, M.F., 2007. Origins and biochemical transformations of inositol stereoisomers and their phosphorylated derivatives in soil, In: Turner, B.L., Richardson, A.E., Mullaney, E.J. (Eds.), Inositol phosphates: linking agriculture and the environment. CABI, Wallingford, pp. 41-60.

L'Annunziata, M.F., Gonzalez, J.I., 1977. Soil metabolic transformations of carbon-14-*myo*-inositol, carbon-14-phytic acid and carbon-14-iron(III) phytate. IAEA, International Atomic Energy Agency (IAEA).

L'Annunziata, M.F., I, J.G., Olivares O, L.A., 1977. Microbial epimerization of *myo*-inositol to *chiro*-inositol in soil. Soil Science Society of America Journal 41, 733-736.

Lal, R., 2004. Soil carbon sequestration impacts on global climate change and food security. Science 304, 1623.

Lang, F., Krüger, J., Amelung, W., Willbold, S., Frossard, E., Bünemann, E.K., Bauhus, J., Nitschke, R., Kandeler, E., Marhan, S., Schulz, S., Bergkemper, F., Schloter, M., Luster, J., Guggisberg, F., Kaiser, K., Mikutta, R., Guggenberger, G., Polle, A., Pena, R., Prietzel, J., Rodionov, A., Talkner, U., Meesenburg, H., von Wilpert, K., Hölscher, A., Dietrich, H.P., Chmara, I., 2017. Soil phosphorus supply controls P nutrition strategies of beech forest ecosystems in Central Europe. Biogeochemistry 136, 5-29.

Lead, J.R., Wilkinson, K.J., 2006. Aquatic colloids and nanoparticles: current knowledge and future trends. Environmental Chemistry 3, 159-171.

Lehmann, J., Kleber, M., 2015. The contentious nature of soil organic matter. Nature 528, 60.

Lévesque, M., 1969. Characterization of model and soil organic matter metal-phosphate complexes. Canadian Journal of Soil Science 49, 365-373.

Lévesque, M., Schnitzer, M., 1967. Organo-metallic interactions in soils: 6. Preparation and properties of fulvic acid-metal phosphates. Soil Science 103.

Leytem, A.B., Maguire, R.O., 2007. Environmental implications of inositol phosphates in animal manures, In: Turner, B.L., Richardson, A.E., Mullaney, E.J. (Eds.), Inositol phosphates: linking agriculture and the environment. CABI, Wallingford, pp. 150-168.

Leytem, A.B., Turner, B.L., Thacker, P.A., 2004. Phosphorus composition of manure from swine fed low-phytate grains. Journal of Environmental Quality 33, 2380-2383.

Li, M., Cozzolino, V., Mazzei, P., Drosos, M., Monda, H., Hu, Z., Piccolo, A., 2018. Effects of microbial bioeffectors and P amendements on P forms in a maize cropped soil as evaluated by ^{31}P–NMR spectroscopy. Plant and Soil 427, 87-104.

Lim, P.E., Tate, M.E., 1973. The phytases. II. Properties of phytase fractions F1 and F2 from wheat bran and the *myo*-inositol phosphates produced by fraction F2. Biochimica et Biophysica Acta (BBA) - Enzymology 302, 316-328.

Lis, L.J., McAlister, M., Fuller, N., Rand, R.P., Parsegian, V.A., 1982. Interactions between neutral phospholipid bilayer membranes. Biophysical journal 37, 657-665.

Madigan, M., Martinko, J., Stahl, D., Clark, D., 2012. Brock biology of microorganisms, 13 (Global Edition) ed. Pearson Education Inc. publishing as Benjamin Cummings, Boston.

Makarov, M.I., Haumaier, L., Zech, W., 2002. Nature of soil organic phosphorus: an assessment of peak assignments in the diester region of ^{31}P NMR spectra. Soil Biology and Biochemistry 34, 1467-1477.

Makarov, M.I., Malysheva, T.I., Haumaier, L., Alt, H.G., Zech, W., 1997. The forms of phosphorus in humic and fulvic acids of a toposequence of alpine soils in the northern Caucasus. Geoderma 80, 61-73.

Martin, C.J., Evans, W.J., 1987. Phytic acid: divalent cation interactions. V. titrimetric, calorimetric, and binding studies with cobalt(ii) and nickel(ii) and their comparison with other metal ions. Journal of Inorganic Biochemistry 30, 101-119.

McCarthy, S.M., Melman, J.H., Reffell, O.K., Gordon-Wylie, S.W., 2020. Chapter 24 - Synthesis and partial characterization of biodiesel via base-catalyzed transesterification, In: Dahiya, A. (Ed.), Bioenergy (Second Edition). Academic Press, pp. 519-524.

McDowell, R.W., Cade-Menun, B., Stewart, I., 2007. Organic phosphorus speciation and pedogenesis: analysis by solution ^{31}P nuclear magnetic resonance spectroscopy. European Journal of Soil Science 58, 1348-1357.

McDowell, R.W., Stewart, I., 2006. The phosphorus composition of contrasting soils in pastoral, native and forest management in Otago, New Zealand: sequential extraction and ^{31}P NMR. Geoderma 130, 176-189.

McDowell, R.W., Stewart, I., Cade-Menun, B.J., 2006. An examination of spin–lattice relaxation times for analysis of soil and manure extracts by liquid state phosphorus-31 nuclear magnetic resonance spectroscopy. Journal of Environmental Quality 35, 293-302.

McKercher, R.B., Anderson, G., 1968a. Characterization of the inositol penta- and hexaphosphate fractions of a number of Canadian and Scottish soils. Journal of Soil Science 19, 302-310.

McKercher, R.B., Anderson, G., 1968b. Content of inositol penta- and hexaphosphates in some Canadian soils. Journal of Soil Science 19, 47-55.

McKercher, R.B., Anderson, G., 1989. Organic phosphate sorption by neutral and basic soils. Communications in Soil Science and Plant Analysis 20, 723-732.

McLaren, T.I., Simpson, R.J., McLaughlin, M.J., Smernik, R.J., McBeath, T.M., Guppy, C.N., Richardson, A.E., 2015a. An assessment of various measures of soil phosphorus and the net accumulation of phosphorus in fertilized soils under pasture. Journal of Plant Nutrition and Soil Science 178, 543-554.

McLaren, T.I., Smernik, R.J., Guppy, C.N., Bell, M.J., Tighe, M.K., 2014. The organic P composition of Vertisols as determined by ^{31}P NMR spectroscopy. Soil Science Society of America Journal 78, 1893-1902.

McLaren, T.I., Smernik, R.J., McLaughlin, M.J., Doolette, A.L., Richardson, A.E., Frossard, E., 2020. Chapter Two - The chemical nature of soil organic phosphorus: a critical review and global compilation of quantitative data, In: Sparks, D.L. (Ed.), Advances in Agronomy. Academic Press, pp. 51-124.

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., 2015b. Complex forms of soil organic phosphorus–A major component of soil phosphorus. Environmental Science & Technology 49, 13238-13245.

McLaren, T.I., Smernik, R.J., Simpson, R.J., McLaughlin, M.J., McBeath, T.M., Guppy, C.N., Richardson, A.E., 2017. The chemical nature of organic phosphorus that accumulates in

fertilized soils of a temperate pasture as determined by solution ^{31}P NMR spectroscopy. Journal of Plant Nutrition and Soil Science 180, 27-38.

McLaren, T.I., Verel, R., Frossard, E., 2019. The structural composition of soil phosphomonoesters as determined by solution ^{31}P NMR spectroscopy and transverse relaxation (T_2) experiments. Geoderma 345, 31-37.

Meiboom, S., Gill, D., 1958. Modified spin-echo method for measuring nuclear relaxation times. Review of Scientific Instruments 29, 688-691.

Meyer, G., Bünemann, E.K., Frossard, E., Maurhofer, M., Mäder, P., Oberson, A., 2017. Gross phosphorus fluxes in a calcareous soil inoculated with Pseudomonas protegens CHA0 revealed by ^{33}P isotopic dilution. Soil Biology and Biochemistry 104, 81-94.

Milliken, G.A., Johnson, D.E., 2009. Analysis of messy data. Volume 1: designed experiments, 2nd ed. ed. Boca Raton, Fla : CRC Press.

Missong, A., Bol, R., Willbold, S., Siemens, J., Klumpp, E., 2016. Phosphorus forms in forest soil colloids as revealed by liquid-state ^{31}P-NMR. Journal of Plant Nutrition and Soil Science 179, 159-167.

Moyer, J.R., Thomas, R.L., 1970. Organic phosphorus and inositol phosphates in molecular size fractions of a soil organic matter extract. Soil Science Society of America Journal 34, 80-83.

Nebbioso, A., Piccolo, A., 2011. Basis of a Humeomics Science: chemical fractionation and molecular characterization of humic biosuprastructures. Biomacromolecules 12, 1187-1199.

Nebbioso, A., Piccolo, A., 2012. Advances in humeomics: Enhanced structural identification of humic molecules after size fractionation of a soil humic acid. Analytica Chimica Acta 720, 77-90.

Nebbioso, A., Vinci, G., Drosos, M., Spaccini, R., Piccolo, A., 2015. Unveiling the molecular composition of the unextractable soil organic fraction (humin) by humeomics. Biology and Fertility of Soils 51, 443-451.

Negassa, W., Kruse, J., Michalik, D., Appathurai, N., Zuin, L., Leinweber, P., 2010. Phosphorus speciation in agro-industrial byproducts: sequential fractionation, solution ^{31}P NMR, and P K- and L2,3-Edge XANES spectroscopy. Environmental Science & Technology 44, 2092-2097.

Neptune, A.M.L., Tabatabai, M.A., Hanway, J.J., 1975. Sulfur fractions and carbon-nitrogen-phosphorus-sulfur relationships in some Brazilian and Iowa soils. Soil Science Society of America Journal 39, 51-55.

Newman, R.H., Tate, K.R., 1980. Soil phosphorus characterisation by ^{31}P nuclear magnetic resonance. Communications in Soil Science and Plant Analysis 11, 835-842.

Noack, S.R., McLaughlin, M.J., Smernik, R.J., McBeath, T.M., Armstrong, R.D., 2012. Crop residue phosphorus: speciation and potential bio-availability. Plant and Soil 359, 375-385.

Ó'Fágáin, C., Cummins, P.M., O'Connor, B.F., 2011. Gel-Filtration Chromatography, In: Walls, D., Loughran, S.T. (Eds.), Protein Chromatography: Methods and Protocols. Humana Press, Totowa, NJ, pp. 25-33.

Oberson, A., Friesen, D.K., Rao, I.M., Bühler, S., Frossard, E., 2001. Phosphorus transformations in an Oxisol under contrasting land-use systems: the role of the soil microbial biomass. Plant and Soil 237, 197-210.

Ognalaga, M., Frossard, E., Thomas, F., 1994. Glucose-1-phosphate and *myo*-inositol hexaphosphate adsorption mechanisms on goethite. Soil Science Society of America Journal 58, 332-337.

Ogner, G., 1983. ^{31}P-NMR spectra of humic acids: a comparison of four different raw humus types in Norway. Geoderma 29, 215-219.

Ohno, T., Zibilske, L.M., 1991. Determination of low concentrations of phosphorus in soil extracts using malachite green. Soil Science Society of America Journal 55, 892-895.

Olk, D.C., Bloom, P.R., Perdue, E.M., McKnight, D.M., Chen, Y., Farenhorst, A., Senesi, N., Chin, Y.-P., Schmitt-Kopplin, P., Hertkorn, N., Harir, M., 2019. Environmental and agricultural relevance of humic fractions extracted by alkali from soils and natural waters. Journal of Environmental Quality 48, 217-232.

Omotoso, T.I., Wild, A., 1970. Occurence of inositol phosphates and other organic phosphate components in an organic complex. Journal of Soil Science 21, 224-232.

Pansu, M., Gautheyrou, J., 2007. Handbook of soil analysis: mineralogical, organic and inorganic methods. Berlin, Heidelberg : Springer.

Pant, H.K., Edwards, A.C., Vaughan, D., 1994. Extraction, molecular fractionation and enzyme degradation of organically associated phosphorus in soil solutions. Biology and Fertility of Soils 17, 196-200.

Parfitt, R.L., 1979. Anion adsorption by soils and soil materials, In: Brady, N.C. (Ed.), Advances in Agronomy. Academic Press, pp. 1-50.

Piccolo, A., 2001. The supramolecular structure of humic substances. Soil Science 166.

Piccolo, A., 2012. Carbon sequestration in agricultural soils: a multidisciplinary approach to innovative methods. Berlin, Heidelberg : Springer.

Piccolo, A., Riccardo, S., Davide, S., Marios, D., Vincenza, C., 2019. The soil humeome: chemical structure, functions and technological perspectives, In: Vaz Jr., S. (Ed.), Sustainable Agrochemistry - A compendium of technologies. Springer International Publishing, pp. 183-222.

Potter, R.S., Snyder, R.S., 1918. The organic phosphorus of soil. Soil Science 6.

Pretsch, E., Bühlmann, P., Badertscher, M., 2010. Spektroskopische Daten zur Strukturaufklärung organischer Verbindungen, 5th ed. 2010. ed. Springer Berlin Heidelberg, Berlin, Heidelberg.

R Core Team, 2020. R: A language and environment for statistical computing, In: Computing, R.F.f.S. (Ed.), Vienna, Austria.

Reusser, J.E., Verel, R., Frossard, E., McLaren, T.I., 2020a. Quantitative measures of *myo*-IP_6 in soil using solution ^{31}P NMR spectroscopy and spectral deconvolution fitting including a broad signal. Environmental Science: Processes & Impacts 22, 1084-1094.

Reusser, J.E., Verel, R., Zindel, D., Frossard, E., McLaren, T.I., 2020b. Identification of lower-order inositol phosphates (IP_5 and IP_4) in soil extracts as determined by hypobromite oxidation and solution ^{31}P NMR spectroscopy. Biogeosciences 17, 5079-5095.

Richardson, A.E., Hadobas, P.A., Hayes, J.E., 2001. Extracellular secretion of *Aspergillus* phytase from *Arabidopsis* roots enables plants to obtain phosphorus from phytate. The Plant Journal 25, 641-649.

Riley, A.M., Trusselle, M., Kuad, P., Borkovec, M., Cho, J., Choi, J.H., Qian, X., Shears, S.B., Spiess, B., Potter, B.V.L., 2006. *scyllo*-Inositol pentakisphosphate as an analogue of *myo*-inositol 1,3,4,5,6-pentakisphosphate: chemical synthesis, physicochemistry and biological applications. ChemBioChem 7, 1114-1122.

Saunders, W.M.H., Williams, E.G., 1955. Observation on the determination of total organic phosphorus in soils. Journal of Soil Science 6, 254-267.

Schmidt-Rohr, K., Spiess, H.W., 1994. Chapter 3 - High-resolution NMR techniques for solids, Multidimensional solid-state NMR and polymers. Academic Press Limited, San Diego, United States, pp. 69-134.

Schmidt, M.W.I., Torn, M.S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I.A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D.A.C., Nannipieri, P., Rasse, D.P., Weiner, S., Trumbore, S.E., 2011. Persistence of soil organic matter as an ecosystem property. Nature 478, 49-56.

Schubert, S., 2018. Pflanzenernährung, 3., vollständig überarbeitete Auflage ed. Verlag Eugen Ulmer, Stuttgart.

Schwertmann, U., 1964. Differenzierung der Eisenoxide des Bodens durch Extraktion mit Ammoniumoxalat-Lösung. Zeitschrift für Pflanzenernährung, Düngung, Bodenkunde 105, 194-202.

Sharma, V.K., 2013. Oxidation of amino acids, peptides, and proteins: kinetics and mechanism. Hoboken : Wiley.

Simpson, M.J., Johnson, P.C.E., 2006. Identification of mobile aliphatic sorptive domains in soil humin by solid-state ^{13}C nuclear magnetic resonance. Environmental Toxicology and Chemistry 25, 52-57.

Simpson, M.J., Simpson, A.J., 2012. The chemical ecology of soil organic matter molecular constituents. Journal of Chemical Ecology 38, 768-784.

Simpson, M.J., Simpson, A.J., 2017. NMR of soil organic matter, In: Lindon, J.C., Tranter, G.E., Koppenaal, D.W. (Eds.), Encyclopedia of spectroscopy and spectrometry (Third Edition). Academic Press, Oxford, pp. 170-174.

Smernik, R.J., Dougherty, W.J., 2007. Identification of phytate in phosphorus-31 nuclear magnetic resonance spectra: the need for spiking. Soil Science Society of America Journal 71, 1045-1050.

Smith, D.H., Clark, F.E., 1951. Anion-exchange chromatography of inositol phosphates from soil. Soil Science 72, 353-360.

Sorensen-Stowell, K., Hengge, A.C., 2005. Examination of P−OR bridging bond orders in phosphate monoesters using ^{18}O isotope shifts in ^{31}P NMR. The Journal of Organic Chemistry 70, 4805-4809.

Spaccini, R., Mbagwu, J.S.C., Conte, P., Piccolo, A., 2006. Changes of humic substances characteristics from forested to cultivated soils in Ethiopia. Geoderma 132, 9-19.

Spaccini, R., Piccolo, A., Haberhauer, G., Gerzabek, M.H., 2000. Transformation of organic matter from maize residues into labile and humic fractions of three European soils as revealed by ^{13}C distribution and CPMAS-NMR spectra. European Journal of Soil Science 51, 583-594.

Spain, A.V., Tibbett, M., Ridd, M., McLaren, T.I., 2018. Phosphorus dynamics in a tropical forest soil restored after strip mining. Plant and Soil 427, 105-123.

Steck, L., 2018. Identification and quantification of phospholipids in soil extracts using solution ^{31}P NMR spectroscopy, Department of Environmental Systems Science. ETH Zurich, Zurich, p. 32.

Stephens, L.R., Irvine, R.F., 1990. Stepwise phosphorylation of *myo*-inositol leading to *myo*-inositol hexakisphosphate in Dictyostelium. Nature 346, 580-583.

Stevenson, F.F., 1986. Cycles of soil: carbon, nitrogen, phosphorus, sulfur, micronutrients. John Wiley and Sons, New York.

Stevenson, F.J., 1982. Organic forms of soil nitrogen. Nitrogen in Agricultural Soils, 67-122.

Stevenson, F.J., 1994. Humus chemistry: genesis, composition, reactions, Second ed. ed. New York [etc.] : Wiley.

Steward, J.H., Tate, M.E., 1971. Gel chromatography of soil organic phosphorus. Journal of Chromatography A 60, 75-82.

Strickland, K.P., 1973. The chemistry of phospholipids, Second, completely revised and enlarged edition. ed. Elsevier Scientific Publ. Company.

Sun, M., Alikhani, J., Massoudieh, A., Greiner, R., Jaisi, D.P., 2017. Phytate degradation by different phosphohydrolase enzymes: contrasting kinetics, decay rates, pathways, and isotope effects. Soil Science Society of America Journal 81, 61-75.

Sun, M., Jaisi, D.P., 2018. Distribution of inositol phosphates in animal feed grains and excreta: distinctions among isomers and phosphate oxygen isotope compositions. Plant and Soil 430, 291-305.

Suzumura, M., Kamatani, A., 1993. Isolation and determination of inositol hexaphosphate in sediments from Tokyo Bay. Geochimica et Cosmochimica Acta 57, 2197-2202.

Swift, R.S., Posner, A.M., 1972. Nitrogen, phosphorus and sulphur contents of humic acids fractionated with respect to molecular weight. Journal of Soil Science 23, 50-57.

Tamburini, F., Pistocchi, C., Helfenstein, J., Frossard, E., 2018. A method to analyse the isotopic composition of oxygen associated with organic phosphorus in soil and plant material. European Journal of Soil Science 69, 816-826.

Tate, K.R., Newman, R.H., 1982. Phosphorus fractions of a climosequence of soils in New Zealand tussock grassland. Soil Biology and Biochemistry 14, 191-196.

Thangarasu, V., Anand, R., 2019. Chapter 17 - Comparative evaluation of corrosion behavior of Aegle Marmelos Correa diesel, biodiesel, and their blends on aluminum and mild steel metals, In: Azad, A.K., Rasul, M. (Eds.), Advanced Biofuels. Woodhead Publishing, pp. 443-471.

Thomas, R.L., Bowman, B.T., 1966. The occurrence of high molecular weight organic phosphorus compounds in soil. Soil Science Society of America Journal 30, 799-801.

Tiessen, H., 2008. Phosphorus in the global environment, In: White, P.J., Hammond, J.P. (Eds.), The Ecophysiology of Plant-Phosphorus Interactions. Springer Netherlands, Dordrecht, pp. 1-7.

Tipping, E., Somerville, C.J., Luster, J., 2016. The C:N:P:S stoichiometry of soil organic matter. Biogeochemistry 130, 117-131.

Turner, B.L., 2006. Organic phosphorus in Madagascan rice soils. Geoderma 136, 279-288.

Turner, B.L., 2007. Inositol phosphates in soil: amounts, forms and significance of the phosphorylated inositol stereoisomers., In: Turner, B.L., Richardson, A.E., Mullaney, E.J. (Eds.), Inositol phosphates: linking agriculture and the environment. CAB International, Wallingford, Oxfordshire, UK, pp. 186-206.

Turner, B.L., 2008. Soil organic phosphorus in tropical forests: an assessment of the NaOH–EDTA extraction procedure for quantitative analysis by solution ^{31}P NMR spectroscopy. European Journal of Soil Science 59, 453-466.

Turner, B.L., 2020. Isolation of inositol hexakisphosphate from soils by alkaline extraction and hypobromite oxidation, In: Miller, G.J. (Ed.), Inositol Phosphates: Methods and Protocols. Springer US, New York, NY, pp. 39-46.

Turner, B.L., Cheesman, A.W., Godage, H.Y., Riley, A.M., Potter, B.V., 2012. Determination of *neo*- and D-*chiro*-inositol hexakisphosphate in soils by solution ^{31}P NMR spectroscopy. Environ Sci Technol 46, 4994-5002.

Turner, B.L., Condron, L.M., Richardson, S.J., Peltzer, D.A., Allison, V.J., 2007a. Soil organic phosphorus transformations during pedogenesis. Ecosystems 10, 1166-1181.

Turner, B.L., Mahieu, N., Condron, L.M., 2003a. Phosphorus-31 nuclear magnetic resonance spectral assignments of phosphorus compounds in soil NaOH–EDTA extracts. Soil Science Society of America Journal 67, 497-510.

Turner, B.L., Mahieu, N., Condron, L.M., 2003b. The phosphorus composition of temperate pasture soils determined by NaOH–EDTA extraction and solution ^{31}P NMR spectroscopy. Organic Geochemistry 34, 1199-1210.

Turner, B.L., Mahieu, N., Condron, L.M., 2003c. Quantification of *myo*-inositol hexakisphosphate in alkaline soil extracts by solution ^{31}P spectroscopy and spectral deconvolution. Soil Science 168, 469-478.

Turner, B.L., Papházy, M.J., Haygarth, P.M., McKelvie, I.D., 2002. Inositol phosphates in the environment. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 357, 449.

Turner, B.L., Richardson, A.E., 2004. Identification of *scyllo*-inositol phosphates in soil by solution phosphorus-31 nuclear magnetic resonance spectroscopy. Soil Science Society of America Journal 68, 802-808.

Turner, B.L., Richardson, A.E., Mullaney, E.J., 2007b. Inositol phosphates: linking agriculture and the environment. CABI, Wallingford.

Turner, B.L., Wells, A., Condron, L.M., 2014. Soil organic phosphorus transformations along a coastal dune chronosequence under New Zealand temperate rain forest. Biogeochemistry 121, 595-611.

Vance, E.D., Brookes, P.C., Jenkinson, D.S., 1987. An extraction method for measuring soil microbial biomass C. Soil Biology and Biochemistry 19, 703-707.

Vaz, M.D.R., Edwards, A.C., Shand, C.A., Cresser, M., 1992. Determination of dissolved organic phosphorus in soil solutions by an improved automated photo-oxidation procedure. Talanta 39, 1479-1487.

Veinot, R.L., Thomas, R.L., 1972. High molecular weight organic phosphorus complexes in soil organic matter: inositol and metal content of various fractions. Soil Science Society of America Journal 36, 71-73.

Vestergren, J., Vincent, A.G., Jansson, M., Persson, P., Ilstedt, U., Gröbner, G., Giesler, R., Schleucher, J., 2012. High-resolution characterization of organic phosphorus in soil extracts using 2D ^{1}H–^{31}P NMR correlation spectroscopy. Environmental Science & Technology 46, 3950-3956.

Vincent, A.G., Schleucher, J., Gröbner, G., Vestergren, J., Persson, P., Jansson, M., Giesler, R., 2012. Changes in organic phosphorus composition in boreal forest humus soils: the role of iron and aluminium. Biogeochemistry 108, 485-499.

Vincent, A.G., Vestergren, J., Gröbner, G., Persson, P., Schleucher, J., Giesler, R., 2013. Soil organic phosphorus transformations in a boreal forest chronosequence. Plant and Soil 367, 149-162.

Vold, R.L., Waugh, J.S., Klein, M.P., Phelps, D.E., 1968. Measurement of spin relaxation in complex systems. The Journal of Chemical Physics 48, 3831-3832.

Volkmann, C.J., Chateauneuf, G.M., Pradhan, J., Bauman, A.T., Brown, R.E., Murthy, P.P.N., 2002. Conformational flexibility of inositol phosphates: influence of structural characteristics. Tetrahedron Letters 43, 4853-4856.

Wang, L., Amelung, W., Willbold, S., 2017. Diffusion-Ordered Nuclear Magnetic Resonance Spectroscopy (DOSY-NMR): A Novel Tool for Identification of Phosphorus Compounds in Soil Extracts. Environmental Science & Technology 51, 13256-13264.

Wang, L., Amelung, W., Willbold, S., 2021. ^{18}O isotope labeling combined with ^{31}P nuclear magnetic resonance spectroscopy for accurate quantification of hydrolyzable phosphorus species in environmental samples. Analytical Chemistry 93, 2018-2025.

Williams, C., Anderson, G., 1968. Inositol phosphates in some Australian soils. Soil Research 6, 121-130.

WRB, I.W.G., 2014. World reference base for soil resources. Food and Agriculture Organization of the United Nations FAO, Rome.

Wrenshall, C.L., Dyer, W.J., 1941. Organic phosphorus in soils: II. the nature of the organic phosphorus compounds. A. Nucleic acid derivatives. B. Phytin. Soil Science 51.

Yoshida, R.K., 1940. Studies on organic phosphorus compounds in soil; isolation of inositol. Soil Science 50.

Zhang, X., Li, Z., Yang, H., Liu, D., Cai, G., Li, G., Mo, J., Wang, D., Zhong, C., Wang, H., Sun, Y., Shi, J., Zheng, E., Meng, F., Zhang, M., He, X., Zhou, R., Zhang, J., Huang, M., Zhang, R., Li, N., Fan, M., Yang, J., Wu, Z., 2018. Novel transgenic pigs with enhanced growth and reduced environmental impact. eLife 7, e34286.

Zimmermann, P., Zardi, G., Lehmann, M., Zeder, C., Amrhein, N., Frossard, E., Bucher, M., 2003. Engineering the root–soil interface via targeted expression of a synthetic phytase gene in trichoblasts. Plant Biotechnology Journal 1, 353-360.

Acknowledgements

First of all, I am very thankful for the great support received from my supervisor Dr Timothy McLaren during my whole PhD. Furthermore, this PhD project would not have been possible without the valuable inputs and feedbacks of Prof Dr Emmanuel Frossard as well as the great help of Dr René Verel from the Laboratory of Inorganic Chemistry and Daniel Zindel and his group from the Laboratory of Physical Chemistry from ETH Zurich.

Many thanks to the group of Plant Nutrition, especially to Dr Laurie Mauclaire Schönholzer, Dr Federica Tamburini, Dr Astrid Oberson, and Monika Macsai, who always helped and supported me without hesitation. I also kindly thank Lena Steck and Dayana Naimid Del Valle Esnarriaga for their great assistance and contribution to my PhD project. A special thank you goes to my former PhD office mates Maja Siegenthaler and Grace Crain for our fruitful and pleasant discussions and chats.

I am very grateful for the collaboration with Dr Andrew Neal and Dr Mark Durenkamp at Rothamsted Research in the UK as well as the collaboration with Prof Dr Alessandro Piccolo and the group of CERMANU at the Università di Napoli Federico II in Italy. Claudia Savarese was very supportive and helped me a lot with the Humeomics sequential chemical fractionation method applied in chapter 4. Furthermore, I am thankful to Dr Giovanni Vinci, Silvana Cangemi, Dr Vincenza Cozzolino, Prof Dr Riccardo Spaccini, and the whole group for warmly welcoming me in Naples and for always supporting me during my stay despite the increasingly difficult times due to the Covid-19 pandemic.

The technical support of Björn Studer from the Physics of Environmental Systems group and the XRD analyses provided by Prof Dr Ruben Kretzschmar are greatly acknowledged. In addition, I am thankful for the soil samples provided by Dr Else Bünemann, Prof Dr Stéphane Burgos, Maurus Fischer, Markus Kipfer, Dr David Lester, Dr Gregor Meyer, Dr Astrid Oberson, Dr Chiara Pistocchi, Jürg Reusser, Markus Schönauer, Daniel Zizek, and Anton Zwimpfer.

My PhD thesis would not have been possible without Prof Dr Rainer Schulin, Dr Michael Scheifele, and the late Werner Rohr, who greatly contributed to my joy in doing research and Soil Sciences in general.

Last but not least I thank my friends and family for their encouragement. Most of all, I am deeply and forever grateful for the help, patience and love of my husband Roger Kohler.

Appendix 1

General soil properties

Table A1-1. Additional information on sampling sites of the 29 soil samples. The six soil samples investigated in depth in this thesis are marked in bold: The Ferralsol (17S5), the Vertisol (17S6), the Cambisol (P) (17S12), the Cambisol (F) (17S22), the Gleysol (17S26) and the Cambisol (A) (17S29).

ID	Site	Country	Coordinates	Elevation (ASL)	Soil type	Sampling depth (cm)	Vegetation
17S1	Ivory	Madagaskar	19°33'S, 46°24'E	900 m	Ferralsol	0-10	arable
17S2	Florencia	Colombia	1°37'N, 75°37'W	280 m	Ferralsol	0-10	arable
17S3	Carimagua	Colombia	4°30'N, 71°19'W	150 m	Ferralsol	0-20	rice
17S4	Carimagua	Colombia	4°30'N, 71°19'W	150 m	Ferralsol	0-20	native grassland
17S5	**Carimagua**	**Colombia**	**4°30'N, 71°19'W**	**150 m**	**Ferralsol**	**0-20**	**grass legume pasture**
17S6	**Queensland**	**Australia**	**27°52'S, 151°37'E**	**402 m**	**Vertisol**	**0-15**	**arable**
17S7	Ersigen	Switzerland	47°06'N, 7°35'E	493 m	Cambisol	0-20	pasture
17S8	Ersigen	Switzerland	47°06'N, 7°35'E	496 m	gleyic Cambisol	0-15	pasture
17S9	Witzwil	Switzerland	46°59'N, 7°05'E	431 m	Fluvisol	0-20	pasture (freshly sown)
17S10	Witzwil	Switzerland	46°59'N, 7°05'E	430 m	Gleysol	0-20	pasture (freshly sown)
17S11	Enggistein	Switzerland	46°55'N, 7°36'E	790 m	Cambisol	0-20	pasture
17S12	**Enggistein**	**Switzerland**	**46°56'N, 7°36'E**	**748 m**	**Cambisol**	**0-20**	**pasture**
17S13	Walkringen	Switzerland	46°57'N, 7°37'E	724 m	Regosol	0-20	pasture
17S14	Walkringen	Switzerland	46°57'N, 7°37'E	687 m	Gleysol	0-20	pasture
17S15	Galmiz	Switzerland	46°58'N, 7°08'E	431 m	Gleysol	0-20	pasture
17S16	Waltenschwil	Switzerland	47°20'N, 8°18'E	426 m	Gleysol	0-20	bare soil
17S17	Aristau	Switzerland	47°17'N, 8°23'E	381 m	Gleysol	0-20	pasture
17S18	Vessertal	Germany	50°36'N, 10°46'E	810 m	Cambisol	0-21	beech forest
17S19	Lüss	Germany	52°50'N, 10°16'E	100 m	Cambisol (A)	0-7	beech forest
17S20	Lüss	Germany	52°50'N, 10°16'E	100 m	Cambisol (Of)	above 0	beech forest
17S21	Lüss	Germany	52°50'N, 10°16'E	100 m	Cambisol (Oh)	above 0	beech forest
17S22	**Bad Brückenau**	**Germany**	**50°21'N, 9°56'E**	**800 m**	**Cambisol (Ah1)**	**0-7**	**beech forest**
17S23	Bad Brückenau	Germany	50°21'N, 9°56'E	800 m	Cambisol (Of)	above 0	beech forest
17S24	Bad Brückenau	Germany	50°21'N, 9°56'E	800 m	Cambisol (Oh)	above 0	beech forest
17S25	Saria	Burkina Faso	12°17'N, 2°9'W	300 m	Ferric Acrisol	0-10	arabe
17S26	**Ruswil**	**Switzerland**	**47°05'N, 8°06'E**	**612 m**	**Gleysol**	**0-10**	**grassland**
17S27	Ruswil	Switzerland	47°05'N, 8°06'E	614 m	Regosol	0-20	grassland
17S28	Damma Gletscher	Switzerland	46°38'N, 8°28'E	1980 m	Cambisol	0-10	grassland
17S29	**Rümlang**	**Switzerland**	**47°26'N, 8°31'E**	**443 m**	**calcaric Cambisol**	**0-20**	**arable**

Table A1-2. Additional information on the soil properties and element concentrations of the 29 soil samples. The six soil samples investigated in depth in this thesis are marked in bold. Microbial P was measured using the liquid hexanol fumigation method after Kouno et al. (1995). Microbial C and N were measured according to the fumigation-extraction method described in Vance et al. (1987). Microbial parameters on soil samples that were received dried were not measured. The other parameters were measured as described in chapters 1 and 2. Values in italic were taken from the literature (Oberson et al., 2001; Meyer et al., 2017).

ID	microbial P (µg P/g)	microbial C (µg C/g)	microbial N (µg N/g)	ratio microbial C:N	NaOH-EDTA extraction			ignition-H_2SO_4 extraction		XRF	Combustion				pH
					total P (µg P/g)	MRP (µg P/g)	MUP (µg P/g)	total P (µg P/g)	P_{org} (µg P/g)	total P (µg P/g)	total N (mg N/g)	total C (mg C/g)	total S (mg S/g)	ratio C:N	
17S1	-	-	-	-	73	27	46	115	95	660	0.9	12	0.1	13	4.1
17S2	-	-	-	-	99	36	63	194	172	962	1.6	18	0.2	11	3.7
17S3	*3*	*72*	*18*	4.1	208	163	45	235	126	471	1.5	22	0.1	15	3.8
17S4	*5*	*145*	*26*	5.5	75	43	32	111	93	228	1.6	25	0.1	16	3.6
17S5	***7***	***172***	***35***	**4.9**	**111**	**66**	**46**	**153**	**109**	**320**	**1.7**	**27**	**0.1**	**15**	**3.6**
17S6	**-**	**-**	**-**	**-**	**447**	**357**	**90**	**1158**	**47**	**1726**	**1.9**	**24**	**0.2**	**13**	**6.1**
17S7	3	14	6	2.1	362	207	155	632	316	1476	1.2	11	0.1	9	5.4
17S8	23	80	27	3.0	438	131	307	553	358	1298	1.7	16	0.2	9	4.8
17S9	23	110	20	5.6	397	237	159	759	315	1256	5.1	73	0.7	14	6.8
17S10	43	373	58	6.4	523	508	14	1502	789	1776	16.2	246	3.4	15	6.6
17S11	20	68	11	6.3	771	542	230	1220	559	2753	2.4	23	0.4	10	5.7
17S12	**31**	**134**	**34**	**3.9**	**827**	**367**	**460**	**1112**	**729**	**2553**	**2.3**	**21**	**0.3**	**9**	**5.1**
17S13	12	92	29	3.1	642	218	424	927	492	1886	2.1	18	0.2	9	4.9
17S14	0	335	85	4.0	1870	957	913	2420	1137	3358	15.4	185	1.9	12	5.6
17S15	69	150	36	4.2	354	205	149	833	462	1201	7.8	102	1.7	13	7.1
17S16	26	93	21	4.4	653	406	248	1268	622	1616	8.4	115	2.2	14	7.0
17S17	53	217	44	5.0	769	362	406	1273	660	1521	8.0	91	1.6	11	7.2
17S18	9	101	24	4.2	576	241	335	737	482	2061	2.7	40	0.4	15	3.7
17S19	4	47	11	4.3	46	22	23	65	58	559	1.0	23	0.2	23	3.7
17S20	32	777	212	3.7	268	147	121	465	349	977	14.0	322	1.9	23	3.6
17S21	69	340	93	3.7	295	147	148	417	317	1191	8.6	175	1.3	20	3.7
17S22	**23**	**219**	**47**	**4.6**	**1579**	**537**	**1042**	**2375**	**1377**	**3841**	**6.6**	**90**	**0.9**	**14**	**3.6**
17S23	78	222	14	16.1	1937	728	1209	2391	1276	3387	15.0	240	1.9	16	3.5
17S24	45	194	43	4.5	1904	726	1177	2866	1689	4063	11.3	174	1.7	15	3.5
17S25	-	-	-	-	22	9	13	79	73	592	0.3	5	0.0	15	6.0
17S26	**213**	**419**	**103**	**4.1**	**1578**	**690**	**888**	**2079**	**939**	**2913**	**10.9**	**148**	**1.7**	**14**	**5.0**
17S27	35	136	36	3.8	1240	683	556	1614	819	2761	6.7	99	1.3	15	6.4
17S28	-	-	-	-	1554	184	1371	1928	1783	3416	9.5	124	1.3	13	3.7
17S29	**-**	***209***	**-**	**-**	**510**	**128**	**382**	**927**	**430**	**1724**	**2.5**	**27.6**	**-**	**11**	***7.7***

Appendix 2

X-ray diffraction results

The soil mineralogy of the six soil samples investigated in this PhD project was determined by Prof Dr Ruben Kretzschmar using X-ray diffraction at the Institute of Biogeochemistry and Pollutant Dynamics (ETH Zurich). In brief, soil samples were ground to powder using a McCrone Micronising Mill (McCrone Scientific Ltd, London, UK) and subsequently packed on a sample holder. The samples were measured on a Bruker D8 Advance diffractometer using Cu $K\alpha_{1,2}$ radiations and a high-resolution energy-dispersive 1D detector (Bruker AXS GmbH, Karlsruhe, Germany). The acquired diffractograms were analysed using Rietveld quantitative phase analysis.

The following results were provided by Prof Dr Ruben Kretzschmar.

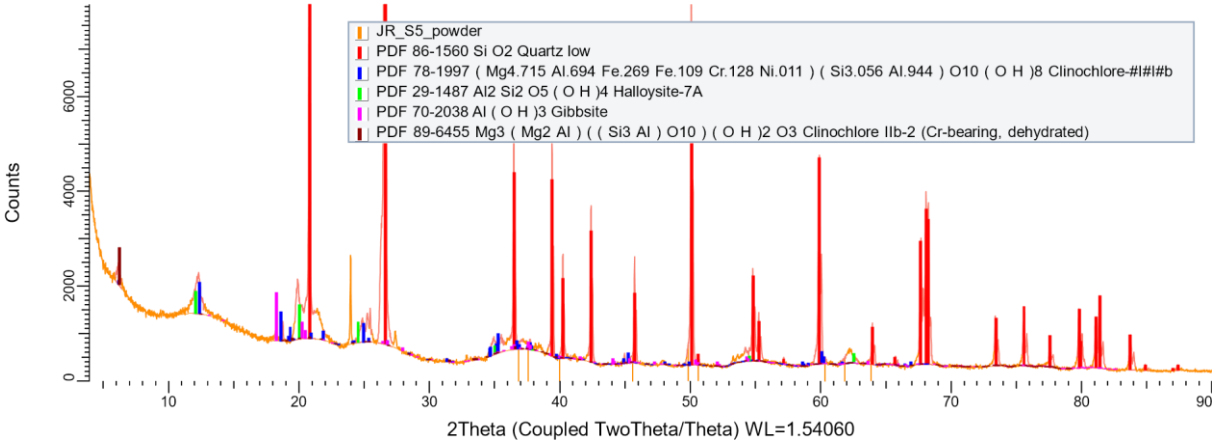
Table A2-1. Soil mineralogy (%) determined by XRD measurements with subsequent Rietveld quantitative phase analysis of the six soils investigated in this thesis.

Mineral	**Ferralsol**	**Vertisol**	**Cambisol (P)**	**Cambisol (F)**	**Gleysol**	**Cambisol (A)**
Quartz (%)	66.6	6.7	68.4	51.3	58.7	63.3
Albites (%)	-	-	12.5	6.7	11.2	9.7
Microclines (%)	-	-	6.3	13.3	6.3	4.2
Clinochlores (%)	0.7	-	5.0	5.7	7.6	7.04
Muscovites (%)	3.7	12.8	7.9	22.9	16.2	15.7
Labradorites (%)	3.0	16.2	-	-	-	-
Montmorillonites (%)	12.4	50.1	-	-	-	-
Kaolinites (%)	13.7	14.2	-	-	-	-

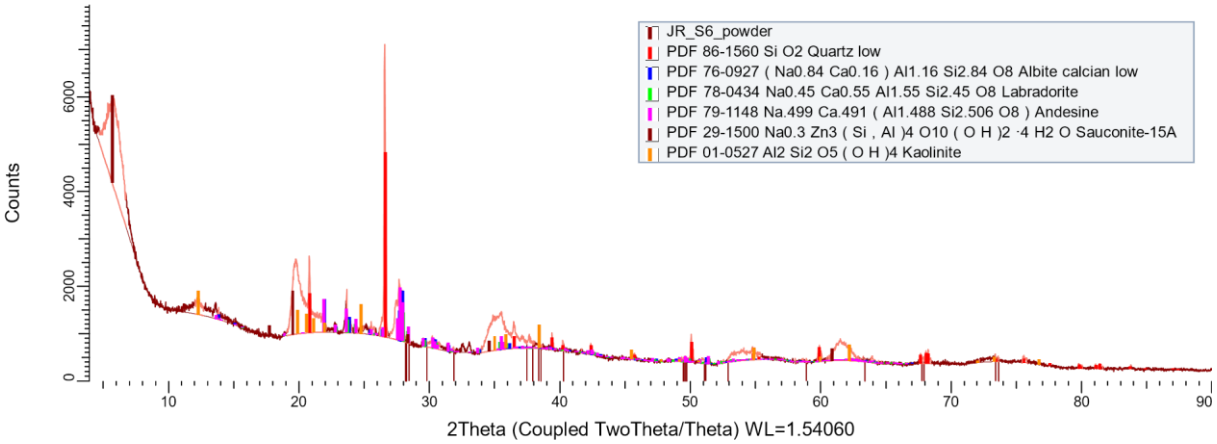
XRD diffractograms: Identification of possible mineral phases

(PDF2 database)

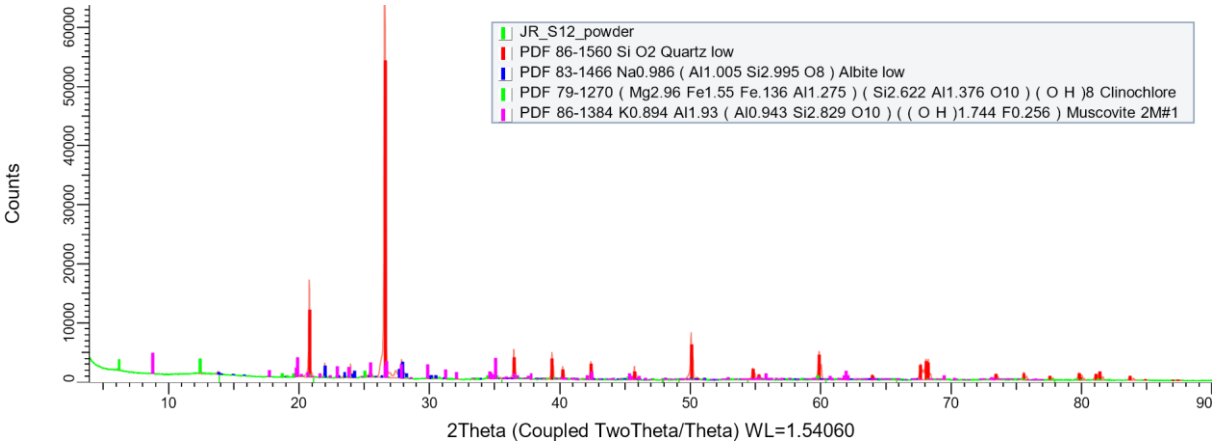
Ferralsol



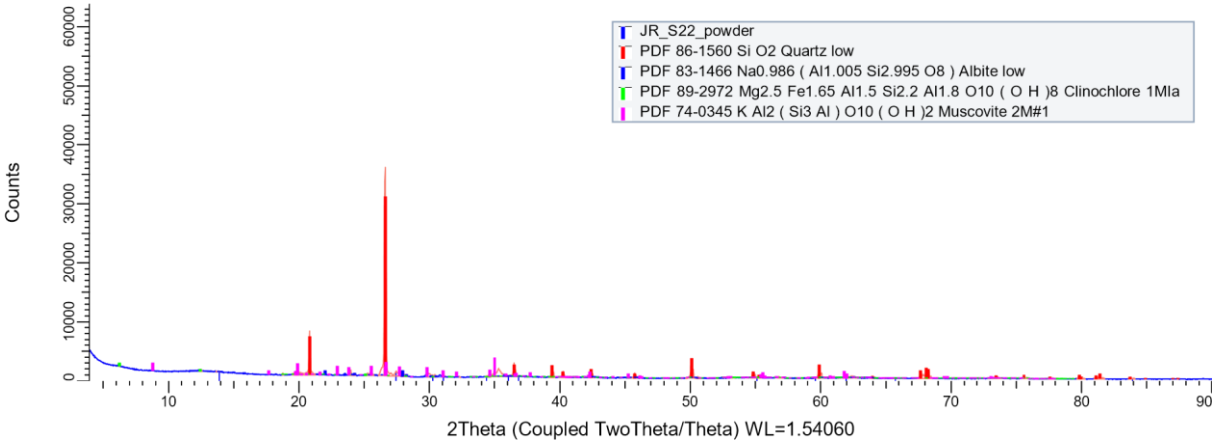
Vertisol



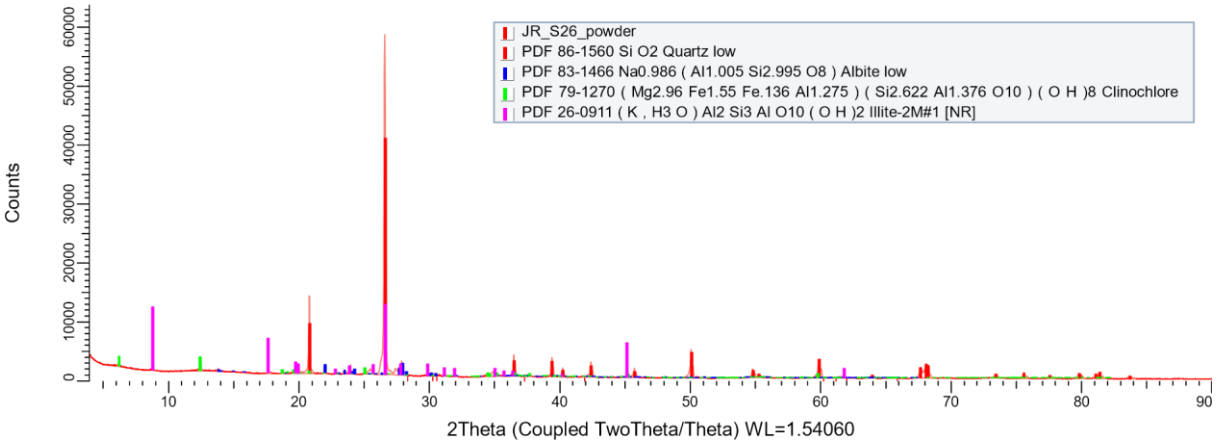
Cambisol (P)



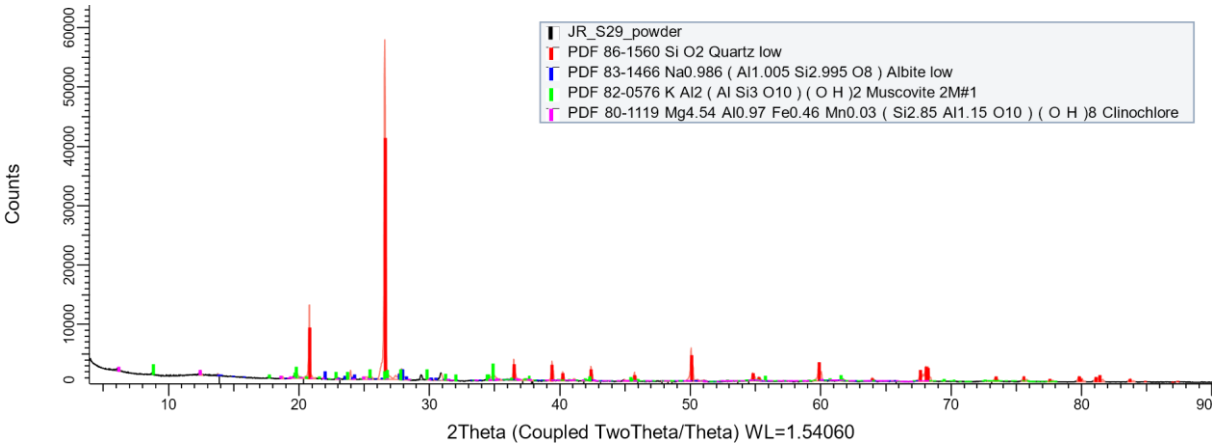
Cambisol (F)



Gleysol

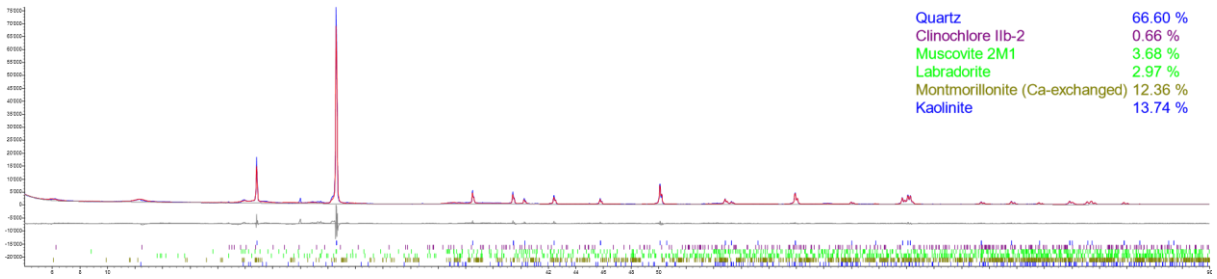


Cambisol (A)

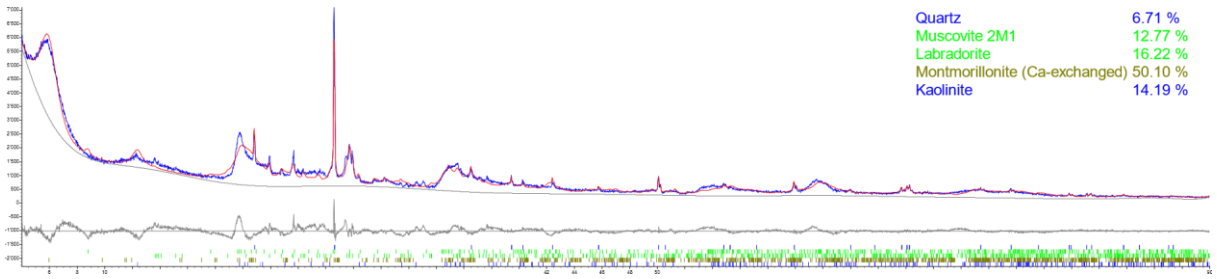


TOPAS: Quantitative Phase Analysis (Rietveld method)

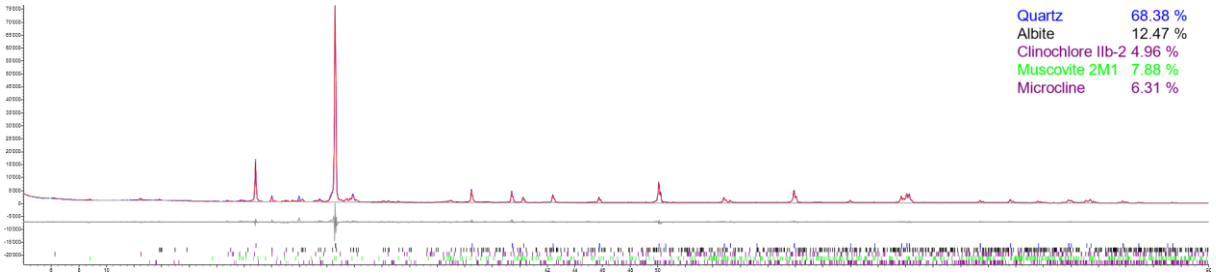
Ferralsol



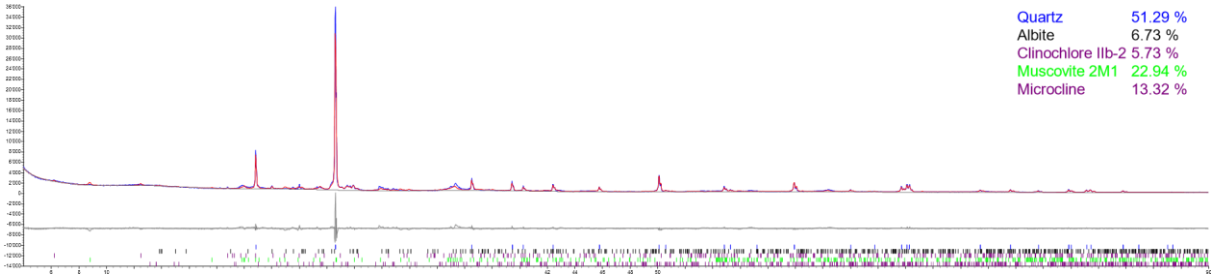
Vertisol



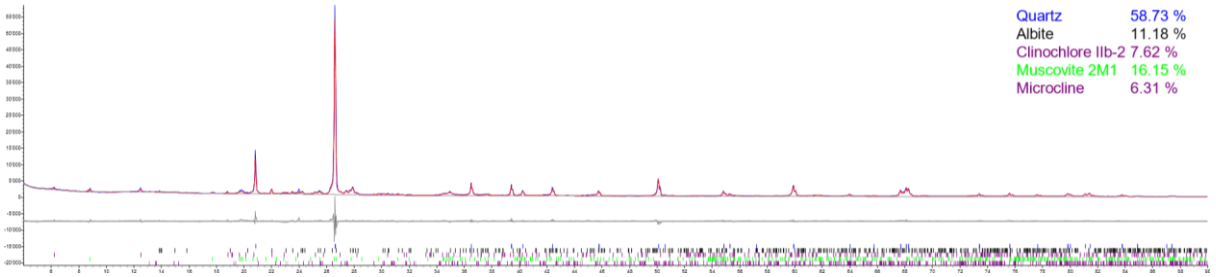
Cambisol (P)



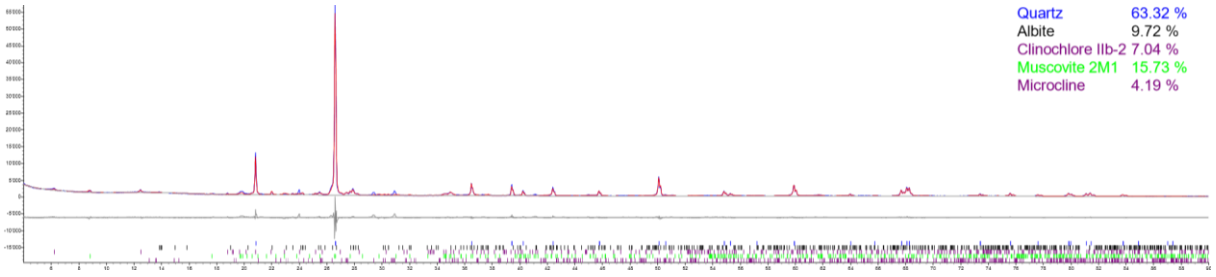
Cambisol (F)



Gleysol



Cambisol (A)



Appendix 3

Pilot study PLFA extraction with subsequent ^{31}P NMR spectroscopic analysis of $CDCl_3$ extracts

Phospholipids were extracted from the Gleysol sample using a mixture of methanol, chloroform and a citrate buffer according to the methods of Bligh and Dyer (1959) and Frostegård et al. (1991). The solution ^{31}P NMR spectrum on the redissolved sample in deuterated chloroform (Figure A3-1) revealed a dominant peak at -17.8 ppm, which represents the triphenylphosphate standard (TPP), and two dominant peaks at chemical shift regions δ -12 to -14 ppm, which could arise from a phosphotriester (Pretsch et al., 2010). Furthermore, signals in the phosphodiester region (δ 2.5 to -3.0 ppm) and a peak at δ 10.8 ppm located in the phosphonate region are visible.

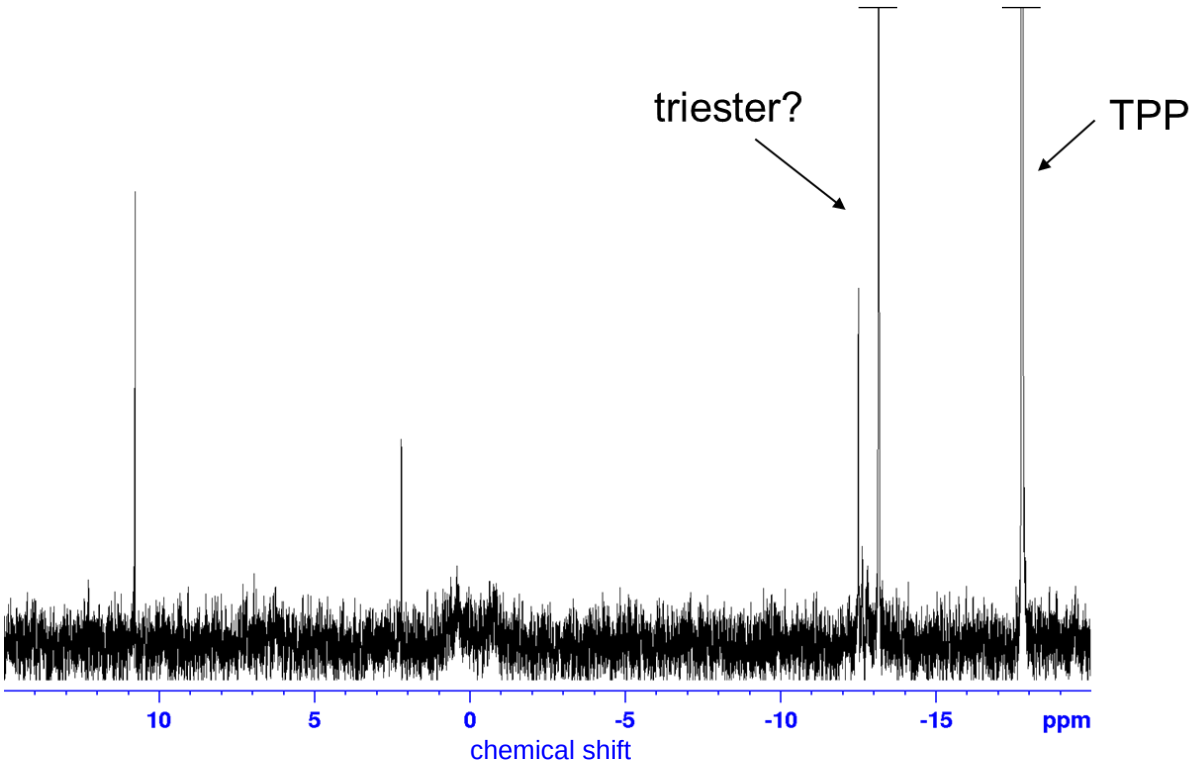


Figure A3-1. Solution ^{31}P NMR spectra (500 MHz) of the in $CDCl_3$ redissolved PLFA extracts. The added triphenylphosphate standard (TPP) is marked.

Appendix 4

Longitudinal relaxation times (T_1) of selected P species

Table A4-1. Longitudinal relaxation times (T_1) of the MDP-standard, the orthophosphate peak and the four *myo*-IP_6 peaks as measured by NMR inversion recovery experiments on 0.25 M NaOH + 0.05 M EDTA soil extracts of chapter 1. The carbon nuclei C1-C6 of the inositol ring on which the phosphate group is attached has been indicated.

Peak	**T_1 (ms)**					
	Ferralsol	**Vertisol**	**Cambisol (P)**	**Cambisol (F)**	**Gleysol**	**Cambisol (A)**
MDP	1489	1867	1226	783	625	1717
Orthophosphate	3374	6052	2174	1399	1053	4953
C1,C3-*myo*-IP_6	-	-	87	890	906	-
C4,C6-*myo*-IP_6	-	-	122	861	954	-

Appendix 5

Pilot study on a phytate standard dissolved in NaOH-EDTA with increasing EDTA concentrations and Fe

According to Figure A5-1, the addition of the paramagnetic ion iron (Fe) to an alkaline solution containing a *myo*-IP_6 standard did not cause a considerable line broadening of the *myo*-IP_6 in the NMR spectra. The Fe concentration in the solution was increased by factor 50, which resulted in an increase of the average line width at half-height over all four peaks of *myo*-IP_6 of a factor 1.5. Hence, the complexation capacity of the added 0.05 M EDTA appears to be high enough to complex even large amounts of Fe. Furthermore, there is no evidence for the development of an additional broad feature in the phosphomonoester region with increasing Fe concentration.

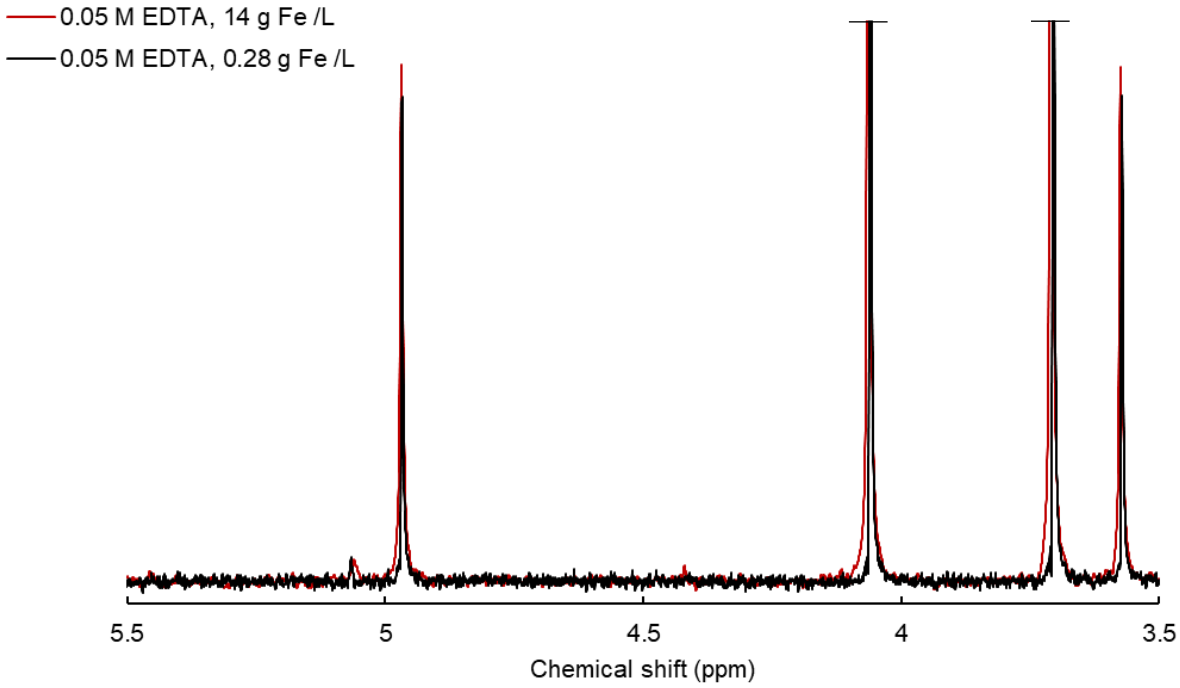


Figure A5-1. Solution ^{31}P NMR spectrum (500 MHz) of a 100 mg P/L *myo*-IP_6 standard (Sigma Aldrich, product no. P5681) dissolved in 0.25 M NaOH + 0.05 M EDTA. The Fe concentration in the NMR solution was adjusted to 14 g Fe/L (red) resp. 0.28 g Fe/L (black) before centrifugation using a 1 kg Fe/L standard solution (in nitric acid). The peak intensities were normalised to the peak intensity of the added MDP standard (times 0.5).

The line width at half-height over all four peaks of *myo*-IP_6 was on average 1.115 Hz for the 14 g Fe/L (red) extract resp. 0.766 Hz for the 0.28 g Fe/L (black) extract.

The EDTA concentrations in an alkaline solution containing a *myo*-IP_6 standard and a known amount of iron was increased in order to assess possible influences on the line width at half-height of the *myo*-IP_6 peaks (Figure A5-1).

The acquired NMR spectra indicate a small shift of the peak position with increasing concentration of EDTA in the solution. This could be explained by changes of the solution's pH due to the increased complexation of ions. Furthermore, the signal-to-noise ratio slightly improves using more than 0.05 M EDTA. However, the effects of different EDTA concentrations on the spectral quality are minor.

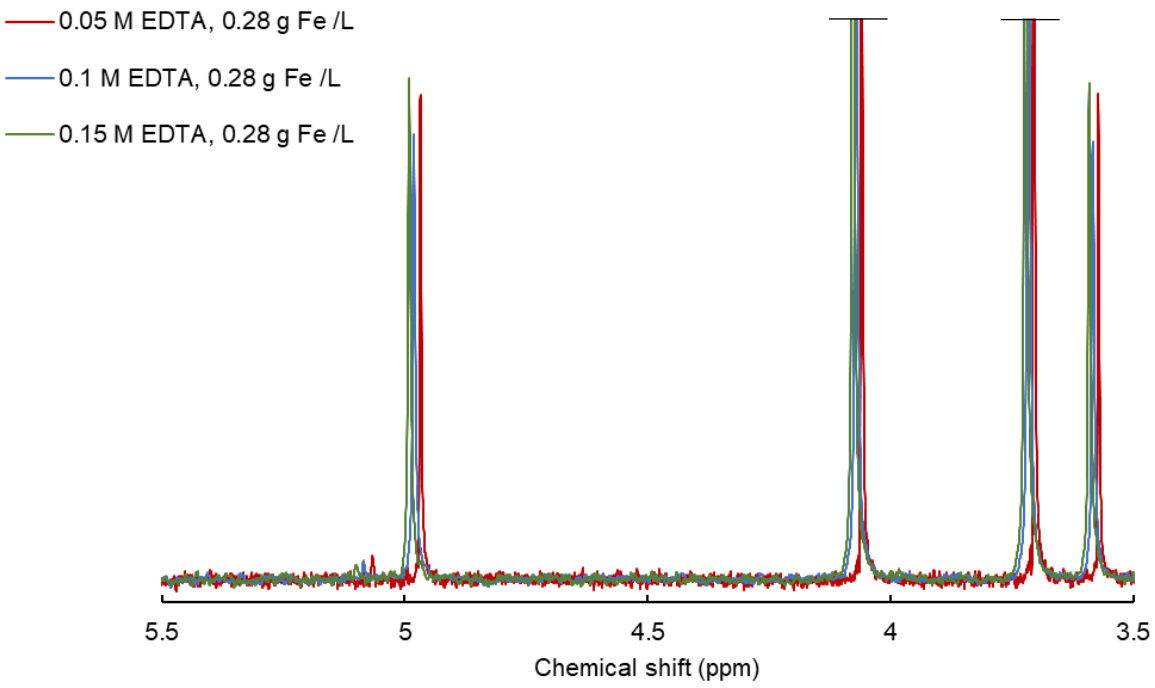


Figure A5-2. Solution ^{31}P NMR spectrum (500 MHz) of a 100 mg P/L *myo*-IP_6 standard (Sigma Aldrich, product no. P5681) dissolved in 0.25 M NaOH + 0.05 M EDTA (red), 0.25 M NaOH + 0.1 M EDTA (blue), 0.25 M NaOH + 0.15 M EDTA (green). The Fe concentration in the NMR solutions was adjusted to 0.28 g Fe/L before centrifugation using a 1 kg Fe/L standard solution (in nitric acid). The peak intensities were normalised to the peak intensity of the added MDP standard (times 0.5).

Appendix 6

Publications

Environmental Science Processes & Impacts

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Environ. Sci.: Processes Impacts*, 2020, **22**, 1084

Quantitative measures of *myo*-IP_6 in soil using solution ^{31}P NMR spectroscopy and spectral deconvolution fitting including a broad signal†

Jolanda E. Reusser, *[a] René Verel, [b] Emmanuel Frossard[a] and Timothy I. McLaren [a]

Received 24th October 2019
Accepted 4th March 2020

DOI: 10.1039/c9em00485h

rsc.li/espi

Inositol phosphates, particularly *myo*-inositol hexakisphosphate (*myo*-IP_6), are an important pool of soil organic phosphorus (P) in terrestrial ecosystems. To measure concentrations of *myo*-IP_6 in alkaline soil extracts, solution ^{31}P nuclear magnetic resonance (NMR) spectroscopy is commonly used. However, overlap of the NMR peaks of *myo*-IP_6 with several other peaks in the phosphomonoester region requires spectral deconvolution fitting (SDF) to partition the signals and quantify *myo*-IP_6. At present, two main SDF approaches are in use; the first fits a Lorentzian/Gaussian lineshape to the *myo*-IP_6 peaks directly to the baseline without an underlying broad signal, and the second fits a Lorentzian/Gaussian lineshape to the *myo*-IP_6 peaks simultaneously with an underlying broad peak. The aim of this study was to compare the recovery of added *myo*-IP_6 to soil extracts using both SDF procedures for six soil samples of diverse origin and differing concentrations of organic P (112 to 1505 mg P per kg_{soil}). The average recovery of total added *myo*-IP_6 was 95% (SD 5) and 122% (SD 32) using SDF with and without an underlying broad signal, respectively. The recovery of individual peaks of *myo*-IP_6 differed, most notably, the C5 phosphate peak of *myo*-IP_6 was overestimated by up to 213% when a broad peak was not included in SDF. Based on the SDF procedure that includes a broad peak, concentrations of *myo*-IP_6 ranged from 0.6 to 90.4 mg P per kg_{soil}, which comprised 1–23% of total phosphomonoesters. Our results demonstrate that the SDF procedure with an underlying broad signal is essential for the accurate quantification of *myo*-IP_6 in soil extracts.

Environmental significance

In terrestrial ecosystems, *myo*-inositol hexakisphosphate (*myo*-IP_6) is considered to be a major pool of organic phosphorus (P) in soil. However, there is disagreement in the literature on how to accurately measure *myo*-IP_6 in soil using solution ^{31}P nuclear magnetic resonance (NMR) spectroscopy. We provide new insights on the use of solution ^{31}P NMR spectroscopy followed by spectral deconvolution fitting for the quantitative measurement of *myo*-IP_6 in soil extracts. Accurate quantification of *myo*-IP_6 is essential for understanding its transformation in soil and hence availability for living organisms as well as its potential contribution to P movement into aquatic ecosystems.

1. Introduction

Phosphorus (P) is an essential macronutrient for all living organisms, which is primarily sourced from the soil environment. It is estimated that between 20 and 80% of the total P (P_{tot}) in soil exists in an organic form.[1,2] A major pool of identifiable organic P (P_{org}) in soil is that of inositol phosphates (IP), of which the *myo* stereoisomer of inositol hexakisphosphate (*myo*-IP_6) is the most abundant.[3,4] Studies have reported that pools of *myo*-IP_6 comprise on average one third of the total P_{org} in soil.[5] The mechanism for its accumulation in soil is thought to be due to its high binding affinity to aluminum and iron (hydro-)oxides.[6]

Solution ^{31}P nuclear magnetic resonance (NMR) spectroscopy has been used since 1980 to identify the chemical nature of P_{org} in soil extracts.[7,8] The majority of P (~80%) in NaOH soil extracts is detected in the phosphomonoester region (organic moiety-O-PO_3) of the NMR spectrum.[5] However, due to many overlapping signals in this region, spectral deconvolution fitting (SDF) procedures are required to partition the NMR signal.[9] The two main SDF approaches applied to soil extracts are that of Turner *et al.*, (2003)[10] or modifications thereof,[11] and Bünemann *et al.*, (2008)[12] or modifications thereof.[13]

[a]*Department of Environmental Systems Science, Group of Plant Nutrition, ETH Zurich, Eschikon 33, CH-8315 Lindau, Switzerland. E-mail: jolanda.reusser@usys.ethz.ch*
[b]*Department of Chemistry and Applied Biosciences, Laboratory of Inorganic Chemistry, ETH Zurich, Vladimir-Prelog-Weg 1-5/10, CH-8093 Zurich, Switzerland*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9em00485h

Turner *et al.*, (2003)[10] were the first to propose a SDF procedure that could partition the NMR signal within the phosphomonoester region and quantify *myo*-IP_6 in soil extracts. The SDF procedure was carried out using the Bruker WinNMR program, and involved fitting a series of sharp signals from the peak maxima of *myo*-IP_6 to the baseline of the spectra. The procedure was applied to 29 soils under grassland in the United Kingdom and the authors tested the efficacy of the procedure by calculating the recovery of added *myo*-IP_6 in a 1 M NaOH solution containing a mixture of P_{org} compounds. The authors reported that the recovery of added *myo*-IP_6 was on average 102%. However, a limitation of this study was that recoveries of added *myo*-IP_6 were determined in non-soil extracts.

In contrast, Bünemann *et al.*, (2008)[12] later proposed a SDF for the quantification of P_{org} compounds (*e.g. myo*-IP_6) that involved fitting a broad feature in the phosphomonoester region, which was then subtracted from the original NMR spectrum and then the overlaying sharp signals were fitted. The authors hypothesized that the broad signal was caused by phosphomonoesters in large and complex molecules, which was later confirmed by McLaren *et al.*, (2015)[14] and McLaren *et al.*, (2019).[13] Bünemann *et al.*, (2008)[12] did not describe the SDF procedure in detail or test its efficacy at the time. However, Doolette *et al.*, (2010)[15] compared the two SDF procedures (with and without a broad signal) to quantitatively recover *myo*-IP_6, which was added to a soil extract. The authors reported that the concentration of *myo*-IP_6 was overestimated by 54% when a broad signal was not included. Doolette *et al.*, (2011)[16] later explained the SDF procedure in more detail. A limitation of this study was that the recovery of added *myo*-IP_6 was only tested on one soil extract. In addition, both SDF methods were not able to identify the recovery of the four individual peaks of *myo*-IP_6 due to poor spectral resolution.

A plethora of studies have been carried out since 2003 to identify the chemical nature of soil P_{org} using solution ^{31}P NMR spectroscopy and SDF procedures.[8] Despite this, there has been no detailed assessment on the efficacy of the two main SDF procedures, which has major consequences for how we understand the composition of soil P_{org}. The implicit assumption of the SDF procedure without a broad signal is that the phosphomonoester region is comprised of an array of sharp peaks (*e.g. myo*-IP_6) with very similar linewidths, which are presumably from small organic molecules.[17,18] In contrast, the implication of the SDF procedure with a broad signal is that it interprets the phosphomonoester region as comprised of a broad signal, which is considered to be P_{org} in the form of large molecular structures and associated with the soil organic matter (SOM),[13,14] and in addition an array of sharp peaks from small organic molecules containing P. It is also important to reconcile these views with previous studies using non-NMR techniques, which often reported large concentrations of P_{org} in large molecular weight fractions.[14,19]

The aim of this study was to assess the efficacy of the two SDF procedures to quantitatively determine *myo*-IP_6 in soil extracts. *myo*-Inositol hexakisphosphate was chosen because it can be easily detected in ^{31}P NMR spectra on soil extracts as four distinct peaks in a 1 : 2 : 2 : 1 ratio. This ratio is caused by the chemical structure of *myo*-IP_6 (see Fig. SI-1 in the ESI†), with the two phosphate groups bound to the C1 and C3 carbon nuclei of the inositol ring being chemically equivalent, and therefore exhibiting the same chemical shift. This is similarly the case for the phosphate groups bound to the C4 and C6 carbon nuclei. The phosphate groups bound to C5 and C2 have distinct chemical environments caused by the conformation of the molecule. The combined four peaks associated with the phosphate groups of *myo*-IP_6 can be used to probe a relatively wide chemical shift range in the phosphomonoester region of the ^{31}P NMR spectrum using a single compound.

2. Materials and methods

2.1. Soil collection and preparation

Six soil samples were collected from the upper horizon of soil profiles including different soil types and land use systems across four countries. We included soils that covered a wide diversity of organic P contents and soil properties: two Cambisols (S3 and S6) and a Gleysol (S5) from Switzerland, a Ferralsol (S1) from Colombia, a Cambisol (S4) from Germany, and a Vertisol (S2) from Australia.[20] Soil S1 was collected in 1997 from the 0–20 cm soil layer of the improved grassland treatment of the long-term Culticore field experiment at the Carimagua Research Station in Colombia.[21] Soil S2 was collected in 2017 from the 0–15 cm soil layer of a field under cropping from southern Queensland, Australia. The site has been under cultivation for the past 25 years and prior to this was shrubland containing sparse *Eucalyptus camaldulensis* L. and native grasses. Soil S3 was collected in 2017 from the 0–20 cm soil layer of a cultivated field, but was under grassland for more than 6 years prior. Soil S4 was collected in 2014 from the 0–7 cm topsoil layer of a beech forest in Bad Brückenau, Germany, as described in Bünemann *et al.*, (2016).[22] Soil S5 was collected in 2017 from the 0–10 cm soil layer of a drained marshland, which has been under grassland for more than 20 years, near Lucerne, Switzerland. Sample S6 was collected in 2013 from the 0–20 cm soil layer from an unfertilized border strip of a cultivated field in Rümlang, Switzerland.[23]

Background information on the studied sites, and some chemical and physical properties of the soils, are reported in Table 1. Soils S3, S4 and S5 were passed through a 5 mm sieve and dried at 60 °C for 5 days. Soils S1 and S6 were received dried and sieved at <2 mm, whereas S2 was received dried (at 40 °C for 2 days) and ground at <2 mm. In order to include soils with a diversity of organic P contents and soil properties, soils were sourced from the field in Switzerland and from previous studies.[21–23] Consequently, there were small differences in soil preparation among soils used in this study, which may slightly affect soil P extraction. Nevertheless, these differences will not affect the application of SDF procedures to the resultant NMR spectra for the quantification of added *myo*-IP_6 in soil extracts. Concentrations of total carbon (C_{tot}) and nitrogen (N_{tot}) in soil were measured using combustion of 50 mg ground soil weighed into tin foil capsules (vario PYRO cube®, Elementar Analysesysteme GmbH). Concentrations of 0.5 M H_2SO_4 extractable P_{tot} and P_{org} in soil were measured using the ignition-H_2SO_4

Table 1 Some background information on the sites and some chemical and physical properties of the soils reported in this study

Parameter	Unit	S1	S2	S3	S4	S5	S6
Soil type	—	Ferralsol	Vertisol	Cambisol	Dystric skeletic Cambisol	Gleysol	Calcaric Cambisol
Country	—	Colombia	Australia	Switzerland	Germany	Switzerland	Switzerland
Coordinates of sampling	—	4°30′ N/71°19′ W	27°52′ S/151°37′ E	46°55′ N/7°36′ E	50°21′ N/9°55′ E	47°05′ N/8°06′ E	47°26′ N/8°31′ E
Elevation	m ASL	150	402	748	800	612	443
Sampling depth	cm	0–20	0–15	0–20	0–7	0–10	0–20
Land use	—	Pasture	Arable	Pasture	Forest	Pasture	Arable
C_{tot}	g C per kg_{soil}	26.7	23.9	21.0	90.3	148.3	27.6
N_{tot}	g N per kg_{soil}	1.7	1.9	2.3	6.6	10.9	2.5
pH in H_2O	—	3.6	6.1	5.1	3.6	5.0	7.7[a]

[a] Meyer *et al.*, (2017).[23]

method of Saunders and Williams (1955)[24] as described in Kuo (1996).[25] Total concentrations of soil P were determined by X-ray fluorescence spectroscopy (SPECTRO XEPOS ED-XRF, AMETEK®) using 4.0 g of ground soil sample mixed with 0.9 g of wax (CEREOX Licowax, FLUXANA®). The XRF instrument was calibrated using commercially available reference soils. Soil pH was measured in H_2O at a soil to solution ratio of 1 : 2.5 (w/w) with a glass electrode.

2.2. Extraction of soil organic P

Concentrations of P_{org} were determined based on the method of Cade-Menun *et al.*, (2002).[26] Briefly, 3.0 g of soil was extracted with 30 mL of 0.25 M NaOH + 0.05 M Na_2EDTA. Soil extracts were shaken for 16 h on a horizontal shaker at 150 rpm at 24 °C, centrifuged for 10 min at 4643*g*, and then the supernatant passed through a Whatman no. 42 filter paper. A 20 mL aliquot of the filtrate was frozen at −80 °C and then lyophilized prior to NMR analysis. This resulted in 420 to 782 mg of lyophilized material across all soils. Concentrations of P_{tot} in the remaining filtrates were measured using inductively coupled plasma-optical emission spectrometry (ICP-OES). Concentrations of molybdate reactive P (MRP) were measured using the malachite green method of Ohno and Zibilske (1991).[27] The difference between P_{tot} and MRP in NaOH-EDTA filtrates is molybdate unreactive P (MUP), which is largely considered to be P_{org},[28] but may also include a small proportion of condensed phosphates (*e.g.* pyrophosphate).[29]

2.3. Preparation of lyophilized material for solution ^{31}P NMR spectroscopy

Preparation of lyophilized material for solution ^{31}P NMR spectroscopy was based on a modification of the methods as reported in Vincent *et al.*, (2013)[30] and Spain *et al.*, (2018).[31] Briefly, 120 mg of lyophilized material was weighed into 1.5 mL microcentrifuge tubes and then a 600 μL aliquot of 0.25 M NaOH + 0.05 M Na_2EDTA solution was added. The solution was briefly vortexed and then allowed to rest overnight in order for complete hydrolysis of RNA and phospholipids.[32–35] This is because the hydrolysis products of RNA (RNA mononucleotides[34]) and phospholipids (α- and β-glycerophosphate[32]) generate peaks in the phosphomonoester region.[32,35] The microcentrifuge tubes were then centrifuged at 10 621*g* for 15 min, and a 500 μL aliquot of the supernatant was transferred to another 1.5 mL microcentrifuge tube, which then received a 25 μL aliquot of a 0.03 M methylenediphosphonic acid (MDP) standard in D_2O (Sigma-Aldrich, product no. M9508) and a 25 μL aliquot of sodium deuteroxide (NaOD) at 40% (w/w) in D_2O (Sigma-Aldrich, product no. 372072). The solution was briefly vortexed and then transferred to 5 mm NMR tubes for analysis.

Subsequent NMR analyses of samples S4 and S5 revealed considerable line-broadening of all peaks in the NMR spectra, which might have been caused by high sample viscosity. Therefore, the preparation of lyophilized material was repeated for these samples but at a wider ratio, as recommended by Cade-Menun and Liu (2014).[8] For these samples, 80 mg of lyophilized material was dissolved in 600 μL of 0.25 M NaOH +

0.05 M Na_2EDTA solution, and then prepared as previously described. This overcame the issue and resulted in high resolution NMR spectra.

2.4. Solution ^{31}P NMR spectroscopy

All NMR analyses were carried out with a Bruker Avance IIIHD 500 MHz NMR spectrometer equipped with a 5 mm liquid-state Prodigy™ CryoProbe (Bruker Corporation; Billerica, MA) at the NMR facility of the Laboratory of Inorganic Chemistry (Hönggerberg, ETH Zürich). Solution ^{31}P NMR spectra were acquired using a ^{31}P frequency of 202.5 MHz, with gated broadband proton decoupling and 90° pulses (duration of 12 μs) for excitation. Careful shimming of the samples resulted in a spectral resolution of <0.1 Hz. The recycle delay of each sample was set based on an inversion recovery experiment.[36] Briefly, the spin–lattice relaxation times (T_1) were calculated from 10 separate experiments with increasing τ values, the time period between the applied pulses in the inversion recovery sequence. Each spectrum was obtained with the collection of 24 scans and a recycle delay of 5 s. Total duration of the inversion recovery experiment for each sample was 56 min. The recycle delay for each sample was calculated by multiplying the longest T_1 value from the inversion recovery experiment by five. This resulted in recycle delays ranging from 6.7 to 31.0 s across all soils. The number of scans was set to 1024 or 4096, depending on the signal to noise ratio of the obtained spectrum.

2.5. Processing of NMR spectra

Spectral processing involved Fourier transformation, phase correction and baseline adjustment using the TopSpin® software of Bruker (Version 3.5 pl 7, Bruker Corporation; Billerica, MA). All NMR spectra were processed with an exponential line-broadening of 0.6 Hz. Since the concentration of added MDP is known, its integral (net peak area) is directly proportional to that of all other NMR signals.[16] Therefore, quantification of P species in NMR spectra was carried out based on spectral integration.[37] In general, integral regions included: (1) phosphonates, in particular the added MDP (δ 16.9 to 16.3 ppm), which includes its two carbon satellite peaks (at δ 16.96 and 16.36 ppm), 2-aminoethylphosphonic acid (δ 19.8 to 19.6 ppm), unknown phosphonate 1 (δ 19.3 to 19.2 ppm), unknown phosphonate 2 (δ 18.3 to 18.1 ppm) and unknown phosphonate 3 (δ 16.5 to 16.4 ppm); (2) the combined orthophosphate and phosphomonoester region (δ 6.0 to 3.0 ppm); (3) phosphodiesters, in particular unknown phosphodiester 1 (δ 2.5 to 2.2 ppm), unknown phosphodiester 2 (δ 0.6 to 0.5 ppm), DNA (δ −0.7 to −1.4 ppm) and unknown phosphodiester 3 (δ −2.3 to −2.4 ppm); and (4) pyrophosphate (δ −4.8 to −5.4 ppm). These integral regions are highlighted in Fig. SI-3 in the ESI.† Due to overlapping signals in the orthophosphate and phosphomonoester region, SDF was needed to partition the NMR signals within this region.

2.6. Deconvolution fitting procedures

Two spectral deconvolution fitting approaches were applied to the orthophosphate and phosphomonoester region. The first approach involved fitting all identifiable sharp peaks (*i.e.* distinguishable from the noise of the spectrum) to the baseline of the spectra, which is based on the method of Turner *et al.*, (2003).[38] The second approach involved fitting all identifiable sharp signals and simultaneously an underlying broad signal in the phosphomonoester region, which has

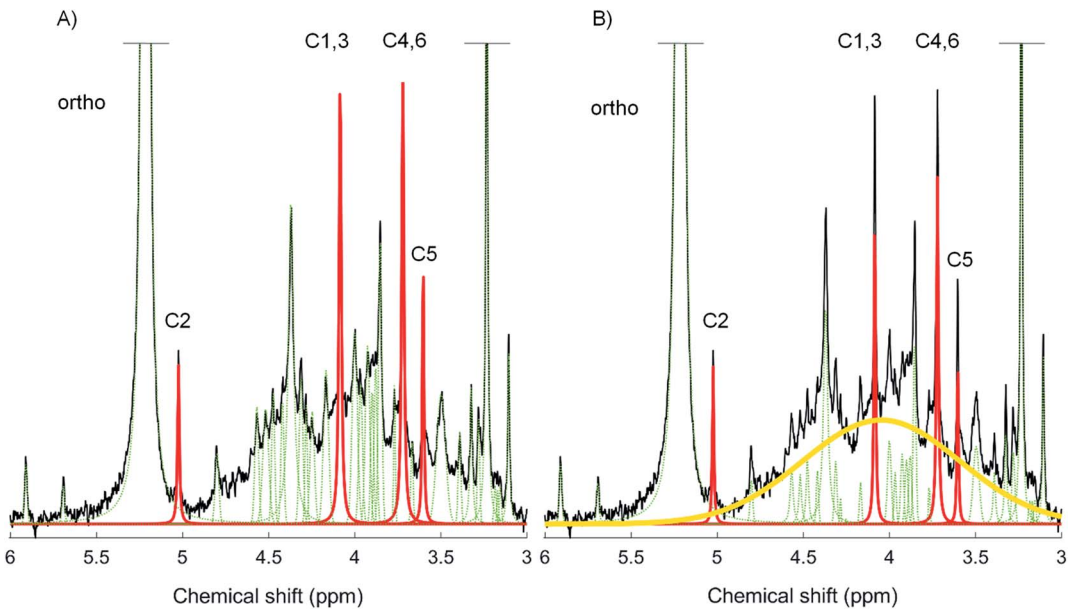


Fig. 1 Solution ^{31}P NMR spectrum of 0.25 M NaOH + 0.05 M EDTA soil extract S6 (black). Graphical representation of the two spectral deconvolution fitting approaches without (A) and with an underlying broad signal (B). All fitted peaks are illustrated; the broad peak (yellow) and the four *myo*-IP_6 peaks (red) have been highlighted. In addition, the phosphate groups bound to carbons C1–6 of *myo*-IP_6 and also that of the orthophosphate (ortho) peak, have been identified.

been described in McLaren *et al.*, (2019),[13] and is based on the method of Bünemann *et al.*, (2008).[17] We used MATLAB® R2017a (The MathWorks, Inc.) scripts containing a non-linear optimization algorithm for the SDF of the NMR spectra. Fig. 1 shows a graphical representation of the two deconvolution fitting procedures used in this study (peak assignment according to[16]). For both SDF procedures, a Lorentzian or Gaussian lineshape model was selected for each peak based on a visual assessment of the most optimized fit and calculated residuals. In general, a Lorentzian lineshape was used to fit the broad signal, signals of *myo*-IP_6 and orthophosphate, whereas a Gaussian lineshape was used to fit the remaining sharp signals, which were generally of lower signal intensity. Each of the identifiable sharp peaks were fitted between an upper and lower bound on both their linewidths at half height and their peak positions at highest intensity. The same was true for the peak position of the underlying broad signal and its linewidth at half height. The upper and lower bound for both parameters of the broad signal being set to δ 3.8 ppm and δ 4.5 ppm, and from 19 Hz to 293 Hz, respectively. Within these visually assessed boundaries, all signals (including the underlying broad signal) were fitted by the non-linear optimization algorithm. The residues of all fitted spectra are plotted in Fig. SI-4 and SI-5 in the ESI.† The goodness of fit parameters of the SDF including the Root-mean-square fitting error and R^2 (coefficient of determination) are listed in Table SI-1 (ESI†). Table SI-1† also includes the results of a reduced Chi^2-test in MATLAB® with the fitted peaks and the according residuals (chemical shift range δ 5.2 to 3.0 ppm). We carried out the Chi^2-test in order to determine if an overfitting of the spectra occurred ($Chi^2 < 1$). Briefly, the results ($Chi^2 > 1$) show that there was no overfitting of the spectra using both SDF approaches, and that the Chi^2 values of the SDF approach with a broad signal were closer to 1 than without a broad signal, which suggests the former approach is a better fit of the spectra. Furthermore, we carried out the Bootstrap sampling function of MATLAB® ($n = 100$) for all peaks within the given boundaries of the peak positions, the area and linewidths at half height. The resulting means of the peak position and linewidths at half height including the standard deviations for the broad signal are listed in Table SI-2 in the ESI.† The calculated standard deviations were very low for all three parameters, indicating a good fit of the broad peak within the given boundaries.

2.7. Spiking experiment

The two deconvolution fitting procedures were assessed by determining the recovery of a known amount of added *myo*-IP_6 to soil extracts. After NMR analysis of the unspiked soil extract, a 10 μL aliquot of a 5.5 mM *myo*-IP_6 standard in D_2O was added to all soil extracts except for soil S5 (Sigma-Aldrich, product no. P5681). For sample S5, a 10 μL aliquot of a 11 mM *myo*-IP_6 standard in D_2O was added. The aim was to add *myo*-IP_6 at a concentration that would result in an increase of peak intensity of approximately 3-times the peak intensity in unspiked extracts.[32] The NMR tube was then sealed with parafilm, inverted several times, and then allowed to rest prior to NMR analysis. The NMR analysis parameters on the spiked soil extract were the same as that carried out on unspiked extracts. Similarly, spectral processing and quantification were carried out as previously described. The recovery of added *myo*-IP_6 was calculated using eqn (1).

$$\text{Recovery of added } myo\text{-IP}_6(\%) = \frac{A(\text{mg P per l}) - B(\text{mg P per l})}{C(\text{mg P per l})} \times 100 \quad (1)$$

where A refers to the concentration of *myo*-IP_6 in the spiked soil extract, B to the concentration of *myo*-IP_6 in the unspiked extract and C to the concentration of the added *myo*-IP_6. Solution ^{31}P NMR recovery of the phytate standard revealed impurities, therefore C represents the actual concentration of *myo*-IP_6 in the standard (see Fig. SI-2 in the ESI†).

2.8. Statistical analyses and graphics

All graphics were created using MATLAB® R2017a (The MathWorks, Inc.). All statistical analyses were carried out using Microsoft® Excel 2016. This included calculating the mean values and standard deviations (SD) of the added *myo*-IP_6

Table 2 Total soil phosphorus (P_{tot}) as measured by X-ray fluorescence (XRF) spectroscopy, and pools of extractable P using the ignition-H_2SO_4 extraction technique of Saunders and Williams (1955)[24] and the NaOH-EDTA extraction technique of Cade-Menun *et al.*, (2002).[26] The percentage of extractable P to that of P_{tot} in soil as measured by XRF is shown in parentheses

	XRF	Ignition-H_2SO_4 extraction	NaOH-EDTA extraction		
Soil	P_{tot}, mg P per kg_{soil}	P_{org}, mg P per kg_{soil} (%)	P_{tot}, mg P per kg_{soil} (%)	MRP[a], mg P per kg_{soil} (%)	MUP[b], mg P per kg_{soil} (%)
S1	320	109 (34)	160 (50)	67 (21)	93 (29)
S2	1726	143 (8)	484 (28)	351 (20)	133 (8)
S3	2553	729 (29)	863 (34)	323 (13)	540 (21)
S4	3841	1377 (36)	1850 (48)	525 (14)	1326 (35)
S5	2913	939 (32)	1490 (51)	610 (21)	880 (30)
S6	1724	430 (25)	510 (30)	128 (7)	382 (22)

[a] Molybdate reactive P (MRP) based on the malachite green method of Ohno and Zibilske (1991).[27] [b] The difference between P_{tot} and MRP is molybdate unreactive P (MUP), which is considered to be P_{org}.

Table 3 Relative concentrations of P classes determined by solution ^{31}P NMR spectroscopy as percentage (%) of total P in NaOH-EDTA soil extracts. Chemical shift regions were attributed to P species according to peak positions in spectra

Soil	Phosphonates[a] (δ 19.8 to 16.5 ppm)	Orthophosphate (δ 5.5 to 5.0 ppm)	P-monoester (δ 6.0 to 3.0 ppm)	P-diesters (δ 2.5 to −2.4 ppm)	Pyrophosphates (δ −4.8 to −5.3 ppm)
S1	1.0	55.0	37.0	5.1	1.9
S2	1.0	85.0	13.3	0.0	0.7
S3	0.2	50.4	47.8	0.4	1.2
S4	1.5	46.7	47.7	2.9	1.3
S5	0.0	48.5	45.3	3.3	2.9
S6	0.2	51.1	46.9	0.0	1.8

[a] The added methylenediphosphonic acid (MDP) standard is not included.

recovery across the six soil samples. The NMR observability was calculated by comparing the concentration of P_{tot} as detected by NMR with that measured by ICP-OES,[16,39] which ranged from 52 to 89% (on average of 68%) across all soils.

3. Results

3.1. Pools of soil P

Concentrations of P_{tot} in soil using XRF ranged from 320 to 3841 mg P per kg_{soil} (Table 2). Concentrations of P_{tot} in NaOH-EDTA extracts comprised 28 to 51% of the P_{tot} in soil as measured by XRF. Concentrations of P_{org} using the ignition-H_2SO_4 extraction technique ranged from 143 to 1377 mg P per kg_{soil}. The concentration of P_{org} in NaOH-EDTA extracts ranged from 93 to 1326 mg P per kg_{soil}, which comprised 8 to 35% (an average value of 24%) of the total soil P using XRF.

3.2. Solution ^{31}P NMR spectra of soil extracts

The majority of NMR signals occurred in the orthophosphate and phosphomonoester region (δ 6.0 to 3.0 ppm), which comprised on average 96% of total NMR signal (Table 3). In general, the peak of greatest intensity across all soils was that of orthophosphate (apart from the added MDP). The largest pool of P_{org} as determined by the integral over the various regions of peaks was that of phosphomonoesters, which accounts on average for 94% of the total NMR signal arising from organic forms. Concentrations of phosphomonoesters range from 36.3 to 501.1 mg P per kg_{soil}. The remaining NMR signal is distributed between phosphodiesters, pyrophosphates and phosphonates.

The phosphomonoester region comprised of two main spectral features based on a visual assessment; the presence of (i) 18 to 47 sharp signals, and (ii) an underlying broad signal. The four peaks of *myo*-IP_6 were observed in the NMR spectra of all soils except in soil S2, in which only the C1,3 and C4,6 peaks could be clearly identified due to the low initial concentration of *myo*-IP_6. In all soils, the C2 peak of *myo*-IP_6 at δ 5.04 ppm in the NMR spectra exhibited little overlap with the broad signal compared to the other peaks of *myo*-IP_6. Some slight overlap between the C2 peak of *myo*-IP_6 and the base of the orthophosphate peak was observed in soils S4 and S5.

3.3. Spike recoveries of total *myo*-IP_6

Spiking the soil extracts with *myo*-IP_6 resulted in a clear increase in its peak intensities at δ 5.04 ppm, 4.10 ppm, 3.72 ppm and 3.61 ppm across all soils (Fig. 2). On average, the increase in peak intensity of *myo*-IP_6 was 3-fold relative to the unspiked samples in terms of absolute intensity from the peak maximum to the baseline of the spectrum. The exception was sample S2, where the increase of the C1,3 and the C4,6 peak was 7-fold relative to the unspiked sample. The increase in peak intensity due to spiking varied for the four individual peaks of *myo*-IP_6. The C2 peak showed the greatest increase of intensity (3.2-fold) whereas the C5 peak showed the least (2.6-fold). The addition of the phytate standard to the soil extract also resulted in an increase in the intensity of other peaks to that of *myo*-IP_6, particularly peaks at δ 3.98, 4.14, 4.17 and 4.57 ppm.

The recovery of total *myo*-IP_6 in the six soil extracts using the SDF procedure with a broad signal was on average 95%, whereas this was on average 122% using the SDF procedure without a broad signal (Table 4). In addition, the variation in recovery of *myo*-IP_6 across the six samples was least using the former approach (SD of 5) compared to the latter approach (SD of 32).

3.4. Spike recoveries of individual peaks of *myo*-IP_6

Spike recoveries for each of the four peaks of *myo*-IP_6 differed between the SDF approaches across all soils (Table 4). Overestimation of spike recoveries using the SDF procedure without a broad peak was greatest (up to 213%) for the C5 peak of *myo*-IP_6 compared to all other peaks. Furthermore, spike recoveries of the C1,3 and C4,6 peaks of *myo*-IP_6 were overestimated on average 8% more than that of the C2 peak. The peak ratios of *myo*-IP_6 in the unspiked soils were on average 1.0 : 2.2 : 1.8 : 0.9 with a broad signal and 1.0 : 2.1 : 1.9 : 0.7 without a broad signal, when the C2 peak was set to 1 (soil S2 was not included due to unreliable measures of the C2 and C5 peaks of *myo*-IP_6).

3.5. Quantification of *myo*-IP_6 (and the broad signal) in soil extracts

Concentrations of total *myo*-IP_6 in soil extracts obtained with SDF with a broad peak ranged from 0.6 to 90.4 mg P per kg_{soil}, which comprised between 1% and 23% of total phosphomonoesters (Table 5). On average, the broad signal accounted for 64% of total phosphomonoesters across all soils (Table 5).

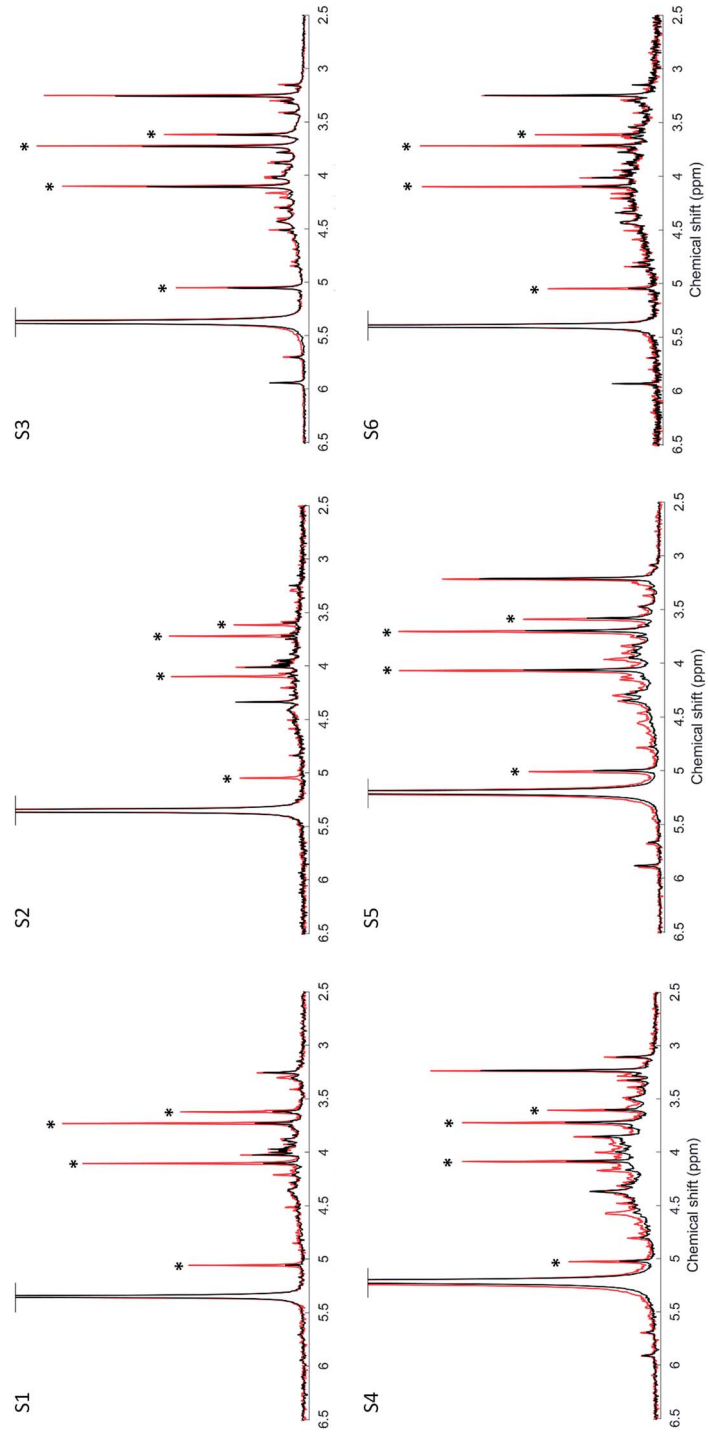


Fig. 2 Solution ^{31}P nuclear magnetic resonance (NMR) spectra of the orthophosphate and phosphomonoester region on unspiked (black) and spiked (red) 0.25 M NaOH + 0.05 M EDTA soil extracts. Samples were spiked with a 10 μL aliquot of 5.5 mm *myo*-IP_6 (Sigma-Aldrich, product no. P5681) except for soil sample S5, which was spiked with a 11 mm *myo*-IP_6-standard solution. Signal intensities were normalized to the MDP peak intensity for direct comparison of unspiked and spiked extracts. However, in order to highlight spectral features, the vertical axes of the spectra have been increased by a factor of 1.7 for soils S3, S4 and S5, and by a factor of 10 for soils S1, S2 and S6. The four peaks generated by *myo*-IP_6 are marked with an asterisk.

Table 4 Calculated recoveries of the added *myo*-IP_6 (Sigma-Aldrich, product no. P5681) in 0.25 M NaOH + 0.05 M EDTA soil extracts across 6 soil samples, and their standard deviation (SD). Concentrations of the added *myo*-IP_6 were obtained from solution ^{31}P NMR spectra using two spectral deconvolution fitting (SDF) procedures; one (i) with a broad underlying signal based on the method of Bünemann *et al.*, (2008),[12] and (ii) one without an underlying broad signal based on the method of Turner *et al.*, (2003).[10] The carbon nuclei C1–C6 of the inositol ring on which the phosphate group is attached has been indicated

SDF procedure	*myo*-IP_6	S1 (%)	S2 (%)	S3 (%)	S4 (%)	S5 (%)	S6 (%)	Average (%)	SD
With broad signal	Total *myo*-IP_6	96	90	99	92	91	105	95	5
	C2-*myo*-IP_6	103	101	73	106	95	104	97	12
	C1,C3-*myo*-IP_6	97	91	108	74	81	106	93	14
	C4,C6-*myo*-IP_6	103	91	103	96	97	116	101	9
	C5-*myo*-IP_6	73	77	96	106	95	81	88	13
Without broad signal	Total *myo*-IP_6	146	120	73	151	94	147	122	32
	C2-*myo*-IP_6	146	113	76	111	101	119	111	23
	C1,C3-*myo*-IP_6	148	109	77	130	83	175	120	38
	C4,C6-*myo*-IP_6	150	127	67	158	91	124	120	35
	C5-*myo*-IP_6	135	133	76	213	117	167	140	46

When concentrations of *myo*-IP_6 as determined by SDF with a broad peak were subtracted from the *myo*-IP_6 values obtained by SDF without a broad peak, the amount of P_{org} that would have been previously attributed to *myo*-IP_6 ranged from 1.3 to 56.1 mg P per kg_{soil}.

4. Discussion

4.1. Extractability of soil organic P and solution ^{31}P NMR spectra

Concentrations of P_{org} in NaOH-EDTA extracts were similar to reported values in previous studies.[19,37] The difference between total soil P as measured by XRF and NaOH-EDTA extractable P is likely due to P_{inorg} held within mineral silicates and other insoluble mineral phases containing P.[40] Organic P extracted with the NaOH-EDTA technique is strongly correlated to pools of soil P_{org}.[40] The extraction method allows for detailed characterization of P_{org} forms using ^{31}P NMR spectroscopy.[8] Solution ^{31}P NMR spectra were highly resolved and exhibited a high signal-to-noise ratio across all samples. The broad classes of P detected by NMR across all soils were phosphonates, orthophosphate, phosphomonoesters, phosphodiesters and pyrophosphate, except for the absence of phosphonates in S5 and phosphodiesters in S2. Their distribution within the total NMR signal is consistent with previous studies, which typically show the majority of NMR signal occurs in the orthophosphate and phosphomonoester region.[14,19,41] The peaks of *myo*-IP_6 were also clearly observed within this region, and therefore allowed for their quantification using SDF.[9]

4.2. Recovery of *myo*-IP_6 in soil extracts

Concentrations of added *myo*-IP_6 were overestimated when the SDF procedure did not include a broad signal. The reason for this is likely twofold: (1) the intensities of the *myo*-IP_6 peaks are higher when a broad signal is not included; and (2) the linewidths of these peaks when fitted to the baseline are greater if a broad signal is not included in the SDF compared to that when a broad signal is included. Therefore, the peak area belonging to compounds other than *myo*-IP_6 is being attributed to *myo*-IP_6, which may result in an overestimation of its concentration in soil. This supports the finding of Doolette *et al.*, (2010)[15] who reported an overestimation of a phytate spike by 54% when a broad signal was not fitted. Whilst the finding of Doolette *et al.*, (2010)[15] was higher than that found in the current study, it is likely due to a greater overlap among sharp peaks in the phosphomonoester region. The authors observed and fitted up to six sharp peaks in the orthophosphate and phosphomonoester region, but the C2 peak of *myo*-IP_6 was not visible because it overlapped with orthophosphate. This was not the case in the current study where all four peaks of *myo*-IP_6 were visible and the C2 peak of *myo*-IP_6 was clearly separated from orthophosphate.

Whilst the recovery of added *myo*-IP_6 was generally overestimated using SDF without a broad signal across all soils, this was not the case for soil S3. The reason for this is unclear but might be due to the high proportion of *myo*-IP_6 to total phosphomonoesters in this soil. It might also be due to an underestimation of the side regions at the base of the peak, as the fitting

Table 5 Concentrations of organic P compounds obtained from solution ^{31}P NMR spectra of 0.25 M NaOH + 0.05 M EDTA soil extracts. The spectral deconvolution fitting of the phosphomonoester region has been carried out with the inclusion of an underlying broad signal

Parameter	Unit	S1	S2	S3	S4	S5	S6
myo-IP_6	mg P per kg_{soil}	4.4	0.6	80.7	46.2	90.4	6.3
Proportion of *myo*-IP_6 to total phosphomonoesters	%	12.0	1.0	23.0	9.0	23.0	4.0
Broad peak	mg P per kg_{soil}	21.6	30.9	186.0	305.8	216.7	110.3
Proportion of broad peak to total phosphomonoesters	%	59.0	79.0	53.0	61.0	54.0	78.0

of neighbouring sharp peaks can influence the partitioning of signals within a particular region. In comparison to SDF with a broad signal of this soil, the fitting of side peak areas of *myo*-IP_6 appeared to improve based on a visual assessment, *i.e.* the peak height was reduced but the linewidth at half height increased. This highlights the sensitivity of peak fitting to small changes in peak shape for the quantification of P species in some soils.

The variability between the two SDF approaches differed in their recovery of added *myo*-IP_6 across all soils, which was least for the SDF approach with a broad signal than that without. The reason for this appears to be the relative proportion of the underlying broad signal to the *myo*-IP_6 peaks among different soils. This is consistent with other compounds that are present within this region, which overlay the broad signal. Doolette *et al.*, (2010)[15] reported that concentrations of glycerophosphate were doubled when SDF was carried out without an underlying broad signal. Of course, since *myo*-IP_6 exhibits 4 peaks in a NMR spectrum, this would vary between the individual peaks of *myo*-IP_6, based on differing proportions of an underlying broad signal.

Individual recoveries of the four peaks of *myo*-IP_6 differed. The greatest difference in the recovery of peaks from *myo*-IP_6 occurred when SDF was carried out without a broad signal, which supports that the baseline of the spectra varies within these chemical shifts.[9] The C1,3 and C4,6 peaks of *myo*-IP_6 have chemical shifts within the phosphomonoester region where the intensity of the broad peak is high relative to that for the C2 and C5 peaks. However, the recovery of the C5 peak of *myo*-IP_6 at δ 3.61 ppm was generally overestimated more than that of the other *myo*-IP_6 peaks. The C5 peak of *myo*-IP_6 is present on the shoulder of the broad signal, which has a maximum intensity at about δ 4.06 ppm. Doolette and Smernik (2015)[9] hypothesized that fitting peaks from the peak maxima to the baseline would result in substantially more signal from an underlying broad peak to the C5 peak of *myo*-IP_6 compared to the C2 peak of *myo*-IP_6. Moreover, the C5 peak appears to be most sensitive to changes in the fitting procedures, as it shows the highest variation for both deconvolution approaches. Another source of variation in the quantification of the C5 peak is that it overlaps with the base of the upfield C4,6 peak of *myo*-IP_6, and possibly peaks from uridine-2′-monophosphate, adenosine-2′-diphosphate, and some unidentified compounds.[33] There is also evidence that the broad peak itself is comprised of more than one component,[9,13] which may result in an imperfect Lorentzian/Gaussian distribution within the phosphomonoester region. Nevertheless, our study shows that the quantification of *myo*-IP_6 using SDF with a broad signal generally results in measures of *myo*-IP_6 that are more accurate and consistent across a diversity of soils compared to that of SDF without a broad signal.

The comparison of the peak ratios to the theoretical value of 1 : 2 : 2 : 1 was used in previous studies to evaluate the accuracy of the applied SDF procedure.[42] In our study, the peak ratios of *myo*-IP_6 were similar between the two SDF approaches and close to the theoretical ratio of 1 : 2 : 2 : 1.[43] Therefore, the ratio of peaks from *myo*-IP_6 cannot be used to assess the efficacy of the SDF for accurate quantification of organic P compounds. Since the peak ratios of *myo*-IP_6 were not a useful assessment of the SDF approaches, it is likely that other organic P compounds exhibiting sharp signals may provide some insight on the validity of the SDF approach. In particular, an overestimation would be likely for the C2,5 peak of *neo*-IP_6 in the 4-eq/2-ax conformation,[3] since its chemical shift is present in the region of the broad peak. Whereas the C1,3,4,6 peak of *neo*-IP_6 is located upfield of the orthophosphate peak and therefore does not overlap with the broad signal. We identified both of these peaks (δ 5.92 and 3.78 ppm) through spiking in soils S3, S4, S5 and S6 (data not shown), calculated their ratios, and compared to the theoretical ratio of 4 : 2.[3] Calculated ratios were on average 4.0 : 6.2 when carrying out SDF without a broad signal and 4.0 : 1.5 with an underlying broad signal. These results provide supporting evidence that a broad signal should be included when carrying out SDF.

In the current study, there was a large number of sharp peaks detected in the phosphomonoester region, which was likely due to optimized extraction techniques and high resolution NMR.[8] Interestingly, the intensity of several unidentified sharp peaks increased when the *myo*-IP_6 standard was added (Fig. 2). The *myo*-IP_6 standard contained many phosphomonoester impurities (see Fig. SI-2 in the ESI†), which are likely those of lower order *myo*-IPs.[44] Doolette and Smernik (2018)[44] investigated the chemical composition of a variety of purchased or synthesized phytate standards using solution ^{31}P NMR spectroscopy. The authors found that the majority of phytate standards were impure and contained a mixture of lower- (and higher-) order IP, and orthophosphate. The authors suggested that thermal degradation during storage was the primary mechanisms in the degradation of higher-order IP to lower-order IP. This suggests the presence of lower order IP in soil extracts that could be detected using solution ^{31}P NMR spectroscopy.

4.3. Concentrations of *myo*-IP_6 (and the broad signal) in soil

Our results demonstrate that a quantitative determination of *myo*-IP_6 in soil extracts using solution ^{31}P NMR spectroscopy requires SDF procedures that include an underlying broad signal. In the current study, concentrations of *myo*-IP_6 and their proportion to the total pool of P_{org} in soil were generally in range or lower than that more broadly reported in the literature. This most likely reflects the majority of published studies that primarily carry out SDF without a broad signal. Clearly, pools of *myo*-IP_6 in soil are an important portion of the soil P_{org}, which are found in the majority of soils across the world.[45] However, they do not account for the majority of P_{org} in soil and sometimes found at negligible concentrations in some soils.[46–48]

The largest pool of soil P_{org} was that of the broad signal across all soils. This is consistent with previous studies, where it generally comprises 40–70% of the total P_{org} in soil.[14,16,19] The exact chemical nature of this pool remains unclear but can be described as phosphomonoesters in the form of large molecular structures,[14] which contain pools of P_{org} resistant to enzymatic hydrolysis,[19] associated with humic fractions,[49] and are structurally complex.[13] The presence of a broad peak is consistent with previous studies using non-NMR techniques that report a large proportion of the P_{org} in soil is unresolved and can occur in large molecular weight fractions.[2,28,48,50]

5. Conclusion

myo-Inositol hexakisphosphate is an important pool of soil organic P. However, its accurate quantification using solution ^{31}P NMR spectroscopy followed by SDF is uncertain. Our aim was to compare the recovery of added *myo*-IP_6 using two SDF procedures in NMR spectra on soil extracts. The average recovery of total added *myo*-IP_6 by SDF with a broad signal was close to 100% and exhibited less variation than that by SDF without a broad signal. The recovery of individual peaks of *myo*-IP_6 differed between its four peaks, which was overestimated for the C5 phosphate peak by up to 140% when a broad peak was not fitted. We recommend that the accurate quantification of *myo*-IP_6 using solution ^{31}P NMR spectroscopy on soil extracts includes a broad signal when carrying out SDF. This is also relevant for other sharp signals in the phosphomonoester region, which overlay the broad signal. Furthermore, our results show that previous studies reporting concentrations of *myo*-IP_6 using SDF without an underlying broad signal may be unreliable. It is essential that pools of *myo*-IP_6 (and the broad signal) are accurately determined for an improved understanding of the abundance and cycling of organic P in soil.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

The authors kindly thank Dr Laurie Paule Mauclaire Schönholzer, Dr Federica Tamburini, Mr Björn Studer, Ms Monika Macsai, and Dr Charles Brearley for their technical support. Furthermore, the authors are grateful to Dr Astrid Oberson, Dr David Lester, Dr Chiara Pistocchi and Dr Gregor Meyer for providing soil samples. The authors are very thankful for the use of the D-*chiro*- and *neo*-IP_6 standards originating from the Dr Dennis Cosgrove collection and the Dr Max Tate collection, respectively. Funding from the Swiss National Science Foundation [grant number 200021_169256] is gratefully acknowledged.

References

1 G. Anderson, in *The role of phosphorus in agriculture*, ed. F. E. Khasawneh, E. C. Sample and E. J. Kamprath, American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI, 1980, pp. 411–431, DOI: 10.2134/1980.roleofphosphorus.c16.

2 H. F. Harrison, *Soil organic phosphorus a review of world literature*, CAB International, Wallingford, 1987.

3 B. L. Turner, A. W. Cheesman, H. Y. Godage, A. M. Riley and B. V. Potter, Determination of neo- and D-chiro-inositol hexakisphosphate in soils by solution ^{31}P NMR spectroscopy, *Environ. Sci. Technol.*, 2012, **46**, 4994–5002.

4 G. C. J. Irving and D. J. Cosgrove, The use of gas-liquid chromatography to determine the proportions of inositol isomers present as pentakis- and hexakisphosphates in alkaline extracts of soils, *Commun. Soil Sci. Plant Anal.*, 1982, **13**, 957–967.

5 T. I. McLaren, R. J. Smernik, M. J. McLaughlin, A. L. Doolette, A. E. Richardson and E. Frossard, in *Advances in Agronomy*, ed. D. L. Sparks, Academic Press, 2020, vol. 160, pp. 51–124.

6 M. Ognalaga, E. Frossard and F. Thomas, Glucose-1-phosphate and myo-inositol hexaphosphate adsorption mechanisms on goethite, *Soil Sci. Soc. Am. J.*, 1994, **58**, 332–337.

7 R. H. Newman and K. R. Tate, Soil phosphorus characterisation by ^{31}P nuclear magnetic resonance, *Commun. Soil Sci. Plant Anal.*, 1980, **11**, 835–842.

8 B. Cade-Menun and C. W. Liu, Solution phosphorus-31 nuclear magnetic resonance spectroscopy of soils from 2005 to 2013: a review of sample preparation and experimental parameters, *Soil Sci. Soc. Am. J.*, 2014, **78**, 19–37.

9 A. L. Doolette and R. J. Smernik, Quantitative analysis of ^{31}P NMR spectra of soil extracts – dealing with overlap of broad and sharp signals, *Magn. Reson. Chem.*, 2015, **53**, 679–685.

10 B. L. Turner, N. Mahieu and L. M. Condron, Quantification of myo-inositol hexakisphosphate in alkaline soil extracts by solution ^{31}P spectroscopy and spectral deconvolution, *Soil Sci.*, 2003, **168**, 469–478.

11 J. E. Hill and B. J. Cade-Menun, Phosphorus-31 nuclear magnetic resonance spectroscopy transect study of poultry operations on the Delmarva Peninsula, *J. Environ. Qual.*, 2009, **38**, 130–138.

12 E. K. Bünemann, R. J. Smernik, P. Marschner and A. M. McNeill, Microbial synthesis of organic and condensed forms of phosphorus in acid and calcareous soils, *Soil Biol. Biochem.*, 2008, **40**, 932–946.

13 T. I. McLaren, R. Verel and E. Frossard, The structural composition of soil phosphomonoesters as determined by solution ^{31}P NMR spectroscopy and transverse relaxation (T2) experiments, *Geoderma*, 2019, **345**, 31–37.

14 T. I. McLaren, R. J. Smernik, M. J. McLaughlin, T. M. McBeath, J. K. Kirby, R. J. Simpson, C. N. Guppy, A. L. Doolette and A. E. Richardson, Complex forms of soil organic phosphorus – a major component of soil phosphorus, *Environ. Sci. Technol.*, 2015, **49**, 13238–13245.

15 A. L. Doolette, R. J. Smernik and W. J. Dougherty, Rapid decomposition of phytate applied to a calcareous soil demonstrated by a solution ^{31}P NMR study, *Eur. J. Soil Sci.*, 2010, **61**, 563–575.

16 A. L. Doolette, R. J. Smernik and W. J. Dougherty, Overestimation of the importance of phytate in NaOH-EDTA soil extracts as assessed by ^{31}P NMR analyses, *Org. Geochem.*, 2011, **42**, 955–964.

17 E. K. Bünemann, R. J. Smernik, A. L. Doolette, P. Marschner, R. Stonor, S. A. Wakelin and A. M. McNeill, Forms of phosphorus in bacteria and fungi isolated from two Australian soils, *Soil Biol. Biochem.*, 2008, **40**, 1908–1915.

18 S. R. Noack, M. J. McLaughlin, R. J. Smernik, T. M. McBeath and R. D. Armstrong, Crop residue phosphorus: speciation and potential bio-availability, *Plant Soil*, 2012, **359**, 375–385.

19 K. A. Jarosch, A. L. Doolette, R. J. Smernik, F. Tamburini, E. Frossard and E. K. Bünemann, Characterisation of soil organic phosphorus in NaOH-EDTA extracts: a comparison of ^{31}P NMR spectroscopy and enzyme addition assays, *Soil Biol. Biochem.*, 2015, **91**, 298–309.

20 FAO and I. W. Group, *World reference base for soil resources 2014*, Food and Agriculture Organization of the United Nations FAO, Rome, 2014.

21 S. Bühler, A. Oberson, S. Sinaj, D. K. Friesen and E. Frossard, Isotope methods for assessing plant available phosphorus in acid tropical soils, *Eur. J. Soil Sci.*, 2003, **54**, 605–616.

22 E. K. Bünemann, S. Augstburger and E. Frossard, Dominance of either physicochemical or biological phosphorus cycling processes in temperate forest soils of contrasting phosphate availability, *Soil Biol. Biochem.*, 2016, **101**, 85–95.

23 G. Meyer, E. K. Bünemann, E. Frossard, M. Maurhofer, P. Mäder and A. Oberson, Gross phosphorus fluxes in a calcareous soil inoculated with Pseudomonas protegens CHA0 revealed by ^{33}P isotopic dilution, *Soil Biol. Biochem.*, 2017, **104**, 81–94.

24 W. M. H. Saunders and E. G. Williams, Observation on the determination of total organic phosphorus in soils, *J. Soil Sci.*, 1955, **6**, 254–267.

25 S. Kuo, in *Methods of soil analysis part 3—Chemical methods*, ed. D. L. Sparks, A. L. Page, P. A. Helmke and R. H. Loeppert, Soil Science Society of America, American Society of Agronomy, Madison, WI, 1996, pp. 869–919, DOI: 10.2136/sssabookser5.3.c32.

26 B. J. Cade-Menun, C. W. Liu, R. Nunlist and J. G. McColl, Soil and litter phosphorus-31 nuclear magnetic resonance spectroscopy, *J. Environ. Qual.*, 2002, **31**, 457–465.

27 T. Ohno and L. M. Zibilske, Determination of low concentrations of phosphorus in soil extracts using malachite green, *Soil Sci. Soc. Am. J.*, 1991, **55**, 892–895.

28 R. A. Bowman and J. O. Moir, Basic EDTA as an extractant for soil organic phosphorus, *Soil Sci. Soc. Am. J.*, 1993, **57**, 1516–1518.

29 M. D. R. Vaz, A. C. Edwards, C. A. Shand and M. Cresser, Determination of dissolved organic phosphorus in soil solutions by an improved automated photo-oxidation procedure, *Talanta*, 1992, **39**, 1479–1487.

30 A. G. Vincent, J. Vestergren, G. Gröbner, P. Persson, J. Schleucher and R. Giesler, Soil organic phosphorus transformations in a boreal forest chronosequence, *Plant Soil*, 2013, **367**, 149–162.

31 A. V. Spain, M. Tibbett, M. Ridd and T. I. McLaren, Phosphorus dynamics in a tropical forest soil restored after strip mining, *Plant Soil*, 2018, **427**, 105–123.

32 A. L. Doolette, R. J. Smernik and W. J. Dougherty, Spiking improved solution phosphorus-31 nuclear magnetic resonance identification of soil phosphorus compounds, *Soil Sci. Soc. Am. J.*, 2009, **73**, 919–927.

33 J. Vestergren, A. G. Vincent, M. Jansson, P. Persson, U. Ilstedt, G. Gröbner, R. Giesler and J. Schleucher, High-resolution characterization of organic phosphorus in soil extracts using 2D ^{1}H–^{31}P NMR correlation spectroscopy, *Environ. Sci. Technol.*, 2012, **46**, 3950–3956.

34 M. I. Makarov, L. Haumaier and W. Zech, Nature of soil organic phosphorus: an assessment of peak assignments in the diester region of ^{31}P NMR spectra, *Soil Biol. Biochem.*, 2002, **34**, 1467–1477.

35 B. L. Turner, N. Mahieu and L. M. Condron, Phosphorus-31 nuclear magnetic resonance spectral assignments of phosphorus compounds in soil NaOH-EDTA extracts, *Soil Sci. Soc. Am. J.*, 2003, **67**, 497–510.

36 R. L. Vold, J. S. Waugh, M. P. Klein and D. E. Phelps, Measurement of spin relaxation in complex systems, *J. Chem. Phys.*, 1968, **48**, 3831–3832.

37 B. L. Turner, Soil organic phosphorus in tropical forests: an assessment of the NaOH-EDTA extraction procedure for quantitative analysis by solution ^{31}P NMR spectroscopy, *Eur. J. Soil Sci.*, 2008, **59**, 453–466.

38 B. L. Turner, N. Mahieu and L. M. Condron, The phosphorus composition of temperate pasture soils determined by NaOH-EDTA extraction and solution ^{31}P NMR spectroscopy, *Org. Geochem.*, 2003, **34**, 1199–1210.

39 W. J. Dougherty, R. J. Smernik and D. J. Chittleborough, Application of spin counting to the solid-state ^{31}P NMR analysis of pasture soils with varying phosphorus content, *Soil Sci. Soc. Am. J.*, 2005, **69**, 2058–2070.

40 T. I. McLaren, R. J. Simpson, M. J. McLaughlin, R. J. Smernik, T. M. McBeath, C. N. Guppy and A. E. Richardson, An assessment of various measures of soil phosphorus and the net accumulation of phosphorus in fertilized soils under pasture, *J. Plant Nutr. Soil Sci.*, 2015, **178**, 543–554.

41 A. L. Doolette, R. J. Smernik and W. J. Dougherty, A quantitative assessment of phosphorus forms in some Australian soils, *Soil Res.*, 2011, **49**, 152–165.

42 A. G. Vincent, J. Schleucher, G. Gröbner, J. Vestergren, P. Persson, M. Jansson and R. Giesler, Changes in organic phosphorus composition in boreal forest humus soils: the role of iron and aluminium, *Biogeochemistry*, 2012, **108**, 485–499.

43 A. J. R. Costello, T. Glonek and T. C. Myers, ^{31}P nuclear magnetic resonance-pH titrations of myo-inositol hexaphosphate, *Carbohydr. Res.*, 1976, **46**, 159–171.

44 A. L. Doolette and R. J. Smernik, Facile decomposition of phytate in the solid-state: kinetics and decomposition pathways, *Phosphorus, Sulfur, Silicon Relat. Elem.*, 2018, **193**, 192–199.

45 B. L. Turner, M. J. Papházy, P. M. Haygarth and I. D. McKelvie, Inositol phosphates in the environment, *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.*, 2002, **357**, 449.

46 B. L. Turner, Organic phosphorus in Madagascan rice soils, *Geoderma*, 2006, **136**, 279–288.

47 C. Williams and G. Anderson, Inositol phosphates in some Australian soils, *Soil Res.*, 1968, **6**, 121–130.

48 J. H. Steward and M. E. Tate, Gel chromatography of soil organic phosphorus, *J. Chromatogr. A*, 1971, **60**, 75–82.

49 Z. He, D. C. Olk and B. J. Cade-Menun, Forms and lability of phosphorus in humic acid fractions of Hord silt loam soil, *Soil Sci. Soc. Am. J.*, 2011, **75**, 1712–1722.

50 R. C. Dalai, in *Advances in Agronomy*, ed. N. C. Brady, Academic Press, 1977, vol. 29, pp. 83–117.

Biogeosciences, 17, 5079–5095, 2020
https://doi.org/10.5194/bg-17-5079-2020

Biogeosciences Open Access EGU

Identification of lower-order inositol phosphates (IP_5 and IP_4) in soil extracts as determined by hypobromite oxidation and solution ^{31}P NMR spectroscopy

Jolanda E. Reusser[1], René Verel[2], Daniel Zindel[2], Emmanuel Frossard[1], and Timothy I. McLaren[1]

[1]Department of Environmental Systems Science, ETH Zurich, Lindau, 8325, Switzerland
[2]Department of Chemistry and Applied Biosciences, ETH Zurich, Zurich, 8093, Switzerland

Correspondence: Jolanda E. Reusser (jolanda.reusser@usys.ethz.ch)

Received: 28 October 2019 – Discussion started: 28 November 2019
Revised: 20 August 2020 – Accepted: 26 August 2020 – Published: 21 October 2020

Abstract. Inositol phosphates (IPs) are a major pool of identifiable organic phosphorus (P) in soil. However, insight into their distribution and cycling in soil remains limited, particularly of lower-order IP (IP_5 and IP_4). This is because the quantification of lower-order IP typically requires a series of chemical extractions, including hypobromite oxidation to isolate IP, followed by chromatographic separation. Here, for the first time, we identify the chemical nature of organic P in four soil extracts following hypobromite oxidation using solution ^{31}P NMR spectroscopy and transverse relaxation (T_2) experiments. Soil samples analysed include A horizons from a Ferralsol (Colombia), a Cambisol and a Gleysol from Switzerland, and a Cambisol from Germany. Solution ^{31}P nuclear magnetic resonance (NMR) spectra of the phosphomonoester region in soil extracts following hypobromite oxidation revealed an increase in the number of sharp signals (up to 70) and an on average 2-fold decrease in the concentration of the broad signal compared to the untreated soil extracts. We identified the presence of four stereoisomers of IP_6, four stereoisomers of IP_5, and *scyllo*-IP_4. We also identified for the first time two isomers of *myo*-IP_5 in soil extracts: *myo*-(1,2,4,5,6)-IP_5 and *myo*-(1,3,4,5,6)-IP_5. Concentrations of total IP ranged from 1.4 to 159.3 mg P per kg soil across all soils, of which between 9 % and 50 % were comprised of lower-order IP. Furthermore, we found that the T_2 times, which are considered to be inversely related to the tumbling of a molecule in solution and hence its molecular size, were significantly shorter for the underlying broad signal compared to for the sharp signals (IP_6) in soil extracts following hypobromite oxidation. In summary, we demonstrate the presence of a plethora of organic P compounds in soil extracts, largely attributed to IPs of various orders, and provide new insight into the chemical stability of complex forms of organic P associated with soil organic matter.

1 Introduction

Inositol phosphates (IPs) are found widely in nature and are important for cellular functions in living organisms. They are found in eukaryotic cells where they operate in ion-regulation processes, as signalling or P storage compounds (Irvine and Schell, 2001). The basic structure of IP consists of a carbon ring (cyclohexanehexol) with one to six phosphorylated centres (IP_{1-6}) and up to nine stereoisomers (Angyal, 1963; Cosgrove and Irving, 1980). An important IP found in nature is *myo*-IP_6, which is used as a P storage compound in plant seeds. Another important species of IP is that of *myo*-(1,3,4,5,6)-IP_5, which is present in most eukaryotic cells at concentrations ranging from 15 to 50 µM (Riley et al., 2006). Species of IP_{1-3} are present in phospholipids such as phosphatidylinositol diphosphates and are an essential structural component of the cell membrane system (Strickland, 1973; Cosgrove and Irving, 1980).

Inositol phosphates have been reported to comprise more than 50 % of total organic phosphorus (P_{org}) in some soils (Cosgrove and Irving, 1980; McDowell and Stewart, 2006; Turner, 2007). Four stereoisomers of IP have been detected in soils, with the *myo* stereoisomer being the most abundant (56 %), followed by *scyllo* (33 %), *neo* and D-*chiro* (11 %;

Cosgrove and Irving, 1980; Turner et al., 2012). The largest input of *myo*-IP_6 to the soil occurs via the addition of plant seeds (Turner et al., 2002). However, the addition of *myo*-IP_6 to soil can also occur via manure input because monogastric animals are mostly incapable of digesting *myo*-IP_6 without the addition of phytases to their diets (Leytem et al., 2004; Leytem and Maguire, 2007; Turner et al., 2007b). An exception to this is pigs, which were found to at least partially digest phytate (Leytem et al., 2004), and transgenic pigs expressing salivary phytase (Golovan et al., 2001; Zhang et al., 2018). The accumulation of *myo*-IP_6 in soil occurs due to the negative charge of the deprotonated phosphate groups, which can coordinate on the charged surfaces of Fe- and Al-(hydro)-oxides (Anderson et al., 1974; Ognalaga et al., 1994), clay minerals (Goring and Bartholomew, 1951), and soil organic matter (SOM; McKercher and Anderson, 1989) or form insoluble precipitates with cations (Celi and Barberis, 2007). These processes lead to the stabilisation of IP in soil resulting in its accumulation and reduced bioavailability (Turner et al., 2002). In contrast, the sources and mechanisms controlling the flux of *scyllo*-, *neo*- and D-*chiro*-IP_6 in soil remain unknown but are thought to involve epimerisation of the *myo* stereoisomer (L'Annunziata, 1975).

Chromatographic separation of alkaline soil extracts revealed the presence of four stereoisomers of IP_6 and lower-order IP_{1-5} (Halstead and Anderson, 1970; Anderson and Malcolm, 1974; Cosgrove and Irving, 1980; Irving and Cosgrove, 1982). Irving and Cosgrove (1981) used hypobromite oxidation prior to chromatography to isolate the IP fraction in alkaline soils. The basis of this approach is that IPs are considered to be highly resistant to hypobromite oxidation, whereas other organic compounds (e.g. phospholipids and nucleic acids) will undergo oxidation (Dyer and Wrenshall, 1941; Turner and Richardson, 2004). The resistance of IP to hypobromite oxidation is thought to be due to the high charge density and steric hindrance, which is caused by the chair conformation of the molecule and the bound phosphate groups, with the P in its highest oxidation state. Hypobromite oxidation of inositol (without phosphate groups) mainly results in the formation of inososes, which have an intact carbon ring (Fatiadi, 1968). Fatiadi (1968) considered that the oxidation of bromine with inositol is stereospecific and comparable to catalytic or bacterial oxidants.

A limitation of chromatographic separation of alkaline extracts is that there is a mixture of unknown organic compounds that can co-elute with IP and result in an overestimation of IP concentrations (Irving and Cosgrove, 1981). However, this can also occur for IP, and historically, studies often reported the combined concentration of IP_6 and IP_5 due to a lack of differentiation in their elution times (McKercher and Anderson, 1968b). More recently, Almeida et al. (2018) investigated how cover crops might mobilise soil IP using hypobromite oxidation on sodium hydroxide–disodium ethylenediaminetetraacetic acid (NaOH-EDTA) extracts followed by chromatographic separation. The authors found that pools of *myo*-IP_6 and "unidentified IP" accounted for 30 % of the total extractable pool of P and hypothesised that the unidentified-IP pool consists solely of lower-order *myo*-IP. Pools of lower-order IP_{1-5} comprise on average 17 % of the total pool of IP in soil and account for an important pool of soil organic P in terrestrial ecosystems (Anderson and Malcolm, 1974; Cosgrove and Irving, 1980; Turner et al., 2002; Turner, 2007).

Since the 1980s, solution ^{31}P nuclear magnetic resonance (NMR) spectroscopy has been the most commonly used technique to characterise the chemical nature of organic P in soil extracts (Newman and Tate, 1980; Cade-Menun and Liu, 2014). An advantage of this technique is the simultaneous detection of all forms of organic P that come into solution, which is brought about by a single-step extraction with alkali and a chelating agent (Cade-Menun and Preston, 1996). However, a limitation of the technique has been the loss of information on the diversity and amount of soil IP compared to that typically obtained prior to 1980 (Smith and Clark, 1951; Anderson, 1955; Cosgrove, 1963). To date, solution ^{31}P NMR spectroscopy on soil extracts has only reported concentrations of *myo*-, *scyllo*-, *chiro*- and *neo*-IP_6. The fact that lower-order IPs were not reported in studies using NMR spectroscopy might be due to overlap of peaks in the phosphomonoester region, which makes peak assignment of specific compounds difficult (Doolette et al., 2009).

Turner et al. (2012) carried out hypobromite oxidation prior to solution ^{31}P NMR analysis of alkaline soil extracts to isolate the IP fraction. This had the advantage of reducing the number of NMR signals in the phosphomonoester region and consequently the overlap of peaks. The authors demonstrated the presence of *neo*- and *chiro*-IP_6 in NMR spectra via spiking of hypobromite oxidised extracts. Interestingly, the authors also reported the presence of NMR signals in the phosphomonoester region that could not be assigned to IP_6 and were resistant to hypobromite oxidation. They were not able to attribute the NMR signals to any specific P compounds but hypothesised based on their resistance to hypobromite oxidation that they were due to lower-order IP.

The aim of this study was to identify and quantify IP in soil extracts following hypobromite oxidation using solution ^{31}P NMR spectroscopy. In addition, the structural composition of phosphomonoesters in soil extracts following hypobromite oxidation was probed using solution ^{31}P NMR spectroscopy and transverse relaxation experiments. We hypothesise that a large portion of sharp peaks in the phosphomonoester region of untreated soil extracts are resistant to hypobromite oxidation, which would indicate the presence of a wide variety of IP. This would have major consequences on our understanding of P cycling in terrestrial (and aquatic) ecosystems, as many more organic P compounds and mechanisms would be involved than previously thought. Furthermore, a better understanding of these organic P compounds in soil would also help improve strategies to increase their biological utilisation, which may reduce the amount of fertiliser needed in

agricultural systems and thus influence the transfer of P to aquatic and marine ecosystems.

2 Experimental section

2.1 Soil collection and preparation

Soil samples were collected from the upper horizon of the profile at four diverse sites. These comprise a Ferralsol from Colombia, a Vertisol from Australia, a Cambisol from Germany, and a Gleysol from Switzerland (FAO and Group, 2014). The four soil samples were chosen from a larger collection based on their diverse concentration of P_{org} and composition of the phosphomonoester region in NMR spectra (Reusser et al., 2020). Background information and some chemical properties of the soils are reported in Table 1. Briefly, the Ferralsol was collected from an improved grassland in 1997 at the Carimagua Research Station's long-term Culticore field experiment in Columbia (Bühler et al., 2003). The Vertisol was collected from an arable field in 2018 located in southern Queensland. The site had been under native shrubland prior to 1992. The Cambisol was collected from a beech forest in 2014 and is part of the SPP 1685 – Ecosystem Nutrition project (Bünemann et al., 2016; Lang et al., 2017). The Gleysol was collected from the peaty top soil layer of a drained marshland in 2017, which has been under grassland for at least 20 years.

Soil samples were passed through a 5 mm sieve and dried at 60 °C for 5 d, except for the Ferralsol (sieved < 2 mm) and the Vertisol (ground < 2 mm), which were received dried. Total concentrations of C and N in soils were obtained using combustion of 50 mg of ground soil (to powder) weighed into tinfoil capsules (vario PYRO cube®, Elementar Analysensysteme GmbH). Soil pH was measured in H_2O with a soil-to-solution ratio of $1 : 2.5$ (w/w) using a glass electrode.

2.2 Soil phosphorus analyses

Total concentrations of soil P were obtained by X-ray fluorescence (XRF) spectroscopy (SPECTRO XEPOS ED-XRF, AMETEK®) using 4.0 g of soil sample ground to powder mixed with 0.9 g of wax (CEREOX Licowax, FLUXANA®). The XRF instrument was calibrated using commercially available reference soils. Concentrations of organic P for NMR analysis were obtained using the NaOH-EDTA extraction technique of Cade-Menun et al. (2002) at a soil-to-solution ratio of 1 : 10, i.e. extracting 4 g of soil with 40 mL of extractant.

2.3 Hypobromite oxidation

Hypobromite oxidation of NaOH-EDTA soil filtrates was carried out based on a modified version of the method described in Suzumura and Kamatani (1993) and Turner et al. (2012). The hypobromite oxidation procedure is similar to that reported in Turner (2020). Briefly, 10 mL of the NaOH-EDTA filtrate (Sect. 2.2) was placed in a three-necked round-bottom flask equipped with a septum, condenser, magnetic stir bar and thermometer (through a Claisen adapter with N_2 adapter). After the addition of 1 mL of 10 M aqueous NaOH and vigorous stirring, an aliquot of 0.6 mL Br_2 (which was cooled prior to use) was added, resulting in an exothermic reaction where some of the soil extracts nearly boiled. The optimal volume of Br_2 for oxidation was assessed in a previous pilot study using 0.2, 0.4, 0.6 and 0.8 mL Br_2 volumes and then observing differences in their NMR spectral features (Fig. S9). The reaction was heated to 100 °C within 10 min and kept at reflux for an additional 5 min. After cooling to room temperature, the solution was acidified with 2 mL of 6 M aqueous HCl solution in order to obtain a $pH < 3$, which was confirmed with a pH test strip. The acidified solution was reheated to 100 °C for 5 min under a stream of nitrogen to vaporise any excess bromine. The pH of the solution was gradually increased to 8.5 using 10 M aqueous NaOH solution. After dilution with 10 mL of H_2O, 5 mL of 50 % (w/w) ethanol and 10 mL of 10 % (w/w) barium acetate solution was added to the solution in order to precipitate any IP (Turner et al., 2012). The solution was then heated and boiled for 10 min and allowed to cool down overnight. The solution was subsequently transferred to a 50 mL centrifuge tube, and a 10 mL aliquot of 50 % (w/w) ethanol was added, manually shaken and centrifuged at 1500 g for 15 min. The supernatant was removed, and a 15 mL aliquot of 50 % (w/w) ethanol was added to the precipitate, shaken and then centrifuged again as before. The supernatant was removed and the process repeated once more to further purify the pool of IP. Afterwards, the precipitate was transferred with 20 mL of H_2O into a 100 mL beaker that contained a 20 mL volume (equating to a mass of 15 g) of Amberlite® IR-120 cation exchange resin beads in the H^+ form (Sigma-Aldrich, product no. 06428). The suspension was stirred for 15 min and then passed through a Whatman no. 42 filter paper. A 9 mL aliquot of the filtrate was frozen at −80 °C and then lyophilised prior to NMR analysis. This resulted in 18–26 mg of lyophilised material across all soils. Concentrations of total P in solutions were obtained using inductively coupled plasma optical emission spectrometry (ICP-OES). Concentrations of molybdate reactive P (MRP) were obtained using the malachite green method of Ohno and Zibilske (1991). The difference in concentrations of total P and MRP in solution is molybdate unreactive P (MUP), which is predominantly organic P for these samples. To assess the effect of hypobromite oxidation on the stability of an IP_6, duplicate samples of the Cambisol and the Gleysol were spiked with 0.1 mL of an 11 mM *myo*-IP_6 standard. The recovery of the added *myo*-IP_6 following hypobromite oxidation was calculated using Eq. (1):

$$\text{Spike recovery}\,(\%) = \frac{C_{\text{spiked}}\left(\frac{\text{mg}}{\text{L}}\right) - C_{\text{unspiked}}\left(\frac{\text{mg}}{\text{L}}\right)}{C_{\text{standard added}}\left(\frac{\text{mg}}{\text{L}}\right)}, \qquad (1)$$

Table 1. General characteristics of soil samples used in this study.

Soil type	Unit	Ferralsol	Vertisol	Cambisol	Gleysol
Country	–	Colombia	Australia	Germany	Switzerland
Coordinates sampling site	–	4°30′ N, 71°19′ W	27°52′ S, 151°37′ E	50°21′ N, 9°55′ E	47°05′ N, 8°06′ E
Elevation	m a.s.l.	150	402	800	612
Sampling depth	cm	0–20	0–15	0–7	0–10
Year of sampling	year	1997	2017	2014	2017
Land use	–	Pasture	Arable field	Forest	Pasture
C_{tot}	g C per kg soil	26.7	23.9	90.3	148.3
N_{tot}	g N per kg soil	1.7	1.9	6.6	10.9
pH in H_2O	–	3.6	6.1	3.6	5.0

where C_{spiked} and $C_{unspiked}$ are the concentrations of *myo*-IP_6 in NaOH-EDTA extracts following hypobromite oxidation of the spiked and unspiked samples, respectively. $C_{standard\ added}$ is the concentration of the added *myo*-IP_6 within the standard. As ^{31}P NMR spectroscopy of the standard revealed impurities, the concentration of *myo*-IP_6 in the standard was calculated based on the ^{31}P NMR spectrum.

2.4 Sample preparation for solution ^{31}P NMR spectroscopy

The lyophilised material of the untreated soil extracts was prepared for solution ^{31}P NMR spectroscopy based on a modification of the methods of Vincent et al. (2013) and Spain et al. (2018). Briefly, 120 mg of lyophilised material was taken and dissolved in 600 µL of 0.25 M NaOH–0.05 M Na_2 EDTA solution (ratio of 1 : 5). However, for the Cambisol sample, this ratio resulted in an NMR spectrum that exhibited significant line broadening. Therefore, this was repeated on a duplicate sample but at a smaller lyophilised-material-to-solution ratio (ratio of 1 : 7.5), as suggested in Cade-Menun and Liu (2014), which resolved the issue of poor spectral quality. The suspension was stored overnight to allow for complete hydrolysis of phospholipids and RNA (Doolette et al., 2009; Vestergren et al., 2012) and was then centrifuged at 10 621 g for 15 min. A 500 µL aliquot of the supernatant was taken, which was subsequently spiked with a 25 µL aliquot of a 0.03 M methylenediphosphonic acid (MDP) standard made in D_2O (Sigma-Aldrich, product no. M9508) and a 25 µL aliquot of sodium deuteroxide at 40 % (w/w) in D_2O (Sigma-Aldrich, product no. 372072). The solution was then mixed and transferred to a 5 mm diameter NMR tube.

A similar procedure was used for the soil extracts that had undergone hypobromite oxidation, except the total mass of lyophilised material (18–26 mg) was dissolved with 600 µL of a 0.25 M NaOH–0.05 M Na_2 EDTA solution. However, for the Cambisol sample, the NMR spectrum exhibited considerable line broadening, and an additional 400 µL aliquot of NaOH-EDTA solution was added to the NMR tube, mixed and then returned to the NMR spectrometer. This resolved the issue of poor spectral quality.

2.5 Solution ^{31}P NMR spectroscopy

Solution ^{31}P NMR analyses were carried out on all untreated and hypobromite oxidised soil extracts at the NMR facility of the Laboratory of Inorganic Chemistry (Hönggerberg, ETH Zurich). All spectra were obtained with a Bruker AVANCE III HD 500 MHz NMR spectrometer equipped with a cryogenic probe (CryoProbe™ Prodigy, Bruker Corporation, Billerica, MA). The ^{31}P frequency for this NMR spectrometer was 202.5 MHz, and gated broadband proton decoupling with a 90° pulse of 12 µs was applied. The spectral resolution under these conditions for ^{31}P was < 1 Hz. Longitudinal relaxation (T_1) times were determined for each sample with an inversion recovery experiment (Vold et al., 1968). This resulted in recycle delays ranging from 8.7 to 30.0 s for the untreated extracts and 7.8 to 38.0 s for the hypobromite oxidised soil extracts. The number of scans for the untreated extracts was set to 1024 or 4096, depending on the signal-to-noise ratio of the obtained spectrum. All hypobromite oxidised spectra were acquired with 3700 to 4096 scans.

2.6 Processing of NMR spectra

All NMR spectra were processed with Fourier transformation, phase correction and baseline adjustment within the TopSpin®software environment (version 3.5 pl 7, Bruker Corporation, Billerica, MA). Line broadening was set to 0.6 Hz. Quantification of NMR signals involved obtaining the integrals of the following regions: (1) up to four phosphonates (δ 19.8 to 16.4 ppm), (2) the added MDP (δ 17.0 to 15.8 ppm) including its two carbon satellite peaks, (3) the combined orthophosphate and phosphomonoester region (δ 6.0 to 3.0 ppm), (4) up to four phosphodiesters (δ 2.5 to -3.0 ppm), and (5) pyrophosphate (δ -4.8 to -5.4 ppm). Due to overlapping peaks in the orthophosphate and phosphomonoester region, spectral-deconvolution fitting (SDF) was applied as described in Reusser et al. (2020). In brief, the SDF procedure involved the fitting of an underlying broad

signal, based on the approach of Bünemann et al. (2008) and McLaren et al. (2019). We carried out the SDF with a non-linear optimisation algorithm in MATLAB® R2017a (The MathWorks, Inc.) and fitted visually identifiable peaks by constraining their linewidths at half height as well as the lower and upper boundary of the peak positions along with an underlying broad signal in the phosphomonoester region. The sharp signals of high intensity (e.g. orthophosphate) and the broad peak were fitted using Lorentzian line shapes, whereas sharp signals of low intensity were fitted using Gaussian line shapes. The NMR observability of total P (P_{tot}) in NaOH-EDTA extracts was calculated using Eq. (2) (Dougherty et al., 2005; Doolette et al., 2011b):

$$\text{NMR observability}\,(\%) = \frac{P_{\text{tot}}\ \text{NMR}}{P_{\text{tot}}\ \text{ICP-OES}} \times 100\,\%, \qquad (2)$$

where P_{tot} NMR refers to the total P content in milligrams of P per kilogram of soil detected in the soil extracts using solution ^{31}P NMR spectroscopy and P_{tot} ICP-OES refers to the total P concentration in milligrams of P per kilogram of soil measured in the soil extracts prior to freeze-drying using ICP-OES.

2.7 Spiking experiments

To identify the presence of IPs in hypobromite oxidised extracts, samples were spiked with a range of standards and then analysed again using NMR spectroscopy. This involved the addition of 5 to 20 µL aliquots of an IP standard solution directly into the NMR tube, which was then sealed with parafilm, manually shaken and then allowed to settle prior to NMR analysis. Each sample extract was consecutively spiked with no more than four IP standards. The NMR spectra of soil extracts after spiking were overlaid with the NMR spectra of unspiked soil extracts to identify the presence of IP across all soil samples. This comparison of NMR spectra was possible due to negligible changes in the chemical shifts of peaks among soil samples. The IP standards used in this study are listed in Table 2.

2.8 Transverse relaxation (T_2) experiments

Due to the presence of sharp and broad signals in the phosphomonoester region of NMR spectra on hypobromite oxidised soil extracts, transverse relaxation (T_2) experiments were carried out to probe their structural composition. The transverse relaxation (originally spin–spin relaxation) describes the loss of magnetisation in the x–y plane. This loss occurs due to magnetic-field differences in the sample, arising either by instrumentally caused magnetic-field inhomogeneities or by local magnetic fields in the sample caused by intramolecular and intermolecular interactions (Claridge, 2016). Generally, small, rapidly tumbling molecules exhibit longer T_2 relaxation times compared to large, slowly tumbling molecules (McLaren et al., 2019).

Briefly, solution ^{31}P NMR spectroscopy with a Carr–Purcell–Meiboom–Gill (CPMG) pulse sequence (Meiboom and Gill, 1958) was carried out on all hypobromite oxidised soil extracts, as described in McLaren et al. (2019). This involved a constant spin-echo delay (τ) of 5 ms, which was repeated for a total of eight iterations (spin-echo periods of 5, 50, 100, 150, 200, 250, 300 and 400 ms). A total of 4096 scans and a recycle delay of 4.75 s were used for all iterations. Transverse relaxation times for the aforementioned integral ranges were calculated using Eq. (3) within the TopSpin® software environment. Due to overlapping peaks in the orthophosphate and phosphomonoester region, spectral deconvolution was carried out to partition the NMR signal, as described in McLaren et al. (2019). The T_2 times of the partitioned NMR signals were calculated using Eq. (3) within RStudio© (version 1.1.442):

$$M\,(t) = M_0 \times e^{\left(-t \times T_2^{-1}\right)}, \qquad (3)$$

where M refers to the net magnetisation derived from the average angular momentum in the x–y plane, τ refers to the spin-echo delay in milliseconds (ms) and T_2 refers to the transverse relaxation time (ms).

2.9 Statistical analyses and graphics

Statistical analyses were carried out using Microsoft® Excel 2016 and MATLAB R2017a (© The MathWorks, Inc.). Graphics were created with Microsoft® Excel 2016 and MATLAB R2017a (© The MathWorks, Inc.). Solution (1D) ^{31}P NMR spectra were normalised to the peak intensity of MDP (δ 16.46 ppm). Spectra from the T_2 experiments were normalised to the peak intensity of *scyllo*-IP_6 (δ 3.22 ppm).

A one-way ANOVA was carried out in MATLAB R2017a (© The MathWorks, Inc.) with a subsequent multicomparison of mean values using Tukey's honestly significant difference procedure based on the Studentised range distribution (Hochberg and Tamhane, 1987; Milliken and Johnson, 2009).

3 Results

3.1 Phosphorus concentrations in soil extracts

Concentrations of total soil P as determined by XRF ranged from 320 to 3841 mg P per kg soil across all soils (Table 3). Concentrations of total P as estimated by the NaOH-EDTA extraction technique ranged from 160 to 1850 mg P per kg soil, which comprised 28 % to 51 % of the total soil P as determined by XRF. Pools of organic P comprised 28 % to 72 % of the total P in NaOH-EDTA untreated soil extracts.

Concentrations of total P in NaOH-EDTA soil extracts following hypobromite oxidation ranged from 77 to 578 mg P per kg soil (Table 3), which accounted for 31 % to 48 % (on average 38 %) of the total P originally present in the extracts.

Table 2. Standard solutions used for the spiking experiment of the hypobromite oxidised soil extracts. All standards were dissolved in 0.25 M NaOH and 0.05 M Na_2 EDTA.

Standard	Product number	Company or origin	Concentration of standard in NaOH-EDTA ($mg\,mL^{-1}$)
myo-IP_6	P5681	Merck (Sigma-Aldrich)	8.10
L-*chiro*-IP_6	Collection of Max Tate		2.39
D-*chiro*-IP_6	CAY-9002341	Cayman Chemical	2.00
neo-IP_6	Collection of Dennis Cosgrove*		4.62
D-*myo*-(1,2,4,5,6)-IP_5	CAY-10008452-1	Cayman Chemical	2.00
myo-(1,2,3,4,6)-IP_5	93987	Merck (Sigma-Aldrich)	2.00
D-*myo*-(1,3,4,5,6)-IP_5	CAY-10009851-1	Cayman Chemical	2.00
D-*myo*-(1,2,3,5,6)-IP_5	CAY-10008453-1	Cayman Chemical	2.00
scyllo-IP_5	Collection of Dennis Cosgrove		2.64
L-*chiro*-IP_5	Collection of Dennis Cosgrove		2.24
neo-IP_5	Collection of Dennis Cosgrove		2.45
myo-IP_4	Collection of Dennis Cosgrove		2.76
scyllo-IP_4	Collection of Dennis Cosgrove		2.41
neo-IP_4	Collection of Dennis Cosgrove		2.33

* Made up in 15 mM HCl.

Similarly, pools of organic P in NaOH-EDTA extracts following hypobromite oxidation were lower, comprising 22 % to 48 % (on average 36 %) of that originally present in untreated NaOH-EDTA extracts across all soils.

3.2 Solution ^{31}P NMR spectra of hypobromite oxidised soil extracts

The most prominent signal in the NMR spectra of untreated NaOH-EDTA soil extracts was that of orthophosphate at δ 5.25 ($\pm$0.25) ppm, followed by the phosphomonoester region ranging from δ 6.0 to 3.0 ppm (Fig. 1). There were also some minor signals due to pyrophosphate δ -5.06 ($\pm$0.19) ppm (all soils); phosphodiesters ranging from δ 2.5 to -2.4 ppm (not detected in the Vertisol); and phosphonates (not including the added MDP) at δ 19.8, 19.2 and 18.3 ppm (not detected in the Gleysol). However, these compounds comprised less than 8 % of the total NMR signal.

Following hypobromite oxidation of NaOH-EDTA extracts, the most prominent NMR signals were found in the orthophosphate (65 % of total NMR signal) and phosphomonoester (35 % of total NMR signal) region across all soils (Fig. 1). Phosphodiesters and pyrophosphates were removed following hypobromite oxidation in the Ferralsol, the Vertisol and the Cambisol (Germany). However, some signal remained in the Gleysol at low concentrations (0.4 % of the total NMR signal). Phosphonates were removed following hypobromite oxidation in the Ferralsol and the Vertisol, but a total of five sharp peaks in the phosphonate region were detected (δ 19.59, 18.58, 17.27 and 9.25 ppm) in the Cambisol. These peaks comprised 0.6 % of the total NMR signal.

The phosphomonoester region of NMR spectra on untreated NaOH-EDTA extracts exhibited two main features: (1) the presence of a broad signal centred at around δ 4.1 ($\pm$0.1) ppm with an average linewidth at half height of 256.12 Hz and (2) the presence of between 19 and 34 sharp signals. This was similarly the case for hypobromite oxidised extracts, except there was a decrease in the intensity of the broad signal and a change in the distribution and intensity of sharp signals. For the Cambisol and Gleysol, the number of sharp signals in the phosphomonoester region approximately doubled (to 40 and 70 sharp signals, respectively) following hypobromite oxidation. In contrast, less than half of the sharp signals remained in the Ferralsol following hypobromite oxidation (i.e. 14 of the 30 peaks originally present in the untreated extract), whereas one peak was removed following hypobromite oxidation in the Vertisol. There was little change (0.23 ppm) in the chemical shifts of peaks between the untreated and hypobromite oxidised extracts.

3.3 Identification and quantification of inositol phosphates (IP_6, IP_5 and IP_4) in soil extracts

Detailed views of the phosphomonoester regions of spiked samples are shown in Figs. S1 to S5 in the Supplement. The number of identified sharp peaks in the phosphomonoester region ranged from 7 (Vertisol) to 33 (Gleysol). *myo*- and *scyllo*-IP_6 were identified in the hypobromite oxidised extracts of all soils (Table 5). On average, 72 % of *myo*-IP_6 and 56 % of *scyllo*-IP_6 present in the untreated extracts remained in the hypobromite oxidised extracts (Table S1 in the Supplement). *neo*-IP_6 was identified in the 2-equatorial–4-axial

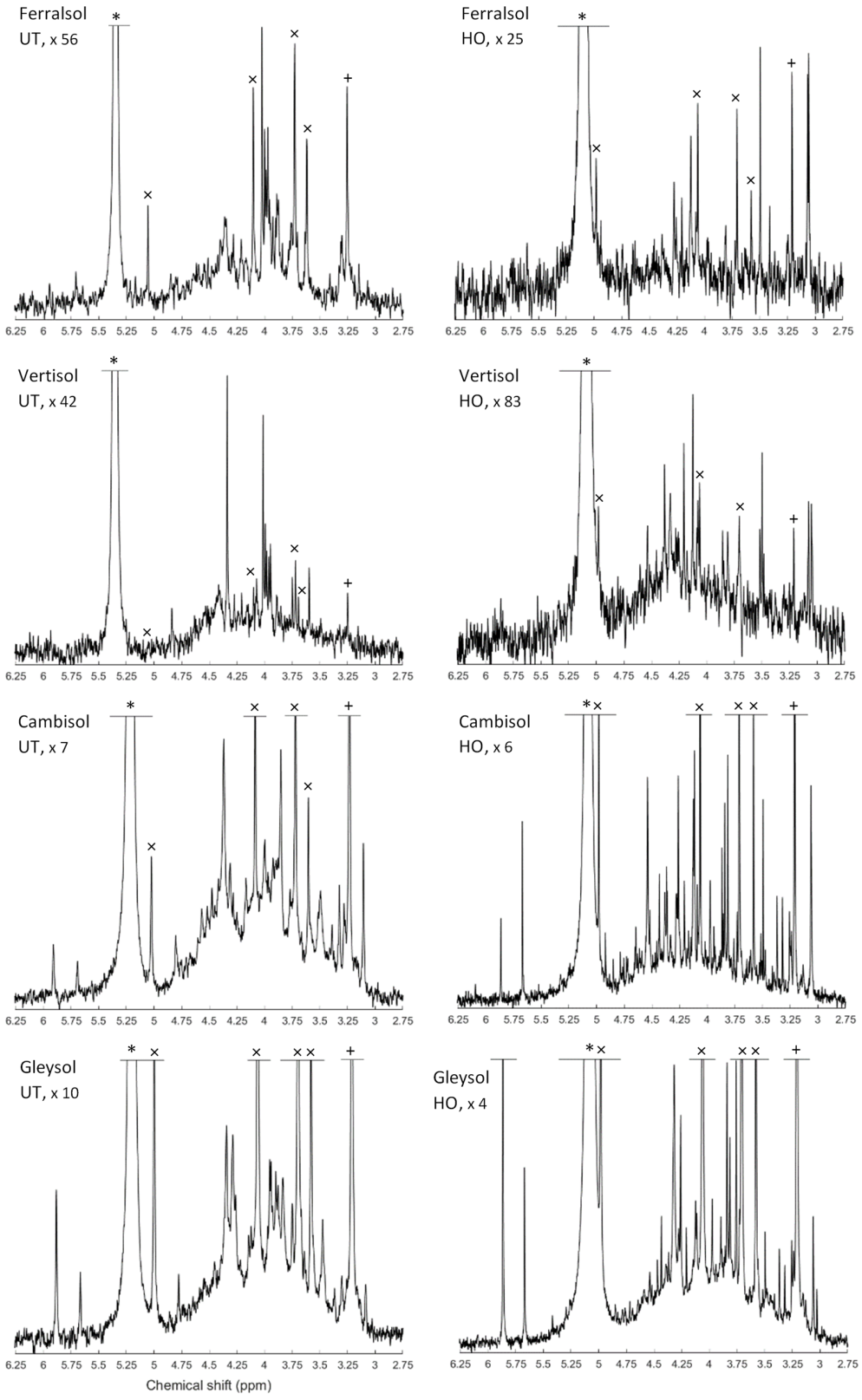


Figure 1. Solution ^{31}P nuclear magnetic resonance (NMR) spectra (500 MHz) of the orthophosphate and phosphomonoester region on untreated (UT) and hypobromite oxidised (HO) 0.25 M NaOH + 0.05 M EDTA soil extracts (Ferralsol, Vertisol, Cambisol and Gleysol). Signal intensities were normalised to the MDP peak intensity. The vertical axes were increased for improved visibility of spectral features, as indicated by a factor. The orthophosphate peak is marked with an asterisk. The symbol "×" marks the four individual peaks of *myo*-IP_6 and "+" the peak of *scyllo*-IP_6.

Table 3. Concentrations of total P as measured by XRF and 0.25 M NaOH + 0.05 M EDTA extractable P before and after hypobromite oxidation of soil extracts. Concentrations of total P in NaOH-EDTA extracts were determined by ICP-OES, whereas the concentration of molybdate reactive P (MRP) was determined by the malachite green method of Ohno and Zibilske (1991). Concentrations of molybdate unreactive P (MUP) were calculated as the difference between total P and MRP.

Measure		Ferralsol	Vertisol	Cambisol	Gleysol
XRF	P_{tot} (mg P per kg soil)	320	1726	3841	2913
NaOH-EDTA extractable P (untreated)	P_{tot} (mg P per kg soil)	160	484	1850	1490
	MRP (mg P per kg soil)	67	351	525	610
	MUP (P_{org}; mg P per kg soil)	93	133	1326	880
NaOH-EDTA extractable P (hypobromite oxidised)	P_{tot} (mg P per kg soil)	77	158	580	578
	MRP (mg P per kg soil)	32	111	283	231
	MUP (P_{org}; mg P per kg soil)	45	47	297	348

Table 4. Concentrations (mg P per kg soil) of P compounds in solution ^{31}P NMR spectra of 0.25 M NaOH + 0.05 M EDTA soil extracts (Ferralsol, Vertisol, Cambisol and Gleysol) before and after hypobromite oxidation (HO). Quantification was based on spectral integration and deconvolution fitting. The proportion of P detected in hypobromite oxidised extracts compared to that in untreated extracts is provided in brackets.

Phosphorus class		Ferralsol	Vertisol	Cambisol	Gleysol
Phosphonates	before HO	1.0	2.6	14.5	–
	after HO	–	–	3.0 (21)	0.2
Orthophosphate	before HO	54.8	221.4	434.3	368.3
	after HO	32.0 (58)	116.6 (53)	329.3 (76)	243.4 (66)
Phosphomonoester	before HO	36.3	39.1	501.1	399.2
	after HO	12.7 (35)	24.2 (62)	210.3 (42)	292.1 (73)
Broad peak in phosphomonoester region	before HO	21.6	30.9	305.8	216.7
	after HO	8.3 (39)	19.3 (63)	99.2 (32)	108.4 (50)
Phosphodiester	before HO	5.1	–	28.2	26.9
	after HO	–	–	–	2.0 (8)
Pyrophosphate	before HO	1.9	1.8	12.9	23.9
	after HO	–	–	–	–

and 4-equatorial–2-axial conformations and *chiro*-IP_6 in the 2-equatorial–4-axial confirmation of the oxidised extracts in the Cambisol and Gleysol but was absent in the Ferralsol and the Vertisol (Figs. S4 and S5 in the Supplement).

The *myo*, *scyllo*, *chiro* and *neo* stereoisomers of IP_5 were identified in various hypobromite oxidised extracts (Table 5). Two isomers of *myo*-IP_5 were identified in some extracts, which included *myo*-(1,2,4,5,6)-IP_5 and *myo*-(1,3,4,5,6)-IP_5. In addition, *scyllo*-IP_4 was detected in all soils except that of the Vertisol. There was insufficient evidence for the presence of *myo*-IP_4 in these soil samples, as only one of the two peaks of this compound was present in the NMR spectra of untreated extracts. This could possibly be due to the partial dephosphorylation of *myo*-IP_4 during the hypobromite oxidation procedure. The reason for the reduced resistance of lower-order IPs to hypobromite oxidation compared to IP_{5+6} might be due to their reduced steric hindrance and charge density, as fewer phosphate groups are bound to the inositol ring.

Concentrations of total IP ranged from 1.4 to 159.3 mg P per kg soil across all soils, which comprised between 1 % (Vertisol) and 18 % (Gleysol) of the organic P in untreated NaOH-EDTA extracts (Table 3). Pools of IP_6 were the most abundant form of IP, which ranged from 0.9 to 144.8 mg P per kg soil across all soils (Table 5). The proportion of IP_6 stereoisomers across all soils was on the order of *myo* (61 %, SD = 12), *scyllo* (29 %, SD = 3), *chiro* (6 %, SD = 8) and *neo* (4 %, SD = 5). Similarly, the *myo* and *scyllo* stereoisomers were also the most predominant forms of IP_5 but comprised between 83 % (Cambisol) and 100 % (Ferralsol and Vertisol) of total IP_5 (Table 5). Trace amounts of *scyllo*-IP_4 were also detected in three of the four soils. The ratio of total IP_6 to IP_5 differed across all soils (Fig. 2).

Table 5. Concentrations of identified inositol phosphates (IPs) in hypobromite oxidised 0.25 M NaOH + 0.05 M EDTA soil extracts (Ferralsol, Vertisol, Cambisol and Gleysol). Concentrations were calculated from solution ^{31}P NMR spectra using spectral-deconvolution fitting including an underlying broad signal. When no concentration is given, the IP compound was not detected in the respective soil extract. Chemical-shift positions are based on the NMR spectrum of the Cambisol extract (Fig. S8 in the Supplement). Peak positions varied by up to +0.018 ppm (Gleysol). Conformation equatorial (eq) and axial (ax) according to Turner et al. (2012).

Phosphorus compound	Chemical shift (δ ppm)	Concentrations (mg P per kg soil)			
		Ferralsol	Vertisol	Cambisol	Gleysol
myo-IP_6	4.97, 4.06, 3.70, 3.57	1.1	0.6	26.3	85.0
scyllo-IP_6	3.20	0.4	0.3	15.6	41.1
neo-IP_6 4-eq–2-ax	5.86, 3.75	–	–	1.4	8.8
neo-IP_6 2-eq–4-ax	4.36, 4.11	–	–	4.0	1.3
D-*chiro*-IP_6 2-eq–4-ax	5.66, 4.25, 3.83	–	–	9.4	8.6
myo-(1,2,4,5,6)-IP_5	4.42, 3.97, 3.72, 3.36, 3.25	–	–	7.0	4.1
myo-(1,3,4,5,6)-IP_5	4.12, 3.60, 3.23	–	–	2.8	1.3
scyllo-IP_5	3.81, 3.31, 3.05	0.7	0.5	10.8	6.1
neo-IP_5	4.64, 4.27, 4.01, 3.87, 3.13	–	–	3.3	2.1
chiro-IP_5	4.61, 3.39	–	–	0.9	–
scyllo-(1,2,3,4)-IP_4	4.12, 3.25	0.8	–	4.3	1.0
Total IP		3.0	1.4	85.9	159.3

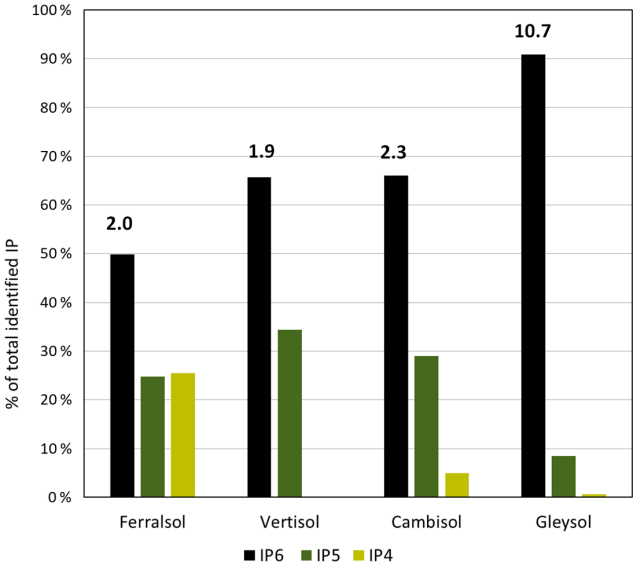


Figure 2. The proportion of total identifiable pools of inositol hexakisphosphates (IP_6), pentakisphosphates (IP_5) or tetrakisphosphates (IP_4) in the total pool of identifiable IP, as determined by solution ^{31}P NMR spectroscopy on four soil extracts (Ferralsol, Vertisol, Cambisol and Gleysol) following hypobromite oxidation. Values located above the IP_6 bar are the ratio of total identifiable IP_6 to total identifiable IP_5 in each soil sample.

If sharp peaks arising from IPs were identified in the NMR spectra on hypobromite oxidised extracts, a comparison was made with that of their corresponding untreated extracts. The sharp peaks of all stereoisomers of IP_6 were present in the untreated extracts. The five peaks of *myo*-(1,2,4,5,6)-IP_5 and the three peaks of *scyllo*-IP_5 were also identified. However, it was not possible to clearly identify other IP_5 compounds in untreated extracts due to overlapping signals. In the Gleysol, all three peaks of *scyllo*-IP_5 were detected, but only two of the possible five peaks could be clearly assigned to *myo*-(1,2,4,5,6)-IP_5. In the Ferralsol, both peaks of *scyllo*-IP_4 were present in the untreated extract, but only two of the three possible peaks could be assigned to *scyllo*-IP_5. In the Vertisol, no IP_5 was identified. Concentrations of IP in untreated extracts assessed by spectral-deconvolution fitting were generally double those measured in hypobromite oxidised extracts. Recoveries of added *myo*-IP_6 in the Gleysol and Cambisol following hypobromite oxidation were 47 % and 20 %, respectively.

3.4 Spin-echo analysis of selected P compounds

Due to the presence of sharp and broad signals in hypobromite oxidised soil extracts, the structural composition of phosphomonoesters was probed. A comparison of the NMR spectra at the lowest ($1^*\tau$) and highest ($80^*\tau$) pulse delays revealed a fast-decaying broad signal for all hypobromite oxidised soil extracts, which was particularly evident in the Gleysol (Fig. 3). Calculated T_2 times of all IP_6 stereoisomers were longer than those of the broad signal (Table 6). The T_2 times of *scyllo*-IP_6 (on average 175.8 ms, SD = 49.7) were generally the longest of all stereoisomers of IP_6. The T_2 time of the orthophosphate peak was the shortest, which was on average 11.5 ms (SD = 4.9).

The average ($n = 4$) T_2 time of the broad peak was significantly different than that of *scyllo*- and *myo*-IP_6 ($p < 0.05$). Significant differences in the T_2 times of *neo*- and D-*chiro*-

Table 6. Transversal relaxation times (T_2) of various P species in the orthophosphate and phosphomonoester regions as determined by solution ^{31}P nuclear magnetic resonance (NMR) spectroscopy and a Carr–Purcell–Meiboom–Gill (CPMG) pulse sequence on hypobromite oxidised soil extracts.

Phosphorus compound	T_2 (ms)			
	Ferralsol	Vertisol	Cambisol	Gleysol
myo-IP_6	163	140	139	121
scyllo-IP_6	250	155	154	144
neo-IP_6	–	–	203	102
D-chiro-IP_6	–	–	108	132
Orthophosphate	14	9	17	6
Broad peak	44	69	89	62

IP_6 were not tested, as these compounds were not detected in the Ferralsol and the Vertisol.

4 Discussion

4.1 Pools of phosphorus in untreated and hypobromite oxidised soil extracts

On average, 44 % of total P (as measured with XRF) was extracted by NaOH-EDTA, which is consistent with previous studies (Turner, 2008; Li et al., 2018; McLaren et al., 2019). The non-extractable pool of P is likely to comprise inorganic P as part of insoluble mineral phases but could also contain some organic P (McLaren et al., 2015a). Nevertheless, the NaOH-EDTA extraction technique is considered to be a measure of total organic P in soil, which can be subsequently characterised by solution ^{31}P NMR spectroscopy (Cade-Menun and Preston, 1996).

Hypobromite oxidation resulted in a decrease in the concentration of inorganic and organic P in NaOH-EDTA extracts across all soils. The decrease of organic P is consistent with previous studies (Turner and Richardson, 2004; Turner et al., 2012; Almeida et al., 2018). However, Almeida et al. (2018) reported an overall increase in the concentration of inorganic P following hypobromite oxidation, which the authors proposed to be caused by the degradation of organic P forms not resistant to hypobromite oxidation. A decrease in the concentration of organic P in NaOH-EDTA extracts following hypobromite oxidation was expected based on the oxidation of organic molecules containing P. The products of hypobromite oxidation are most probably carbon dioxide, simple organic acids from the oxidative cleavage of the phosphoesters and orthophosphate (Irving and Cosgrove, 1981; Sharma, 2013).

Overall, hypobromite oxidation of NaOH-EDTA soil extracts resulted in a considerable increase in the number of sharp peaks and a decrease in the broad underlying peak in the phosphomonoester region compared to untreated soil extracts. This was particularly the case for the Cambisol and the Gleysol, which had high concentrations of extractable organic P. Since the broad peak is thought to be closely associated with the SOM (Dougherty et al., 2007; Bünemann et al., 2008; McLaren et al., 2015b), its decrease in soil extracts following hypobromite oxidation is consistent with that observed for other organic compounds (Turner et al., 2012). Our results indicate that a majority of sharp peaks present in the phosphomonoester region of untreated soil extracts are stable to hypobromite oxidation and are therefore likely to be IPs.

Across all soils, 5 to 15 peaks in the phosphomonoester region were removed following hypobromite oxidation compared to those in untreated extracts, which are likely due to the oxidation of α- and β-glycerophosphate (Doolette et al., 2009; McLaren et al., 2015b), RNA mononucleotides (8 peaks; Vincent et al., 2013), glucose 6-phosphate, phosphocholine, glucose 1-phosphate, or phosphorylethanolamine (Cade-Menun, 2015).

4.2 Phosphorus assignments of sharp peaks in hypobromite oxidised extracts

The detection of *myo-, scyllo-*, *chiro* and *neo*-IP_6 in untreated and hypobromite oxidised soil extracts is consistent with previous studies using chromatography (Irving and Cosgrove, 1982; Almeida et al., 2018) and NMR (Turner and Richardson, 2004; Doolette et al., 2011a; Vincent et al., 2013; Jarosch et al., 2015; McLaren et al., 2015b). Turner et al. (2012) suggested that hypobromite oxidised extracts only contained *neo*-IP_6 in the 4-equatorial–2-axial conformation due to the absence of signals from the 2-equatorial–4-axial conformation. In the current study, both conformations could be identified in two of the four soil extracts, which is likely due to improved spectral resolution and sensitivity. The relative abundances of the four identified stereoisomers of IP_6 in soil extracts were similar to previous studies (Irving and Cosgrove, 1982; Turner et al., 2012).

Several studies have shown overlap of peaks relating to RNA mononucleotides and those of α- and β-glycerophosphate, which are the alkaline hydrolysis products of RNA and phospholipids, respectively. However, in the current study, several sharp peaks were present in hypobromite oxidised extracts which are in the chemical-shift range of RNA mononucleotides and α- and β-glycerophosphate. Whilst a peak at δ 4.36 ppm would be assigned to α-glycerophosphate based on spiking experiments in the untreated extracts of the Cambisol and the Gleysol, hypobromite oxidation revealed the presence of the 2-equatorial–4-axial C2,5 peak of *neo*-IP_6 at δ 4.37 ppm and also an unidentified peak at δ 4.36 ppm in the Cambisol. Therefore, the assignment and concentration of α-glycerophosphate may be unreliable in some soils of previous studies.

For the first time, we identified lower-order IP (IP_5 and IP_4) in soil extracts using solution ^{31}P NMR spectroscopy.

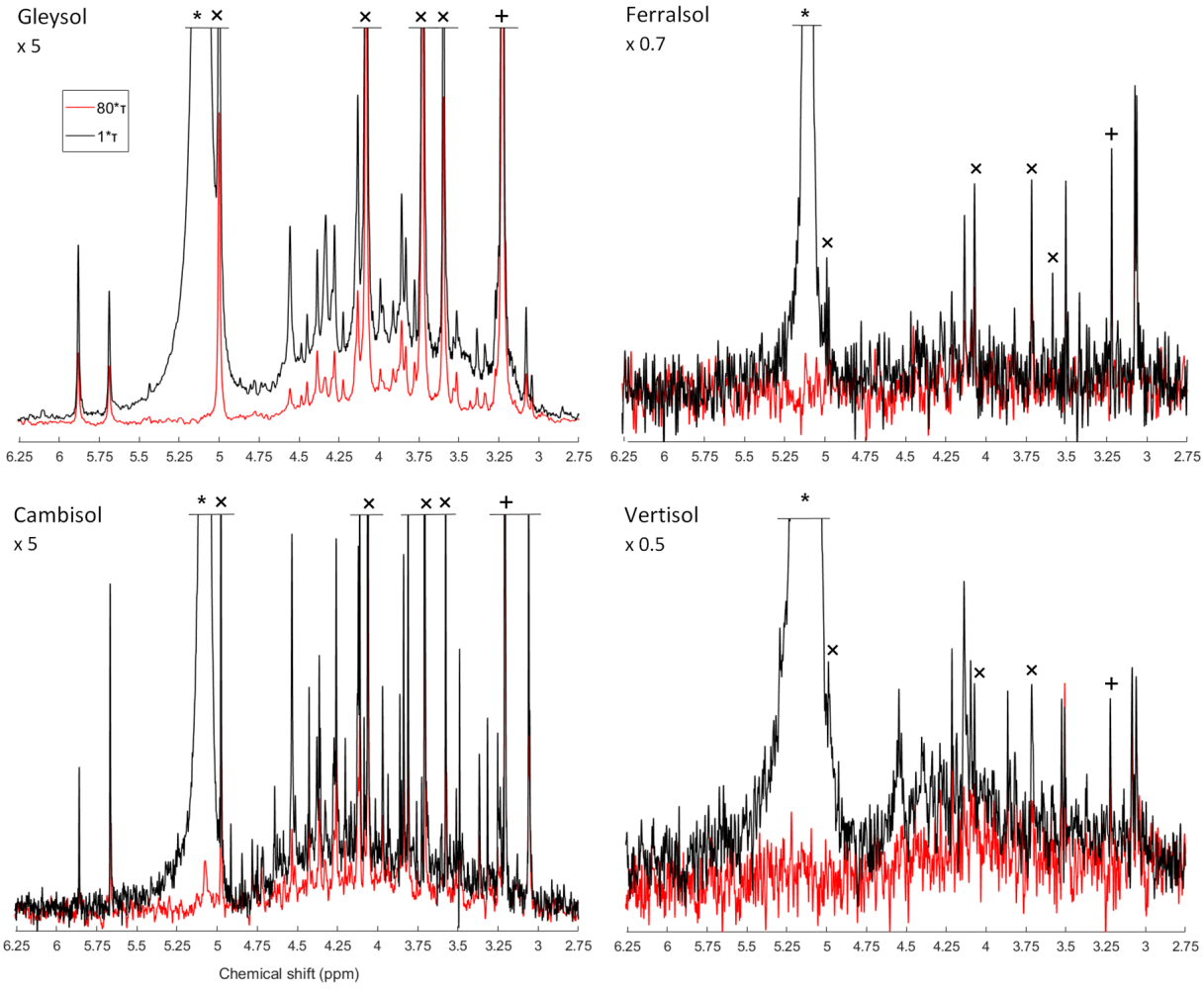


Figure 3. Solution ^{31}P NMR spectra of hypobromite oxidised soil extracts acquired with a CPMG pulse sequence with $1^{*}\tau$ (black) and $80^{*}\tau$ (red) spin-echo delays. The orthophosphate (*), *scyllo*-IP_6 (+) and *myo*-IP_6 peaks (×) are marked accordingly. Spectra were normalised to the maximum *scyllo*-IP_6 peak intensity in the $1^{*}\tau$ spectrum for each soil. The vertical axes were increased or decreased for better visualisation by an indicated factor.

Smith and Clark (1951) were the first to suggest the presence of IP_5 in soil extracts using anion-exchange chromatography, which was later confirmed (Anderson, 1955; Cosgrove, 1963; McKercher and Anderson, 1968b). Halstead and Anderson (1970) reported the presence of all four stereoisomers (*myo*, *scyllo*, *neo* and *chiro*) in the lower ester fractions (IP_2–IP_4) as well as in the higher ester fractions (IP_5, IP_6) isolated from soil, with the *myo* stereoisomer being the main form in all fractions. In the current study, all four stereoisomers of IP_5 could be detected in the hypobromite oxidised soil extracts, of which the *myo* and *scyllo* stereoisomers were the most abundant. The relative abundances of IP_5 stereoisomers are consistent with the findings of Irving and Cosgrove (1982) using gas-liquid chromatography on the combined $IP_6 + IP_5$ fraction. The detection of all four stereoisomers of IP_5 in NMR spectra provides direct spectroscopic evidence for their existence in soil extracts.

In addition to the four stereoisomers of IP_5, we were able to identify the presence of two isomers of *myo*-IP_5 in the Cambisol and Gleysol, i.e. *myo*-(1,2,4,5,6)-IP_5 and *myo*-(1,3,4,5,6)-IP_5. These two isomers have not yet been detected in soil extracts. A distinction of different *myo*-IP_5 isomers was not reported in earlier studies using chromatographic separation. In non-soil extracts, *myo*-(1,2,4,5,6)-IP_5 was detected by Doolette and Smernik (2016) in grapevine canes and *myo*-(1,3,4,5,6)-IP_5 as the thermal-decomposition product of a phytate standard (Doolette and Smernik, 2018). It is possible that an abiotic transformation of *myo*-IP_6 to *myo*-(1,3,4,5,6)-IP_5 occurs, which could then be adsorbed by soil constituents. Stephens and Irvine (1990) reported *myo*-(1,3,4,5,6)-IP_5 as an intermediate in the synthesis of IP_6 from *myo*-IP in the cellular slime mould *Dictyostelium*. Therefore, *myo*-(1,3,4,5,6)-IP_5 could have been biologically added to the soil. Furthermore, *myo*-(1,3,4,5,6)-IP_5 was present in dif-

ferent animal feeds and manures (Sun and Jaisi, 2018). Sun et al. (2017) reported *myo*-(1,3,4,5,6)-IP_5 and *myo*-(1,2,4,5,6)-IP_5 as intermediates in the minor pathways and major pathways, respectively, of *Aspergillus niger* phytase and acid phosphatase (potato) phytate degradation. The presence of *myo*-(1,2,3,4,6)-IP_5 could not be confirmed as NMR analyses on the compound itself exhibited a broad NMR signal (Fig. S7 in the Supplement). This is because in solutions with a pH of 9.5 or above, the 1-axial–5-equatorial and 5-axial–1-equatorial forms of *myo*-(1,2,3,4,6)-IP_5 are in a dynamic equilibrium, which can cause broadening (Volkmann et al., 2002). According to Turner and Richardson (2004) and Chung et al. (1999), the two identified *scyllo*-IP_4 peaks (signal pattern 2 : 2) can be attributed to the *scyllo*-(1,2,3,4)-IP_4 isomer. However, the two peaks of *scyllo*-IP_4 overlapped in the Cambisol and Gleysol with the peak at the furthest upfield chemical shift of *myo*-(1,2,4,5,6)-IP_5 at δ 3.25 ppm and with the peak at the furthest downfield chemical shift of *myo*-(1,3,4,5,6)-IP_5 at δ 4.12 ppm. Turner and Richardson (2004) reported NMR signals for two other *scyllo*-IP_4 isomers, which could not be tested for in this study due to the lack of available standards.

Whilst on average 48 % of the sharp peaks in the phosphomonoester region of soil extracts following hypobromite oxidation could be attributed to IP_6, IP_5 and *scyllo*-IP_4, the identities of many sharp peaks remain unknown. An unidentified peak at δ 4.33 ppm is present in all soil samples except in the Ferralsol, with concentrations of up to 10 mg P per kg soil (Cambisol). Other unidentified peaks at δ 3.49, 3.86, 4.20 and 3.91 ppm were detected in all soils, with concentrations ranging from 1 to 2 mg P per kg soil. Interestingly, two peaks upfield of *scyllo*-IP_6 became more prominent (at δ 3.08 and 3.05 ppm) following hypobromite oxidation, which was particularly the case in the Vertisol soil. The diversity of organic P species in the Vertisol soil appears to be much greater than previously reported (McLaren et al., 2014). We hypothesise that many of these unidentified peaks arise from other isomers of *myo*- and *scyllo*-IP_5, based on the higher abundance of their IP_6 counterparts.

The ratio of IP_6 to lower-order IP varied across soils, which ranged in decreasing order: Gleysol $\gg$ Cambisol > Vertisol > Ferralsol. McKercher and Anderson (1968a) found a higher ratio of IP_6 to IP_5 in some Scottish soils (ratio 1.8 to 4.6) compared to some Canadian soils (0.9 to 2.4). The authors attributed this difference to the greater stabilisation of IP_6 relative to lower esters in the Scottish soils, possibly due to climatic reasons or effects of different soil properties. In a subsequent study, McKercher and Anderson (1968b) observed increased IP contents with increasing total organic P content. Studies of organic P speciation along chronosequences found that *myo*-IP_6 concentrations declined in older soils (McDowell et al., 2007; Turner et al., 2007a, 2014). Similarly, Baker (1976) found that the $IP_6 + IP_5$ concentrations in the Franz Josef chronosequence increased until reaching 1000 years, followed by a rapid decline. In our soil samples, the highest IP_6-to-IP_5 ratio was found in the soil with the highest SOM content, suggesting a possible stabilisation of IP_6 due to association with SOM (Borie et al., 1989; Makarov et al., 1997). In contrast, the Ferralsol sample containing high amounts of Fe and Al showed the smallest IP_6-to-IP_5 ratio, even though IP_6 is known to strongly adsorb to sesquioxides (Anderson and Arlidge, 1962; Anderson et al., 1974). However, the production, input and mineralisation rates of IP_6 and IP_5 are not known for our soil samples. Further research is needed to understand the mechanisms controlling the flux of lower-order IP in soil.

In the Ferralsol and the Cambisol, there was an overall decrease in the concentration of IP_6 and IP_5 following hypobromite oxidation compared to the untreated extracts. Since the main cause of resistance of IP to hypobromite oxidation is that of steric hindrance, which generally decreases with a decreasing phosphorylation state and conformation of the phosphate groups (axial vs. equatorial), we assume that low recoveries of added *myo*-IP_6 are due to losses of precipitated P_{org} compounds during the precipitation and dissolution steps. This is supported by the decrease in the concentration of orthophosphate following hypobromite oxidation compared to untreated extracts. Therefore, quantities of IP as reported in the current study should be considered as conservative.

4.3 Structural composition of phosphomonoesters in hypobromite oxidised soil extracts

The NMR spectra on hypobromite oxidised soil extracts revealed the presence of sharp and broad signals in the phosphomonoester region. Transverse relaxation experiments revealed a rapid decay of the broad signal compared to the sharp peaks of IP_6, which support the hypothesis that the compounds causing the broad signal arise from P compounds other than IP. These findings are consistent with those of McLaren et al. (2019), who probed the structural composition of phosphomonoesters in untreated soil extracts. Overall, measured T_2 times in the current study on hypobromite oxidised extracts were markedly longer compared to those of untreated extracts reported in McLaren et al. (2019). This could be due to removal of other organic compounds by hypobromite oxidation in the matrix and therefore a decrease in the viscosity of the sample. This would result in an overall faster tumbling of the molecules and hence an increased T_2 relaxation time. As reported by McLaren et al. (2019), calculations of the broad signal's linewidth based on the T_2 times were considerably lower compared to those of the standard deconvolution fitting (SDF). When applying the same calculations to our samples, the linewidth of the broad signal at half height is on average 5.2 Hz based on the T_2 times. In contrast, the linewidths acquired from the SDF average 256.1 Hz. McLaren et al. (2019) suggested that the broad signal is itself comprised of more than one compound. Our re-

sults are consistent with this view, and therefore it is likely that the main cause of the broad signal is a diversity of P molecules of differing chemical environments within this region, rather than the slow tumbling of just one macromolecule. Nebbioso and Piccolo (2011) reported that high-molecular-weight material of organic matter in soil is an association of smaller organic molecules. We suggest that these associations would still cause a broad signal in the phosphomonoester region of soil extracts and could be a reason that some organic molecules containing P are protected from hypobromite oxidation.

Since a portion of the broad signal is resistant to hypobromite oxidation, this suggests the organic P is complex and in the form of polymeric structures. The chemical resistance of the broad signal to hypobromite oxidation may also indicate a high stability in soil (Jarosch et al., 2015). Annaheim et al. (2015) found that concentrations of the broad signal remained unchanged across three different organic fertiliser strategies after 62 years of cropping. In contrast, the organic P compounds annually added with the fertilisers were completely transformed or lost in the slightly acidic topsoil of the field trial. The large proportion of the broad signal in the total organic P pool demonstrates its importance in the soil P cycle.

Unexpectedly, the transverse relaxation times of orthophosphate were shorter than those of the broad signal. This was similarly the case in an untreated NaOH-EDTA extract of a forest soil with the same origin as the Cambisol as reported in McLaren et al. (2019). The authors hypothesised that this might be due to the sample matrix (i.e. high concentration of metals and organic matter). Whilst these factors are likely to affect T_2 times, they do not appear to be the main cause as the hypobromite oxidised extracts in the current study contained low concentrations of organic matter and metals as a consequence of the isolation procedure. The fast decay of orthophosphate was found across all four soil extracts with a diverse array of organic P concentrations and compositions of organic P in the phosphomonoester region. Therefore, another possible explanation could be a matrix effect or an association with large organic P compounds causing the broad signal (McLaren et al., 2019). It is known that dynamic intramolecular processes such as ring inversion and intermolecular processes such as binding of small-molecule ligands to macromolecules can cause a broadening or a doubling of resonances (Claridge, 2016). When the smaller molecule is bound to the larger molecule, it experiences slower tumbling in the solution and hence a shorter T_2 time. It is possible that a chemical exchange of the orthophosphate with a compound in the matrix or an organic P molecule could result in the short T_2 time of the orthophosphate peak. We carried out a T_2 experiment on a pure solution of monopotassium phosphate (described in the Supplement), in which the matrix effects should be considerably reduced compared to the soil extracts. We found that the T_2 time of orthophosphate (203 ms) in the pure solution was considerably longer than that reported in soil extracts following hypobromite oxidation.

5 Conclusions

Inositol phosphates are an important pool of organic P in soil, but information on the mechanisms controlling their flux in soil remains limited due in part to an inability to detect them using solution ^{31}P NMR spectroscopy. For the first time, we identified six different lower-order IPs in the solution ^{31}P NMR spectra on soil extracts. Solution ^{31}P NMR spectra on hypobromite oxidised extracts revealed the presence of up to 70 sharp peaks, of which about 50 % could be identified. Our results indicate that a majority of the sharp peaks in solution ^{31}P NMR soil spectra were resistant to hypobromite oxidation and therefore suggest the presence of diverse IPs. Our study highlights the great diversity and abundance of IPs in soils and therefore their importance in terrestrial P cycles. Further research on the mechanisms and processes involved in the cycling of this wide variety of IPs in soil will have implications for our understanding of organic P turnover as well as plant availability and possibly help improve fertiliser strategies in agricultural systems.

Furthermore, we provide new insight into the large pool of phosphomonoesters represented by the broad signal, of which a considerable portion was resistant to hypobromite oxidation. Further research is needed to understand the chemical composition of the broad signal and the mechanisms controlling its flux in terrestrial ecosystems.

Data availability. All data presented in this study and the Supplement are also available by request from the corresponding author.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/bg-17-5079-2020-supplement.

Author contributions. The experimental design was planned by JER, TIM, DZ, RV and EF. The experiments were carried out by JER under the supervision of TIM, DZ and RV. RV provided the MATLAB code for the standard deconvolution fitting of the NMR spectra. The data were processed, analysed and interpreted by JER with support from TIM, DZ and RV. JER prepared the manuscript with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors are grateful to Laurie Paule Schönholzer, Federica Tamburini, Björn Studer, Monika Macsai and Charles Brearley for technical support.

Furthermore, the authors thank Astrid Oberson, David Lester, Chiara Pistocchi and Gregor Meyer for providing soil samples. This study would not have been possible without the IP standards originating from the late Dennis Cosgrove collection and Max Tate collection, which we highly appreciate.

Financial support. This research has been supported by the Swiss National Science Foundation (grant no. 200021_169256).

Review statement. This paper was edited by Sébastien Fontaine and reviewed by three anonymous referees.

References

Almeida, D. S., Menezes-Blackburn, D., Turner, B. L., Wearing, C., Haygarth, P. M., and Rosolem, C. A.: Urochloa ruziziensis cover crop increases the cycling of soil inositol phosphates, Biol. Fert. Soils, 54, 935–947, https://doi.org/10.1007/s00374-018-1316-3, 2018.

Anderson, G.: Paper chromatography of inositol phosphates, Nature, 175, 863–864, https://doi.org/10.1038/175863b0, 1955.

Anderson, G. and Arlidge, E. Z.: The adsorption of inositol phosphates and glycerophosphate by soil clays, clay minerals, and hydrated sesquioxides in acid media, J. Soil Sci., 13, 216–224, https://doi.org/10.1111/j.1365-2389.1962.tb00699.x, 1962.

Anderson, G. and Malcolm, R. E.: The nature of alkali-soluble soil organic phosphates., J. Soil Sci., 25, 282–297, https://doi.org/10.1111/j.1365-2389.1974.tb01124.x, 1974.

Anderson, G., Williams, E. G., and Moir, J. O.: A comparison of the sorption of inorganic orthophosphate and inositol hexaphosphate by six acid soils, J. Soil Sci., 25, 51–62, https://doi.org/10.1111/j.1365-2389.1974.tb01102.x, 1974.

Angyal, S. J.: Cyclitols, in: Comprehensive Biochemistry, edited by: Florkin, M. and Stotz, E. H., Elsevier, chap. VIII, 53 Vanderbilt Avenue, New York 17, NY, 297–303, 1963.

Annaheim, K. E., Doolette, A. L., Smernik, R. J., Mayer, J., Oberson, A., Frossard, E., and Bünemann, E. K.: Long-term addition of organic fertilizers has little effect on soil organic phosphorus as characterized by ^{31}P NMR spectroscopy and enzyme additions, Geoderma, 257–258, 67–77, https://doi.org/10.1016/j.geoderma.2015.01.014, 2015.

Baker, R. T.: Changes in the chemical nature of soil organic phosphate during pedogenesis., J. Soil Sci., 27, 504–512, https://doi.org/10.1111/j.1365-2389.1976.tb02020.x, 1976.

Borie, F., Zunino, H., and Martínez, L.: Macromolecule associations and inositol phosphates in some chilean volcanic soils of temperate regions, Commun. Soil Sci. Plan., 20, 1881–1894, https://doi.org/10.1080/00103628909368190, 1989.

Bühler, S., Oberson, A., Sinaj, S., Friesen, D. K., and Frossard, E.: Isotope methods for assessing plant available phosphorus in acid tropical soils, Eur. J. Soil Sci., 54, 605–616, https://doi.org/10.1046/j.1365-2389.2003.00542.x, 2003.

Bünemann, E. K., Smernik, R. J., Marschner, P., and McNeill, A. M.: Microbial synthesis of organic and condensed forms of phosphorus in acid and calcareous soils, Soil Biol. Biochem., 40, 932–946, https://doi.org/10.1016/j.soilbio.2007.11.012, 2008.

Bünemann, E. K., Augstburger, S., and Frossard, E.: Dominance of either physicochemical or biological phosphorus cycling processes in temperate forest soils of contrasting phosphate availability, Soil Biol. Biochem., 101, 85–95, https://doi.org/10.1016/j.soilbio.2016.07.005, 2016.

Cade-Menun, B. J.: Improved peak identification in ^{31}P-NMR spectra of environmental samples with a standardized method and peak library, Geoderma, 257–258, 102–114, https://doi.org/10.1016/j.geoderma.2014.12.016, 2015.

Cade-Menun, B. and 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 Sci. Soc. Am. J., 78, 19–37, https://doi.org/10.2136/sssaj2013.05.0187dgs, 2014.

Cade-Menun, B. J. and Preston, C. M.: A comparison of soil extraction procedures for ^{31}P NMR spectroscopy, Soil Sci., 161, 770–785, 1996.

Cade-Menun, B. J., Liu, C. W., Nunlist, R., and McColl, J. G.: Soil and litter phosphorus-31 nuclear magnetic resonance spectroscopy, J. Environ. Qual., 31, 457–465, https://doi.org/10.2134/jeq2002.4570, 2002.

Celi, L. and Barberis, E.: Abiotic reactions of inositol phosphates in soil, in: Inositol phosphates: linking agriculture and the environment, edited by: Turner, B. L., Richardson, A. E., and Mullaney, E. J., CABI, Wallingford, 207–220, 2007.

Chung, S.-K., Kwon, Y.-U., Chang, Y.-T., Sohn, K.-H., Shin, J.-H., Park, K.-H., Hong, B.-J., and Chung, I.-H.: Synthesis of all possible regioisomers of *scyllo*-Inositol phosphate, Bioorgan. Med. Chem., 7, 2577–2589, https://doi.org/10.1016/S0968-0896(99)00183-2, 1999.

Claridge, T. D. W.: Introducing High-Resolution NMR, in: High-Resolution NMR techniques in organic chemistry, 3rd edn., edited by: Claridge, T. D. W., Elsevier, Boston, chap. 2, 11–59, 2016.

Cosgrove, D.: The chemical nature of soil organic phosphorus. I. Inositol phosphates, Soil Res., 1, 203–214, https://doi.org/10.1071/SR9630203, 1963.

Cosgrove, D. J. and Irving, G. C. J.: Inositol phosphates: their chemistry, biochemistry and physiology, Studies in organic chemistr, Elsevier, Amsterdam, 1980.

Doolette, A. L. and Smernik, R. J.: Phosphorus speciation of dormant grapevine (Vitis vinifera L.) canes in the Barossa Valley, South Australia, Aust. J. Grape Wine R., 22, 462–468, https://doi.org/10.1111/ajgw.12234, 2016.

Doolette, A. L. and Smernik, R. J.: Facile decomposition of phytate in the solid-state: kinetics and decomposition pathways, Phosphorus Sulfur, 193, 192–199, https://doi.org/10.1080/10426507.2017.1416614, 2018.

Doolette, A. L., Smernik, R. J., and Dougherty, W. J.: Spiking improved solution phosphorus-31 nuclear magnetic resonance identification of soil phosphorus compounds, Soil Sci. Soc. Am. J., 73, 919–927, https://doi.org/10.2136/sssaj2008.0192, 2009.

Doolette, A. L., Smernik, R. J., and Dougherty, W. J.: A quantitative assessment of phosphorus forms in some Australian soils, Soil Res., 49, 152–165, https://doi.org/10.1071/SR10092, 2011a.

Doolette, A. L., Smernik, R. J., and Dougherty, W. J.: Overestimation of the importance of phytate in NaOH–EDTA soil extracts as assessed by 31P NMR analyses, Org. Geochem., 42, 955–964, https://doi.org/10.1016/j.orggeochem.2011.04.004, 2011b.

Dougherty, W. J., Smernik, R. J., and Chittleborough, D. J.: Application of spin counting to the solid-state ^{31}P NMR analysis of pasture soils with varying phosphorus content, Soil Sci. Soc. Am. J., 69, 2058–2070, 10.2136/sssaj2005.0017, 2005.

Dougherty, W. J., Smernik, R. J., Bünemann, E. K., and Chittleborough, D. J.: On the use of hydrofluoric acid pretreatment of soils for phosphorus-31 nuclear magnetic resonance analyses, Soil Sci. Soc. Am. J., 71, 1111–1118, https://doi.org/10.2136/sssaj2006.0300, 2007.

Dyer, W. J. and Wrenshall, C. L.: Organic phosphorus in soils: III. The decomposition of some organic phosphorus compounds in soil cultures, Soil Sci., 51, 323–329, 1941.

FAO and Group, I. W.: World reference base for soil resources 2014, World soil resources reports, Food and Agriculture Organization of the United Nations FAO, Rome, 2014.

Fatiadi, A. J.: Bromine oxidation of inositols for preparation of inosose phenylhydrazones and phenylosazones, Carbohyd. Res., 8, 135–147, https://doi.org/10.1016/S0008-6215(00)80149-4, 1968.

Golovan, S. P., Meidinger, R. G., Ajakaiye, A., Cottrill, M., Wiederkehr, M. Z., Barney, D. J., Plante, C., Pollard, J. W., Fan, M. Z., Hayes, M. A., Laursen, J., Hjorth, J. P., Hacker, R. R., Phillips, J. P., and Forsberg, C. W.: Pigs expressing salivary phytase produce low-phosphorus manure, Nat. Biotechnol., 19, 741–745, https://doi.org/10.1038/90788, 2001.

Goring, C. A. I. and Bartholomew, W. V.: Microbial products and soil organic matter: III. Adsorption of carbohydrate phosphates by clays, Soil Sci. Soc. Am. J., 15, 189–194, https://doi.org/10.2136/sssaj1951.036159950015000C0043x, 1951.

Halstead, R. L. and Anderson, G.: Chromatographic fractionation of organic phosphates from alkali, acid, and aqueous acetylacetone extracts of soils, Can. J. Soil Sci., 50, 111–119, https://doi.org/10.4141/cjss70-018, 1970.

Hochberg, Y. and Tamhane, A. C.: Multiple comparison procedures, Wiley series in probability and mathematical statistics. Applied probability and statistics, Wiley, New York, 1987.

Irvine, R. F. and Schell, M. J.: Back in the water: the return of the inositol phosphates, Nat. Rev. Mol. Cell Bio., 2, 327, https://doi.org/10.1038/35073015, 2001.

Irving, G. C. J. and Cosgrove, D. J.: The use of hypobromite oxidation to evaluate two current methods for the estimation of inositol polyphosphates in alkaline extracts of soils, Commun. Soil Sci. Plan., 12, 495–509, https://doi.org/10.1080/00103628109367169, 1981.

Irving, G. C. J. and Cosgrove, D. J.: The use of gasliquid chromatography to determine the proportions of inositol isomers present as pentakis and hexakisphosphates in alkaline extracts of soils, Commun. Soil Sci. Plan., 13, 957–967, https://doi.org/10.1080/00103628209367324, 1982.

Jarosch, K. A., Doolette, A. L., Smernik, R. J., Tamburini, F., Frossard, E., and Bünemann, E. K.: Characterisation of soil organic phosphorus in NaOH-EDTA extracts: a comparison of ^{31}P NMR spectroscopy and enzyme addition assays, Soil Biol. Biochem., 91, 298–309, https://doi.org/10.1016/j.soilbio.2015.09.010, 2015.

Lang, F., Krüger, J., Amelung, W., Willbold, S., Frossard, E., Bünemann, E. K., Bauhus, J., Nitschke, R., Kandeler, E., Marhan, S., Schulz, S., Bergkemper, F., Schloter, M., Luster, J., Guggisberg, F., Kaiser, K., Mikutta, R., Guggenberger, G., Polle, A., Pena, R., Prietzel, J., Rodionov, A., Talkner, U., Meesenburg, H., von Wilpert, K., Hölscher, A., Dietrich, H. P., and Chmara, I.: Soil phosphorus supply controls P nutrition strategies of beech forest ecosystems in Central Europe, Biogeochemistry, 136, 5–29, https://doi.org/10.1007/s10533-017-0375-0, 2017.

L'Annunziata, M. F.: The origin and transformations of the soil inositol phosphate isomers, Soil Sci. Soc. Am. J., 39, 377–379, https://doi.org/10.2136/sssaj1975.03615995003900020041x, 1975.

Leytem, A. B. and Maguire, R. O.: Environmental implications of inositol phosphates in animal manures, in: Inositol phosphates: linking agriculture and the environment, edited by: Turner, B. L., Richardson, A. E., and Mullaney, E. J., CABI, Wallingford, 150–168, 2007.

Leytem, A. B., Turner, B. L., and Thacker, P. A.: Phosphorus composition of manure from swine fed low-phytate grains, J. Environ. Qual., 33, 2380–2383, https://doi.org/10.2134/jeq2004.2380, 2004.

Li, M., Cozzolino, V., Mazzei, P., Drosos, M., Monda, H., Hu, Z., and Piccolo, A.: Effects of microbial bioeffectors and P amendements on P forms in a maize cropped soil as evaluated by ^{31}P–NMR spectroscopy, Plant Soil, 427, 87–104, https://doi.org/10.1007/s11104-017-3405-8, 2018.

Makarov, M. I., Malysheva, T. I., Haumaier, L., Alt, H. G., and Zech, W.: The forms of phosphorus in humic and fulvic acids of a toposequence of alpine soils in the northern Caucasus, Geoderma, 80, 61–73, https://doi.org/10.1016/S0016-7061(97)00049-9, 1997.

McDowell, R. W. and Stewart, I.: The phosphorus composition of contrasting soils in pastoral, native and forest management in Otago, New Zealand: Sequential extraction and ^{31}P NMR, Geoderma, 130, 176–189, https://doi.org/10.1016/j.geoderma.2005.01.020, 2006.

McDowell, R. W., Cade-Menun, B., and Stewart, I.: Organic phosphorus speciation and pedogenesis: analysis by solution ^{31}P nuclear magnetic resonance spectroscopy, Eur. J. Soil Sci., 58, 1348–1357, https://doi.org/10.1111/j.1365-2389.2007.00933.x, 2007.

McKercher, R. B. and Anderson, G.: Characterization of the inositol penta- and hexaphosphate fractions of a number of Canadian and Scottish soils, J. Soil Sci., 19, 302–310, https://doi.org/10.1111/j.1365-2389.1968.tb01542.x, 1968a.

McKercher, R. B. and Anderson, G.: Content of inositol penta- and hexaphosphates in some Canadian soils, J. Soil Sci., 19, 47–55, https://doi.org/10.1111/j.1365-2389.1968.tb01519.x, 1968b.

McKercher, R. B. and Anderson, G.: Organic phosphate sorption by neutral and basic soils, Commun. Soil Sci. Plan., 20, 723–732, https://doi.org/10.1080/00103628909368112, 1989.

McLaren, T. I., Smernik, R. J., Guppy, C. N., Bell, M. J., and Tighe, M. K.: The organic P composition of vertisols as determined by ^{31}P NMR spectroscopy, Soil Sci. Soc. Am. J., 78, 1893–1902, https://doi.org/10.2136/sssaj2014.04.0139, 2014.

McLaren, T. I., Simpson, R. J., McLaughlin, M. J., Smernik, R. J., McBeath, T. M., Guppy, C. N., and Richardson, A. E.: An assessment of various measures of soil phosphorus and the net accumulation of phosphorus in fertilized soils under pasture, J. Plant Nutr. Soil Sci., 178, 543–554, https://doi.org/10.1002/jpln.201400657, 2015a.

McLaren, T. I., Smernik, R. J., McLaughlin, M. J., McBeath, T. M., Kirby, J. K., Simpson, R. J., Guppy, C. N., Doolette, A. L., and Richardson, A. E.: Complex forms of soil organic phosphorus–A major component of soil phosphorus, Environ. Sci. Technol., 49, 13238–13245, https://doi.org/10.1021/acs.est.5b02948, 2015b.

McLaren, T. I., Verel, R., and Frossard, E.: The structural composition of soil phosphomonoesters as determined by solution ^{31}P NMR spectroscopy and transverse relaxation (T_2) experiments, Geoderma, 345, 31–37, https://doi.org/10.1016/j.geoderma.2019.03.015, 2019.

Meiboom, S. and Gill, D.: Modified spinecho method for measuring nuclear relaxation times, Rev. Sci. Instrum., 29, 688–691, https://doi.org/10.1063/1.1716296, 1958.

Milliken, G. A. and Johnson, D. E.: Analysis of messy data. Volume 1: designed experiments, 2nd edn., CRC Press, Boca Raton, Florida, 2009.

Nebbioso, A. and Piccolo, A.: Basis of a Humeomics science: chemical fractionation and molecular characterization of humic biosuprastructures, Biomacromolecules, 12, 1187–1199, https://doi.org/10.1021/bm101488e, 2011.

Newman, R. H. and Tate, K. R.: Soil phosphorus characterisation by ^{31}P nuclear magnetic resonance, Commun. Soil Sci. Plan., 11, 835–842, https://doi.org/10.1080/00103628009367083, 1980.

Ognalaga, M., Frossard, E., and Thomas, F.: Glucose-1-phosphate and *myo*-inositol hexaphosphate adsorption mechanisms on goethite, Soil Sci. Soc. Am. J., 58, 332–337, https://doi.org/10.2136/sssaj1994.03615995005800020011x, 1994.

Ohno, T. and Zibilske, L. M.: Determination of low concentrations of phosphorus in soil extracts using malachite green, Soil Sci. Soc. Am. J., 55, 892–895, https://doi.org/10.2136/sssaj1991.03615995005500030046x, 1991.

Reusser, J. E., Verel, R., Frossard, E., and McLaren, T. I.: Quantitative measures of *myo*-IP_6 in soil using solution ^{31}P NMR spectroscopy and spectral deconvolution fitting including a broad signal, Environ. Sci. – Proc. Imp., 22, 1084–1094, https://doi.org/10.1039/C9EM00485H, 2020.

Riley, A. M., Trusselle, M., Kuad, P., Borkovec, M., Cho, J., Choi, J. H., Qian, X., Shears, S. B., Spiess, B., and Potter, B. V. L.: *scyllo* Inositol pentakisphosphate as an analogue of *myo* inositol 1,3,4,5,6: Chemical synthesis, physicochemistry and biological applications, ChemBioChem, 7, 1114–1122, https://doi.org/10.1002/cbic.200600037, 2006.

Sharma, V. K.: Oxidation of amino acids, peptides, and proteins: kinetics and mechanism, Wiley series of reactive intermediates in chemistry and biology, Wiley, Hoboken, 2013.

Smith, D. H. and Clark, F. E.: Anion-exchange chromatography of inositol phosphates from soil, Soil Sci., 72, 353–360, 1951.

Spain, A. V., Tibbett, M., Ridd, M., and McLaren, T. I.: Phosphorus dynamics in a tropical forest soil restored after strip mining, Plant Soil, 427, 105–123, https://doi.org/10.1007/s11104-018-3668-8, 2018.

Stephens, L. R. and Irvine, R. F.: Stepwise phosphorylation of *myo*-inositol leading to *myo*-inositol hexakisphosphate in Dictyostelium, Nature, 346, 580–583, https://doi.org/10.1038/346580a0, 1990.

Strickland, K. P.: The chemistry of phospholipids, in: Form and Function of Phospholipids, 2nd edn., edited by: Ansell, G. B., Hawthorne, J. N., and Dawson, R. M. C., Elsevier Scientific Publ. Company, Amsterdam, the Netherlands, 1973.

Sun, M. and Jaisi, D. P.: Distribution of inositol phosphates in animal feed grains and excreta: distinctions among isomers and phosphate oxygen isotope compositions, Plant Soil, 430, 291–305, https://doi.org/10.1007/s11104-018-3723-5, 2018.

Sun, M., Alikhani, J., Massoudieh, A., Greiner, R., and Jaisi, D. P.: Phytate degradation by different phosphohydrolase enzymes: contrasting kinetics, decay rates, pathways, and isotope effects, Soil Sci. Soc. Am. J., 81, 61–75, https://doi.org/10.2136/sssaj2016.07.0219, 2017.

Suzumura, M. and Kamatani, A.: Isolation and determination of inositol hexaphosphate in sediments from Tokyo Bay, Geochim. Cosmochim. Ac., 57, 2197–2202, https://doi.org/10.1016/0016-7037(93)90561-A, 1993.

Turner, B. L.: Inositol phosphates in soil: Amounts, forms and significance of the phosphorylated inositol stereoisomers., in: Inositol phosphates: Linking agriculture and the environment., edited by: Turner, B. L., Richardson, A. E., and Mullaney, E. J., CAB International, Wallingford, Oxfordshire, UK, 186–206, 2007.

Turner, B. L.: Soil organic phosphorus in tropical forests: an assessment of the NaOH–EDTA extraction procedure for quantitative analysis by solution ^{31}P NMR spectroscopy, Eur. J. Soil Sci., 59, 453–466, https://doi.org/10.1111/j.1365-2389.2007.00994.x, 2008.

Turner, B. L.: Isolation of inositol hexakisphosphate from soils by alkaline extraction and hypobromite oxidation, in: Inositol Phosphates: Methods and Protocols, edited by: Miller, G. J., Springer US, New York, NY, 39–46, 2020.

Turner, B. L. and Richardson, A. E.: Identification of *scyllo*-inositol phosphates in soil by solution phosphorus-31 nuclear magnetic resonance spectroscopy, Soil Sci. Soc. Am. J., 68, 802–808, https://doi.org/10.2136/sssaj2004.8020, 2004.

Turner, B. L., Papházy, M. J., Haygarth, P. M., and McKelvie, I. D.: Inositol phosphates in the environment, Philos. T. Roy. Soc. B, 357, 449–469, https://doi.org/10.1098/rstb.2001.0837, 2002.

Turner, B. L., Condron, L. M., Richardson, S. J., Peltzer, D. A., and Allison, V. J.: Soil organic phosphorus transformations during pedogenesis, Ecosystems, 10, 1166–1181, https://doi.org/10.1007/s10021-007-9086-z, 2007a.

Turner, B. L., Richardson, A. E., and Mullaney, E. J.: Inositol phosphates: linking agriculture and the environment, CABI, Wallingford, xi + 288 pp., 2007b.

Turner, B. L., Cheesman, A. W., Godage, H. Y., Riley, A. M., and Potter, B. V.: Determination of *neo*- and D-*chiro*-inositol hexakisphosphate in soils by solution ^{31}P NMR spectroscopy, Environ. Sci. Technol., 46, 4994–5002, https://doi.org/10.1021/es204446z, 2012.

Turner, B. L., Wells, A., and Condron, L. M.: Soil organic phosphorus transformations along a coastal dune chronosequence under New Zealand temperate rain forest, Biogeochemistry, 121, 595–611, https://doi.org/10.1007/s10533-014-0025-8, 2014.

Vestergren, J., Vincent, A. G., Jansson, M., Persson, P., Ilstedt, U., Gröbner, G., Giesler, R., and Schleucher, J.: High-resolution characterization of organic phosphorus in soil extracts using 2D ^{1}H–^{31}P NMR correlation spectroscopy, Environ. Sci. Technol., 46, 3950–3956, https://doi.org/10.1021/es204016h, 2012.

Vincent, A. G., Vestergren, J., Gröbner, G., Persson, P., Schleucher, J., and Giesler, R.: Soil organic phosphorus transformations

in a boreal forest chronosequence, Plant Soil, 367, 149–162, https://doi.org/10.1007/s11104-013-1731-z, 2013.

Vold, R. L., Waugh, J. S., Klein, M. P., and Phelps, D. E.: Measurement of spin relaxation in complex systems, J. Chem. Phys., 48, 3831–3832, https://doi.org/10.1063/1.1669699, 1968.

Volkmann, C. J., Chateauneuf, G. M., Pradhan, J., Bauman, A. T., Brown, R. E., and Murthy, P. P. N.: Conformational flexibility of inositol phosphates: influence of structural characteristics, Tetrahedron Lett., 43, 4853–4856, https://doi.org/10.1016/S0040-4039(02)00875-4, 2002.

Zhang, X., Li, Z., Yang, H., Liu, D., Cai, G., Li, G., Mo, J., Wang, D., Zhong, C., Wang, H., Sun, Y., Shi, J., Zheng, E., Meng, F., Zhang, M., He, X., Zhou, R., Zhang, J., Huang, M., Zhang, R., Li, N., Fan, M., Yang, J., and Wu, Z.: Novel transgenic pigs with enhanced growth and reduced environmental impact, eLife, 7, e34286, https://doi.org/10.7554/eLife.34286, 2018.