Fast and cost-efficient approaches for ^{17}O-isotopic labeling of carboxylic groups in biomolecules: from free amino acids to peptide chains

J. Špačková[1*], I. Goldberga[1], R. Yadav[1], G. Cazals[2], A. Lebrun[2], P. Verdié[2], T.-X. Métro[2*] and D. Laurencin[1*]

[1] ICGM, Université de Montpellier, CNRS, ENSCM, Montpellier, France

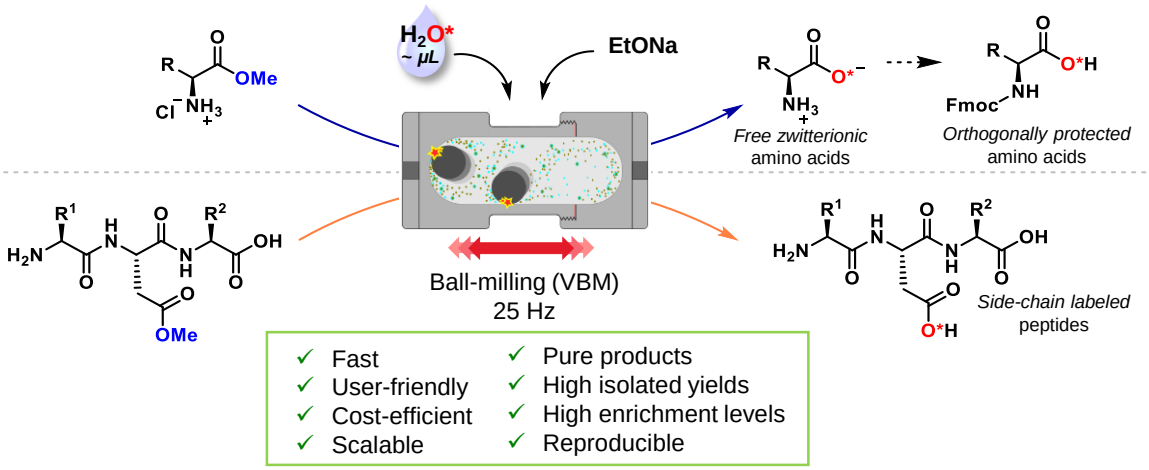
[2] IBMM, Université de Montpellier, CNRS, ENSCM, Montpellier, France

[*] jessica.spackova@umontpellier.fr, thomas-xavier.metro@umontpellier.fr, danielle.laurencin@umontpellier.fr

ABSTRACT:

^{17}O NMR spectroscopy is a powerful analytical technique, which enables to access unique information regarding the structure and reactivity of biomolecules, such as peptides and proteins. However, due to the exceedingly low natural abundance of ^{17}O (0.04 %), it is necessary to work with ^{17}O-enriched samples, which are not easily accessible because of the experimental constraints and high costs associated with the traditional enrichment procedures. Here, we present simple, fast and cost-efficient labeling strategies for ^{17}O-enrichment of amino acids and peptides. First, using mechanochemical saponification, a variety of unprotected amino acids were enriched within 30 min of milling under ambient conditions, consuming only microliter amounts of costly labeled water, and producing pure molecules with high enrichment levels (up to ~ 40 %), and in medium to high yields (~ 60 - 85 %) without the loss of their optical purity (*ee* > 99%). The labeling efficiency of the mechanochemical protocol was then compared to a re-optimised enrichment strategy based on acid-catalysed oxygen exchange. Subsequently, ^{17}O-enriched Fmoc/*t*Bu-protected amino acids were produced on a 1 g/day scale with high enrichment levels (~ 40 %), and in high synthetic yields (~ 75 - 85 %), by scaling up the mechanochemical enrichment followed by a Fmoc-protection step. Lastly, a direct site-selective ^{17}O-labeling of carboxylic functions in peptide side-chains was developed and applied to the RGD and GRGDS peptides, reaching up to 29% enrichment level. Producing highly enriched molecules enabled to record ^{17}O solid-state NMR spectra at 14.1 T in reasonable analytical times. Overall, this work represents an important step forward in providing easy access to highly ^{17}O-enriched peptides and proteins to be subsequently studied by high-resolution ^{17}O NMR spectroscopy.

Graphical abstract:



INTRODUCTION:

Peptides and proteins are present in every cell and tissue of the human body. They are involved in numerous essential biological processes serving, for example, as messengers (peptide hormones such as insulin), structure-forming agents (e.g. collagen, fibrin) or an immunity defence line in the form of antibodies (immunoglobulins). Their biological functions are determined by a specific sequence of amino acids (AA) interconnected *via* peptide bonds and by the 2D and 3D structure of the long protein chains.[1] Due to the irreplaceable roles of peptides and proteins, a lot of effort has been invested to explore their structure and to better understand their reactivity, as a crucial prerequisite for the development of new drugs[2] or biomimetic materials.[3]

A possible way to address some of the key questions regarding the structure and functional properties of peptides and proteins, is to prepare them in an isotopically labeled version, starting from amino acids labeled in stable isotopes, such as ^{2}H, ^{13}C, ^{15}N, or $^{17}O/^{18}O$.[4] Such labeled species have been involved in many studies where both qualitative and/or quantitative information were extracted from MS,[5-6] IR[7-8] or NMR analyses.[9-10] Producing peptides/proteins labeled in $^{17}O/^{18}O$ is particularly interesting, with oxygen atoms being frequently involved in hydrogen bonds stabilizing complex macromolecular protein structures, which are directly linked to their biological functions, or which mediate various protein-ligand interactions when located at binding sites. ^{18}O-labeled proteins have been used in high-resolution FTIR (Fourier transform infrared) studies providing side-specific structural information and/or to follow rapid structural changes and protein dynamics,[8, 11-13] or as MS standards for the purpose of quantitative proteomics.[6, 14] For such applications in IR and MS, it is required to work with almost quantitatively ^{18}O-labeled products (with ^{18}O-enrichment level > 90 %).

The great potential of ^{17}O-labeled products can be exploited through high-resolution solution and solid-state NMR studies.[15-16] Indeed, multinuclear NMR spectroscopy has proven to be a versatile analytical technique capable of providing information on the structure, function and dynamics of peptides and proteins by probing the local environment of a given nucleus, typically ^{1}H, ^{13}C or ^{15}N.[17-20] In this context, ^{17}O NMR spectroscopy appears as particularly advantageous as it can bring unique and very precise information regarding the structure and reactivity of proteins, due to the very high sensitivity of ^{17}O NMR parameters to the oxygen local environment, with NMR signals spanning over a ~ 1000 ppm wide chemical shift scale.[10, 21] For example, a transmembrane protein WALP23 containing ^{17}O-labeled alanine was

characterized by ^{17}O NMR (at 14.1 and 18.8 T) in its lyophilised form, and once incorporated in hydrated vesicles. In both cases, the C=O bond lengths at the labeled sites were estimated from these analyses.[22] In a more recent study of ion channel polypeptide Gramicidine A, high-resolution ^{17}O NMR experiments performed at 35.2 T enabled to detect selective hydrogen bonds between protons belonging to a "water wire" and isotopically labeled carbonyl oxygens of glycine and leucine lining the channel pore, and to observe for the first time a millisecond stability of the water wire orientation in the gramicidine A pore.[23] Furthermore, in terms of larger biomolecules, a recombinant protein, yeast ubiquitin (made of 76-residues), was prepared *via* expression with *E.coli* and analysed by ^{17}O NMR spectroscopy, including measurements at 35.2 T.[24] Both side-chain signals as well as carbonyl backbone signals were identified in the solution and solid-state ^{17}O NMR spectra, respectively, and parameters were extracted for each oxygen site.

Compared to the other NMR active nuclei, ^{17}O NMR studies are still rare due to the constraints associated with performing ^{17}O NMR analyses. This is mainly because of the very low natural abundance of the only NMR active oxygen isotope, oxygen-17 (0.04%), which requires working with ^{17}O-enriched biomolecules in order to obtain high-resolution data with sufficient sensitivity. In contrast with the applications of ^{18}O-labeled products, the ^{17}O-enrichment levels of the amino acids of interest within ^{17}O-labeled peptides are typically in the 15 - 60% range, which is sufficient for performing 1D and 2D ^{17}O NMR experiments at high magnetic fields. For most peptides involved in high-resolution NMR analyses, the ^{17}O-label was introduced in the sequence during a solid-phase peptide synthesis (SPPS) using ^{17}O-labeled fluorenylmethyloxycarbonyl/*tert*-butyl (Fmoc/*t*Bu) orthogonally protected amino acids as building blocks. This approach enables to label selectively only the chosen amino acid positions. In contrast, in the recently reported study of ubiquitin, the recombinant protein expression method employed (using free, unprotected amino acids, labeled in ^{17}O), resulted in labeling of all the glycine, phenylalanine or tyrosine units within the protein.[24]

With only two ^{17}O-enriched *free* amino acids currently available on the market (L-glutamic acid and L-tyrosine enriched on the phenolic group, to the best of our knowledge),[25] the laboratory synthesis of ^{17}O-labeled *free* and/or Fmoc/*t*Bu-*protected* amino acids is inevitable for nearly all studies of ^{17}O-enriched peptides and proteins. Nevertheless, their preparation is not straightforward due to the lack of user-friendly and cost-efficient ^{17}O-labeling protocols, which consequently hinders the widespread use of ^{17}O NMR spectroscopy in biomolecular applications. Indeed, the majority of oxygen labeling schemes available in the literature are developed and applied for ^{18}O-enrichment using large excess of ^{18}O-enriched

water. However, due to the considerably higher cost of ^{17}O-enriched water (1 mL of 90% ^{17}O-enriched H_2O* costs 1800 - 2900 € compared to 60 - 150 € for 1 mL of > 97% ^{18}O-enriched H_2O*), application of these protocols for ^{17}O-enrichment would be prohibitively expensive. In a typical labeling procedure, the ^{18}O-enrichment of *free* amino acids is based on acid-catalysed oxygen exchange between the amino acid and isotopically enriched water (introduced in excess, typically ~ 30 – 100 equivalents) leading to highly enriched molecules (with enrichment level above 90 %, using 95 - 98% $H_2{}^{18}O$).[26-30] In addition to using a large excess of enriched water, this labeling scheme suffers from other important drawbacks, such as very long reaction times (up to several days) at elevated temperatures (> 60 ºC), and the necessity to use a strong acid (such as HCl) for catalysis, in order to accelerate the conversion and reach equilibrium. In most cases, toxic and corrosive HCl gas is bubbled through the reaction medium for this purpose, which adds experimental constraints regarding the safety of the labeling process. Furthermore, most of the time, the *free* amino acids thus enriched are isolated under their hydrochloride salt form, and an additional anion-exchange procedure is then required to provide the pure zwitterionic form of the labeled product.[24] Due to the high cost of the ^{17}O-enriched H_2O*, and the time and experimental constraints of these protocols, it is desirable to develop more cost-efficient and user-friendly procedures for ^{17}O-enrichment of *free* amino acids.

Regarding the synthesis of enriched Fmoc-*protected* amino acids (Fmoc-AA-OH), acid-catalysed oxygen isotope exchange schemes like the one described above have been typically employed, providing the labeled species in one-step starting from non-labeled Fmoc-AA-OH.[13, 31] However, these protocols are only applicable to residues that do not require the presence of acid-labile *t*Bu-based side-chain protection groups for subsequent SPPS, thereby limiting the scope of accessible ^{17}O/^{18}O-labeled protected amino acids. More problematically, due to the large excess of enriched water employed, these procedures have only been used for ^{18}O-labeling, and even then the enrichment medium had to be recovered for the procedure to be more cost-efficient. Concerning the ^{17}O/^{18}O-enriched Fmoc/*t*Bu *orthogonally protected* amino acids, their preparation is more challenging, due to the simultaneous presence of the Fmoc-group protecting the α-NH_2 functions that is labile under basic conditions, and the side-chain protection group (*t*Bu) that is labile under strongly acidic conditions. The development of alternative labeling schemes working under mild reaction conditions has been therefore a focus of recent research,[32-34] yet still with room for improvement, as will be discussed later.

Overall, for both *free* and Fmoc-*protected* amino acids, the traditional oxygen labeling procedures tend to be expensive, long and not broadly applicable due to the harsh conditions frequently employed. More problematically, in many of these protocols, the synthetic procedure

is not properly described missing some important details, such as the number of water equivalents, the work-up conditions, isolated yield, enrichment level, or the purity of the isolated enriched compounds, making it difficult for other researchers to use or adapt the labeling schemes for their needs. With ^{17}O-enriched H_2O* as the most practical but at the same time very expensive source of ^{17}O isotope, the development of cost-efficient, reliable and broadly applicable labeling schemes is therefore of outmost importance.

In order to facilitate the access to ^{17}O-labeled amino acids and peptides, we recently turned our attention to mechanochemistry - a rapidly developing green synthetic methodology with increasing numbers of applications in organic synthesis.[35-36] The mechanochemical reactions are performed in reaction vessels (jars) containing reagents and milling beads vigorously agitated by a ball-mill, without the need for a bulk solvent. During ball-milling, the mechanical energy is transferred from the milling beads to the reagents by impact and shear forces, which most often result in a highly efficient mixing of the reagents, a decrease of solid particles' size, and the creation of fresh surfaces for subsequent reaction.[37] The lack of solvent presence has a beneficial impact not only on decreasing the solvent waste generated, but also on the reaction time which can be significantly reduced due to the high concentration of the reaction medium.[38] In some cases, the beneficial effect of adding a small amount of liquid (less than 1 μL per 1 mg of substrates) was observed.[39-40] This is referred to as "Liquid-assisted grinding" (LAG).[41] In our previous studies, we have shown that ^{17}O-enriched water can be used for LAG, by playing the role of both a "liquid assistant" and a ^{17}O-labeled reagent at the same time, producing labeled carboxylic acid-containing molecules (including fatty acids) *via* either a two-step CDI-activation/hydrolysis protocol or a one-pot saponification reaction.[42-44] It is both the rapidity and the minimal consumption of costly ^{17}O-enriched water that makes the use of ball-milling advantageous compared to the traditional solution-based labeling schemes.

In this work, we present at first how ^{17}O-enriched *free* amino acids could be obtained easily in a fast, user-friendly, cost- and atom-efficient manner using mechanochemical saponification, consuming only microliter amounts of costly labeled water. The versatility of this approach is demonstrated in the enrichment of diverse amino acids selected based on their different physicochemical properties and reactivity, related to their respective side-chains. In particular, we focused on: Glycine (Gly, G), as the simplest and the most studied amino acid; L-Leucine (Leu, L) and L-Phenylalanine (Phe, F) as representatives of amino acids with non-polar side-chains; L-Aspartic acid (Asp, D) with an additional carboxylic group in the side-chain; and finally L-Lysine (Lys, K) having a basic side-chain (Figure 1a). This study will show how the labeled amino acids could be isolated in their zwitterionic forms reproducibly with

high enrichment levels and with high enantiomeric excesses. Comparison to previously reported procedures based on acid-catalysed oxygen exchange will also be proposed.

Subsequently, a two-step protocol for synthesising ^{17}O-enriched Fmoc/*t*Bu *orthogonally protected* amino acids will be introduced, combining mechanochemistry with classical solution Fmoc-protection synthesis, and leading to highly enriched, essentially pure Fmoc-AA-O*H (Figure 1b). Furthermore, for the first time, the application of mechanochemical saponification for the site-selective ^{17}O-enrichment of a whole peptide chain will be described, in the case of two short peptides widely encountered in biological systems, namely RGD and GRGDS (Figure 1c). Following the synthetic part, ^{17}O solid-state NMR spectra of thus ^{17}O-enriched biomolecules will be presented and discussed.

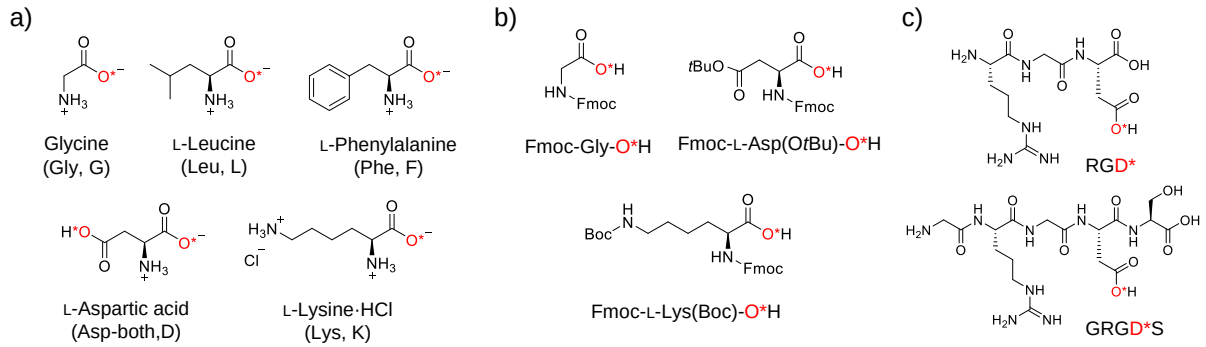


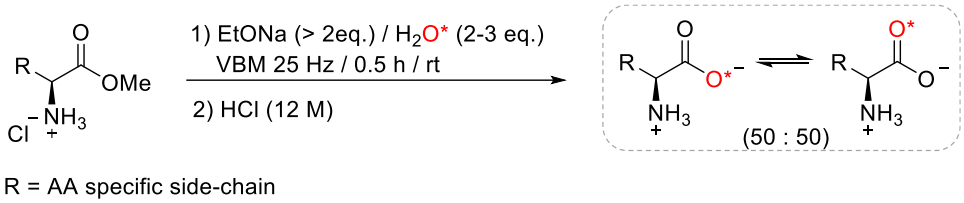
Figure 1: Scope of ^{17}O-labeled molecules prepared in this work: a) *free* amino acids, b) Fmoc/*t*Bu-*protected* amino acids and c) peptides. Only one oxygen is labeled per carboxylic function using mechanochemistry, but globally, the label is distributed between both oxygen sites (*i.e.* C=O* and C-O*H, 50 : 50), as evident from ^{17}O NMR analyses.

RESULTS AND DISCUSSION:

Synthesis of ^{17}O-enriched free amino acids by mechanochemistry

The ^{17}O-enriched *free* amino acids were prepared by a mechanochemical saponification route (Scheme 1), which had been developed previously in our laboratory as an efficient labeling scheme for the ^{17}O/^{18}O-enrichment of fatty acids.[44] Here, the saponification reactions were performed using a vibratory ball-mill (VBM) Retsch MM400 equipped with a stainless-steel jar (10 mL inner volume) containing two stainless-steel beads (1 cm diameter), and the following reagents: an amino acid methyl ester hydrochloride salt (H-AA-OMe·HCl), sodium ethoxide (NaOEt, 2.5 – 3.5 equivalents), and labeled water (H_2O*, 2 - 3 equivalents per carboxylic group to be labeled) (Scheme 1). After 30 min of milling at 25 Hz under ambient conditions, complete conversion was confirmed by ATR-IR analyses by looking at the shift of carbonyl stretching vibration band (from ~ 1740 to ~ 1560 cm^{-1}, see ESI).

Scheme 1: Mechanochemical saponification procedure used to enrich *free* unprotected amino acids.

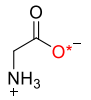
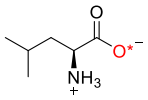
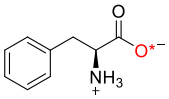
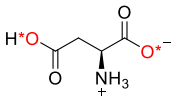
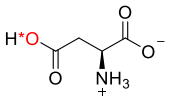
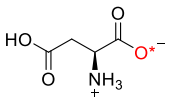
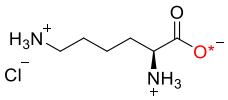


This one-pot synthesis was performed with minimal excess of reagents (especially H_2O*) with respect to the amino acid substrate. The excess of NaOEt (> 2 equivalents) was required due to the substrate form (used here as a hydrochloride salt), hence, one NaOEt equivalent was consumed simply by reacting with the hydrochloride. Enriched hydroxide anions were formed *in-situ* by an acid-base reaction between H_2O* (used in 2 - 3 equivalents excess per ester-derivatized carboxylic function) and remaining NaOEt, and enabled to hydrolyse the ester function in the substrate producing the labeled molecule (in a carboxylate form) and small amounts of ethanol and methanol (side-products easy to eliminate by evaporation). The reaction mixture was recovered from the jar using minimal amounts of non-labeled water, and the pH was adjusted accordingly to obtain the labeled amino acids in their zwitterionic form (except for Lys, which was obtained as the hydrochloride salt, see Table 1). The work-up procedures were optimised to isolate each of the labeled species free of the NaCl side-product, while maintaining the highest yields possible in a reproducible manner. No loss of enrichment was observed upon the work-up conditions applied. Their detailed descriptions

can be found in the Supplementary Information (ESI). For each amino acid, the labeling schemes were first optimised using ^{18}O-enriched water due to its lower cost, before producing the ^{17}O-labeled molecules.

The synthetic and enrichment yields, as well as the enrichment levels and enantiomeric excesses for each free amino acid $^{17}O/^{18}O$-enriched by this mechanochemical saponification procedure are summarized in Table 1. Here, as a starting point, the initial mass of amino acid methyl ester reagents was set to 150 mg (0.6 - 1.2 mmol), referred to as a "low scale" in this text. With this scale, ~ 60 - 110 mg of pure oxygen labeled free amino acids could be produced in one reaction while consuming only ~ 25 - 55 µL of labeled H_2O*. The corresponding isolated yield varied from around 60 % for the highly water-soluble amino acids (Gly, Lys), to approximately 80 % for the less soluble molecules (Asp, Leu) (see Table 1). An increase in isolated yield of ~ 11 % was observed, on average, when the reaction was scaled-up, producing up to ~ 200 mg of labeled amino acid in a single run. The purity of labeled compounds was confirmed by ^{1}H and ^{13}C solution NMR and powder XRD analyses. Moreover, the enantiomeric excess of all optically active enriched amino acids was verified by chiral HPLC analysis of their Fmoc-protected derivatives, and no loss of enantiomeric excess upon the mechanochemical saponification was observed (see the ESI).

Table 1: Summary of results for $^{17}O/^{18}O$-labeled *free* amino acids prepared using mechanochemical saponification.

Enriched amino acid		Low scale[a] Mass/Yield	Scale-up[b] Mass/Yield	EY[c]	^{18}O-EL[c]	^{17}O-EL[c]	Enantiomeric excess (*ee*)
	Gly (G)	58 ± 3 mg 64 ± 3 %	133 ± 3 mg 73 ± 1 %	89 ± 2 %[d]	44 ± 1 %[d]	41 %[d]	-
	Leu (L)	86 ± 3 mg 78 ± 3 %	N.P.	82 ± 1 %	41 ± 1 %	29 %	> 99 %
	Phe (F)	71 ± 3 mg 61 ± 2 %	N.P.	92 ± 2 %	46 ± 1 %	32 %	> 99 %
	Asp-both[e] (D)	80 ± 2 mg 76 ± 1 %	174 ± 5 mg 84 ± 3 %	88 ± 1 %	44 ± 1 %	31 %	> 99 %
	Asp-side[e] (D)	77 ± 7 mg 70 ± 6 %	N.P.	85 ± 1 %	42 %	31 %	N.D.
	Asp-main[e] (D)	113 ± 12 mg 83 ± 11 %	N.P.	95 ± 1 %	47 %	34 %	N.D.
	Lys·HCl[f] (K)	88 ± 5 mg 62 ± 4 %	221 ± 6 mg 78 ± 2 %	90 ± 2 %	46 ± 1 %	31 %	> 99 %

[a] Low scale corresponds to a reaction starting from 150 mg of H-AA-OMe·xHCl substrate; [b] Scale-up corresponds to a reaction starting from 300 mg of H-AA-OMe·xHCl; [c] EY corresponds to $^{17}O/^{18}O$-enrichment yield, ^{18}O-EL and ^{17}O-EL correspond to enrichment level of oxygen-18 and oxygen-17 isotopes per carboxylic oxygen using 99% ^{18}O-enriched or 70% ^{17}O-enriched water, respectively, and were calculated from HRMS analyses; [d] EL estimated from measurement of protected Fmoc-Gly-O*H,[45] 90% ^{17}O-enriched water was used here for the ^{17}O-labeling; [e] Asp-both, Asp-side and Asp-main correspond to L-Asp labeled either on both carboxylic functions, on the side-chain only or on main-chain only; [f] L-lysine isolated as monohydrochloride dihydrate. Amino acids were enriched using 2 or 3 eq. excess of $^{17}O/^{18}O$-enriched H_2O* per ester-derivatized carboxylic function. The synthetic and enrichment yields correspond to the average values of both ^{18}O and ^{17}O-enriched products (when available), the enrichment levels correspond to the products produced at the low-scale. Error bars are calculated from repeated experiments (n = 2 - 4). One labeled oxygen is introduced per carboxylic function, but globally, the label is distributed between both oxygen sites (*i.e.* C=O* and C-O*H). N.P. = Not Performed, N.D. = Not Determined.

Regarding the labeling efficiency of the proposed protocols, the enrichment yield of the labeling scheme (EY) and average enrichment level per carboxylic oxygen (EL) were calculated from high-resolution mass analyses (HRMS) of isolated products. Due to the reaction mechanism of saponification, only one labeled oxygen can be introduced per carboxylic

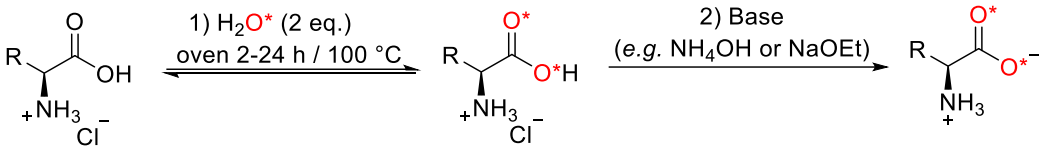
function, which results in maximal theoretical average enrichment level per carboxylic function ≤ 50 %. For the mechanochemical saponification, the enrichment yield was typically around 90 % using only a small excess of labeled water (2 - 3 equivalents per carboxylic function), which enabled to obtain the ^{17}O-labeled amino acids with high enrichment levels (29 - 34 % using 70% ^{17}O-enriched H_2O*, see Table 1). This allowed ^{17}O solid-state NMR spectra to be recorded with good signal-to-noise ratios in very short times (details provided later in this text and ESI). In the case of L-aspartic acid (Asp), which contains two carboxylic functions, the site-selective labeling could also be achieved, producing Asp enriched either on both carboxylic sites or only on one, by varying the substrate used for the saponification (di-methyl ester or a specific mono-methyl ester derivative, respectively).

Up to date, only a few examples of ^{17}O-labeling schemes based on saponification have been reported in the literature.[28, 46-47] These saponification reactions were performed in solution on both, the Boc-protected amino acid esters and on *N*-unprotected esters. In general, only a small excess of labeled water (2 - 5 equivalents) was used, and the ester function was hydrolysed in the presence of sodium hydroxide within several hours of stirring at room temperature (up to 12 h). However, the isolated yields were typically not reported, and an additional Boc-removal step was required to access the *free* labeled amino acids, thereby prolonging the overall reaction time.[28, 47] Here in this work, mechanochemistry enables to access highly enriched molecules (up to 41 % in ^{17}O), under ambient conditions, rapidly (30 min reaction time), without a need for aggressive reagents, cost-efficiently, and consuming only microliter quantities of costly labeled water (~ 25 – 55 μL per reaction at the low scale).

Synthesis of ^{18}O-enriched free amino acids by acid-catalysed exchange

To evaluate the efficiency of the mechanochemical saponification approach, a comparison was made with an acid-catalysed oxygen exchange enrichment procedure, which, as mentioned previously, is the one used the most in the literature for enriching amino acids. More specifically, we adapted a procedure reported by Mears *et al.* in 1938 for producing ^{18}O-enriched glycine, starting from the glycine hydrochloride salt (Gly·HCl) and heating it in the presence of ^{18}O-enriched H_2O* for 24 h (Scheme 2).[48] The use of Gly·HCl salt as a substrate is indeed particularly interesting, as it avoids having to produce and engage HCl gas, which is experimentally constraining, or having to add aqueous HCl to the amino acid, which would reduce the labeling efficiency due to the presence of additional non-labeled water.

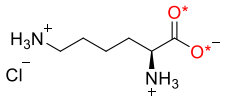
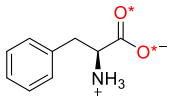
Scheme 2: Acid-catalysed oxygen exchange procedure used for the enrichment of *free* amino acids. Both oxygen atoms per carboxylic function can become labeled by this procedure.



R = AA specific side-chain

Here, for comparison purposes, the molar amounts and stoichiometry of reagents in the acid-catalysed oxygen exchange procedure of glycine were adjusted to be identical to the low scale conditions used for the mechanochemical saponification, and the work-up conditions were likewise adapted to provide the same forms of labeled products. All tests were performed here using ^{18}O-enriched water only. The reaction was carried out in a vial containing the Gly·HCl substrate and ^{18}O-enriched H_2O* (2 equivalents), which was closed with a screw cap and placed in an oven heated at 100 ºC for 24 h (*i.e.* using the same heating time as initially reported by Mears *et al.*).[48] The results are presented in Table 2. The enrichment level achieved for Gly was estimated to be around 49 %,[45] which is slightly higher compared to the enrichment level achieved using saponification (44 %), and corresponds to almost quantitative conversion (the maximum enrichment level expected was ~ 50 %, see ESI for calculation details). However, this enrichment level was reached after 24 h of heating at 100 °C in contrast to the VBM saponification, where the reaction was complete within 30 min of milling at ambient temperature. To further approach the experimental times used in the mechanochemical saponification, the heating duration was reduced to 2 hours. Under these conditions, the conversion determined in the isolated phase was already at 91 %, which corresponds to an enrichment level of 46 %, and suggests that the kinetics of the oxygen isotopic exchange are in fact very fast for this substrate under the conditions applied.

Table 2: Summary of results for ^{18}O-labeled *free* amino acids prepared *via* acid-catalysed exchange.

Enriched amino acid		Reaction time	EY[a]	^{18}O-EL[a]	Enantiomeric excess (*ee*)
	Gly (G)	2 h	91 ± 3 %[b]	46 ± 1 %[b]	-
		24 h	97 ± 1 %[b]	49 ± 1 %[b]	-
	Lys·HCl (K)	2 h	84 ± 3 %	42 ± 1 %	N. D.
		24 h	90 ± 1 %	45 ± 1 %	> 99 %
	Phe (F)	24 h	77 ± 5 %	39 ± 2 %	N. D.

[a] EY corresponds to ^{18}O-enrichment yield, ^{18}O-EL corresponds to enrichment level of oxygen-18 using 99% ^{18}O-enriched water, calculated from HRMS analyses; [b] EL estimated from measurement of protected Fmoc-Gly-O*H.[45] All amino acids were enriched using 2 eq. of ^{18}O-enriched H_2O* per carboxylic function. Error bars are calculated from repeated experiments (n = 2), N.D. = Not Determined.

Due to the simplicity of this labeling scheme for glycine, its extension for oxygen enrichment of other amino acids commercially available as HCl salts, namely L-lysine and L-phenylalanine, was further investigated. First, in the case of L-lysine (Lys), Lys·2HCl was used as a substrate as no enrichment was observed when starting from the monohydrochloride salt. Using 2 equivalents of H_2O*, a 45% average enrichment level per carboxylic oxygen was determined after 24 h of heating at 100 °C, which is equivalent to what could be reached by VBM saponification (46 %) within 30 min of milling at room temperature. Just like for glycine, when the reaction time was reduced to 2 hours, only a small decrease in the final enrichment level was detected (42 %) (see Table 2). Second, in the case of L-phenylalanine hydrochloride (Phe·HCl), a first study of its reactivity towards oxygen exchange was performed, and the preliminary results are also presented in Table 2. Despite the lower solubility of the Phe·HCl substrate in water, which remained substantially non-dissolved after 24 h of heating at 100 °C (compared to Gly·HCl and Lys·2HCl, which were completely dissolved), an average 39% enrichment level was achieved. Overall, these results suggest that this labeling scheme, starting from HCl salts, may be applicable to other amino acid hydrochlorides.

As not all amino acids are commercially available as HCl salts (notably aspartic acid), an alternative related approach was also tested as part of this work. Here, the enrichment of L-Asp was performed by adding to the free amino acid LiCl, to catalyse the oxygen isotopic exchange using Li^+ as Lewis acid. After 24 h of heating at 100 °C in the presence of 4 equivalents of ^{18}O-enriched H_2O* (*i.e.* 2 equivalents per carboxylic function), and

1 equivalent of LiCl, a minor enrichment level of around 5 % was detected in the isolated product. Reaction conditions were not further optimized at this stage. Yet, most importantly, it is worth highlighting that no enrichment was observed when the same reaction was performed in the absence of LiCl. To the best of our knowledge, this test is the first application of Lewis acid catalysis for oxygen isotope exchange of carboxylic groups in amino acids, which could offer an interesting alternative for some substrates.

Overall, the acid-catalysed isotopic exchange presented here enabled obtaining ^{18}O-enriched amino acids easily without the need for toxic and corrosive HCl gas. Moreover, the reaction time could be significantly reduced to only 2 hours (compared to the initially reported 24 h),[48] still providing highly enriched molecules (up to 46 %, using 99% ^{18}O-enriched H_2O*). Therefore, this can be seen as an attractive cost-efficient alternative to the VBM saponification. On the other hand, its weakness lies in the commercial availability of starting hydrochloride salts, which are not as easily accessible as the amino acid esters used for saponification. Furthermore, unlike in the VBM approach, the acid-catalysed oxygen exchange protocol does not allow straightforward site-selective labeling of amino acids containing additional carboxylic groups in their side-chains (*e.g.* aspartic or glutamic acid).

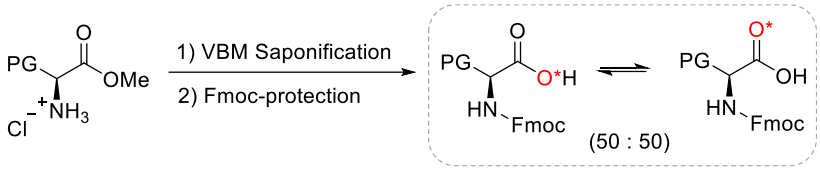
Synthesis of ^{17}O/^{18}O-enriched Fmoc-protected amino acids

In contrast to the recombinant protein synthesis, which uses ^{17}O/^{18}O-labeled *free* amino acids, SPPS enables to selectively enrich only the chosen carbonyl positions within a peptide chain using ^{17}O/^{18}O-labeled Fmoc/*t*Bu *orthogonally protected* amino acids as building blocks. This approach is particularly convenient for subsequent ^{17}O NMR studies where direct information from the specific oxygen environment can be obtained. Since none of the oxygen labeled *protected* amino acids is commercially available (to the best of our knowledge), their in-house preparation is an indispensable prerequisite.

In this work, three Fmoc-protected amino acids were enriched in ^{17}O (Figure 1b): the Fmoc-Gly-O*H molecule with no side-chain protection group, and the two orthogonally protected molecules Fmoc-L-Asp(O*t*Bu)-O*H and Fmoc-L-Lys(Boc)-O*H. All were obtained in two steps combining mechanochemistry (used for enrichment) with classical solution-based chemistry (used for Fmoc-protection) (Scheme 3). For each substrate, the reaction conditions were first optimised using ^{18}O-enriched water due to its lower purchasing price before producing the ^{17}O-labeled equivalents. In the first step, the amino acids were enriched by mechanochemical saponification starting from their methyl ester derivatives (as described in the previous section). However, in order to increase the overall isolated yield, the labeled amino

acid intermediates were not isolated. Instead, the reaction medium after saponification was dissolved in non-labeled water, and recovered from the jar. After a minor pH adjustment, the subsequent Fmoc-protection step was directly performed in solution following published procedures (for synthetic details, see ESI). No loss of enrichment was observed upon the Fmoc-protection steps. Notably, the *tert*-butyl group at the *β*-carboxylate of Asp as well as the *tert*-butyloxycarbonyl (Boc) protection group at the side-chain of Lys were found to be stable under the basic saponification conditions when using exactly 2 equivalents of NaOEt, with only limited side-chain deprotection observed in the case of Asp.

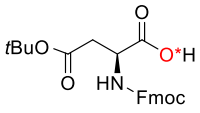
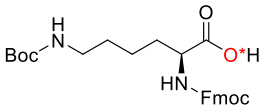
Scheme 3: Two-step procedure used for the synthesis of $^{17}O/^{18}O$-enriched Fmoc/*t*Bu-*protected* amino acids.



PG = side-chain of a given amino acid protected with *t*Bu-group

Yields and enrichment levels of $^{17}O/^{18}O$-enriched Fmoc/*t*Bu-*protected* amino acids are reported in Table 3. Because SPPS synthesis is generally more efficient when using large excesses of Fmoc-protected amino acids, the production of Fmoc-Gly-O*H and Fmoc-L-Asp(O*t*Bu)-O*H was also scaled-up as part of this work, providing up to ~ 1 g of labeled products within 1 day (work-up included).[49] Importantly, the final products were isolated essentially pure, in high yields, and with high enrichment levels (up to ~ 40 % using 90% ^{17}O-enriched water).

Table 3: Summary of results for $^{17}O/^{18}O$-labeled Fmoc/*t*Bu-*protected* amino acids.

Enriched Fmoc/*t*Bu protected amino acid	Low scale[a] Mass/Yield	Scale-up[b] Mass/Yield	^{18}O-EL[c]	^{17}O-EL[c]
 Fmoc-Gly-O*H	225 ± 13 mg 75 ± 4 %	496 ± 7 mg 83 ± 1 %	43 ± 1 %	38 ± 1 %
 Fmoc-L-Asp(O*t*Bu)-O*H	165 ± 7 mg 71 ± 3 %	342 ± 3 mg 74 ± 1 %	43 ± 2 %	40 ± 1 %[d]
 Fmoc-L-Lys(Boc)-O*H	180 ± 24 mg 76 ± 10 %	N. P.	46 ± 1 %	32 %[e]

[a] Low scale corresponds to a reaction starting from 150 mg of H-AA-OMe·xHCl substrate; [b] Scale-up corresponds to a reaction starting from ~ 300 mg of H-AA-OMe·xHCl; [c] ^{18}O-EL and ^{17}O-EL correspond to enrichment level of oxygen-18 and oxygen-17 isotopes per carboxylic oxygen, using 99% ^{18}O-enriched or 90% ^{17}O-enriched water, calculated from HRMS analyses; [d] ^{17}O-EL of the products prepared under scaled-up conditions; [e] 70% ^{17}O-enriched water was used. The synthetic yields correspond to the average values of both ^{18}O and ^{17}O-enriched products (when available), the enrichment levels correspond to the products produced at the low-scale. Error bars are calculated from repeated experiments (n = 2 - 5). One labeled oxygen is introduced per carboxylic function, but globally, the labeled is distributed between both oxygen sites (*i.e.* C=O* and C-O*H). N.P. = Not Performed.

The two-step procedure described above was then compared to the protocols previously reported in the literature for the ^{17}O-enrichment of Fmoc/*t*Bu-orthogonally protected amino acids. To the best of our knowledge, only two methods working under mild conditions compatible with the presence of side-chain protection groups have been described so far.[32-33] In the work of Yamada *et al.*,[33] Fmoc/*t*Bu protected amino acids were enriched *via* hydrolysis of their pentafluorophenyl esters at room temperature and nearly neutral pH, using a small excess of ^{17}O-labeled H_2O* (5 equivalents), but with long reaction times (up to several days). In another approach developed by Seyfried *et al.*,[32] the protected amino acids were enriched based on multiple turnover *in-situ* activation/hydrolysis cycles in the presence of coupling reagent (EDC·HCl) and ^{18}O-labeled H_2O*. The final molecules were isolated with high enrichment levels (up to 95 % using 95% ^{18}O-enriched H_2O*) but to the expense of using a large excess of labeled water (50 equivalents) and very long reaction times (up to 40 h).[32] Recently, this method

was adapted and optimized to be suitable for ^{17}O-enrichment.[34, 50-51] The reaction was performed within 8 h, and the labeled Fmoc-AA-O*H products were isolated with enrichment level around 40 % (using 40% ^{17}O-enriched H_2O*), which is sufficient for most high-resolution ^{17}O NMR studies.[34] Nevertheless, despite the optimizations, the excess of ^{17}O-enriched water involved remained rather high (30 equivalents). Moreover, neither the synthetic yields nor the purity of isolated ^{17}O-labeled Fmoc-AA-O*H were reported, which together with the water consumption, will affect the final labeling cost.

In contrast to the previously published enrichment schemes, our methodology provides ^{17}O-enriched Fmoc-protected products (up to ~ 40 % enrichment) in high isolated yields (up to ~ 85 %) in just one day, consuming only 2 - 3 equivalents of ^{17}O-enriched water per amino acid. As a result, the labeling scheme presented here is exceptionally efficient in terms of the labeling cost, which is a critical factor especially for the ^{17}O-enrichment. In cases when up to 1 g quantities of Fmoc-protected amino acid were produced within 1 day, the overall water consumption per 100 mg of *isolated product* was only ~ 18 μL (90% ^{17}O-enriched H_2O*). This corresponds to a labeling cost of 50 € / 100 mg of *isolated* ^{17}O-labeled Fmoc-AA-O*H with ~ 40% ^{17}O-enrichment level (see supporting information for full details on the calculation of the enrichment cost).[52] Compared to the very recent work from Ha *et al.*,[50] where the labeling cost was calculated for 100 mg of *starting Fmoc-AA-OH* based on kinetics experiments, and considering a negotiated labeled water price,[53] a different approach for calculating the labeling cost was reported here, which, we believe, brings a more straightforward and informative estimation for other researchers interested in ^{17}O-enrichment of biomolecules.

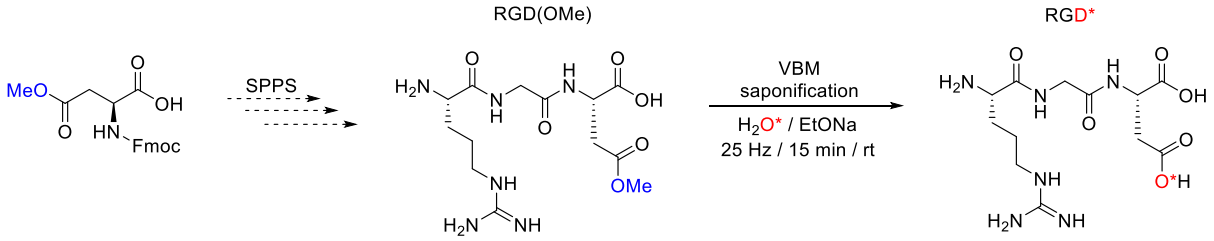
Site-selective ^{17}O-labeling of peptide side-chains using mechanochemistry

Using ^{17}O-enriched Fmoc/*t*Bu-*protected* amino acids (like those prepared as described in the previous section), it is possible to produce ^{17}O-labeled peptides by SPPS. Yet, when starting from the precursors reported in Table 3, only the carbonyl oxygen in the peptide back-bone chain will be labeled. As the ^{17}O-labeling of peptide side-chains is also important to gain information on their three-dimensional structure and/or reactivity, we looked into the possibility of enriching the *β*-carboxylic function in aspartic acid side-chain by performing the mechanochemical saponification on a whole peptide chain. As model substrates, a short peptide sequence Arginyl-glycyl-aspartic acid (Arg-Gly-Asp, RGD) and its longer derivative Gly-Arg-Gly-Asp-Ser (GRGDS) were selected. The RGD sequence serves indeed as a key structural recognition motif for cell transmembrane protein receptors (integrins), which facilitate cell-cell and cell-extracellular matrix adhesion,[54] and the RGD-related peptides are involved in many

biological applications (such as drug delivery systems or coatings of implants).[55-57] Moreover, both peptides are commercially available, which highlights their general interest for the research community.

Here, the RGD tri-peptide and GRGDS penta-peptide were site-selectively ^{17}O-labeled at the *β*-carboxylic group of the aspartic acid unit, by mechanochemical saponification of the corresponding peptides containing an aspartic acid side-chain initially protected with methyl ester (prepared by SPPS, for synthetic details see ESI). The mechanochemical saponification was performed on only 8 mg of the methyl ester protected peptides in a stainless-steel jar (1.5 mL inner volume) containing two stainless steel beads (5 mm diameter) in the presence of NaOEt (3.5 equivalents) and ^{17}O-labeled water (7 equivalents, 90% in ^{17}O) (Scheme 4). After 15 min of milling at 25 Hz, the content of the jar was recovered using non-labeled water and then acidified to pH ~ 3. Finally, ^{17}O-labeled peptides were isolated in a mixture with NaCl side-product *via* lyophilisation (~ 15 mg of a peptide-NaCl mixture), and characterized by LCMS, ^{1}H solution NMR and ^{17}O solid-state NMR analyses (for details, see ESI). No degradation of the peptide chains upon the milling treatment was observed.

Scheme 4: Synthesis of site-selectively ^{17}O-enriched RGD* tri-peptide by VBM saponification of methyl ester protected peptide. Only one oxygen is labeled per ester-derivatized carboxylic function using mechanochemistry, but globally, the labeled is distributed between both oxygen sites (*i.e.* C=O* and C-O*H, 50 : 50).



The enrichment levels of ^{17}O-labeled RGD* and GRGD*S peptides calculated from HRMS analyses of the isolated phases were 29 and 27 %, respectively, which enabled to record the ^{17}O MAS NMR spectra at 14.1 T (Figure 2b). The fact that a spectrum could be recorded on only ~ 3 mg of peptide-NaCl mixture underscores the high labeling efficiency of VBM saponification producing molecules with enrichment levels high enough for ^{17}O NMR analyses, when incorporated in more complex biological systems. Overall, the mechanochemical saponification approach represents a straightforward and cost-efficient access to side-chain ^{17}O-labeled peptides, which may be easily adapted for enrichment of other peptides containing aspartic or glutamic acid residues that are key to biological processes, due to the presence of their pending carboxylic functions.

^{17}O solid-state NMR analyses of the ^{17}O-enriched amino acids and peptides

As developed in more detail in previous sections, using mechanochemical saponification enabled obtaining ^{17}O-enriched free amino acids, their Fmoc/*t*Bu protected derivatives and also selectively side-chain labeled short peptides, all with high enrichment levels (up to ~ 40 %, using 90% ^{17}O-enriched H_2O*). Notably, the high labeling efficiency of VBM saponification allowed ^{17}O MAS NMR spectra of amino acids to be recorded in just a few hours at 14.1 T (see ESI for acquisition details).

The majority of ^{17}O-labeled products prepared in this work were in fact characterized in their zwitterionic form by ^{17}O solid-state NMR spectroscopy for the first time. Indeed, the ^{17}O MAS NMR spectra reported in the literature generally correspond to the amino acid HCl salts resulting from the HCl catalysed oxygen exchange labeling scheme traditionally employed for oxygen enrichment.[30, 58] Some of the recorded ^{17}O MAS NMR spectra are presented in Figure 2, while the rest can be found in ESI (section VI). Below, a more detailed analysis of the ^{17}O NMR data obtained for Gly, Lys, Asp, and the labeled peptides is provided.

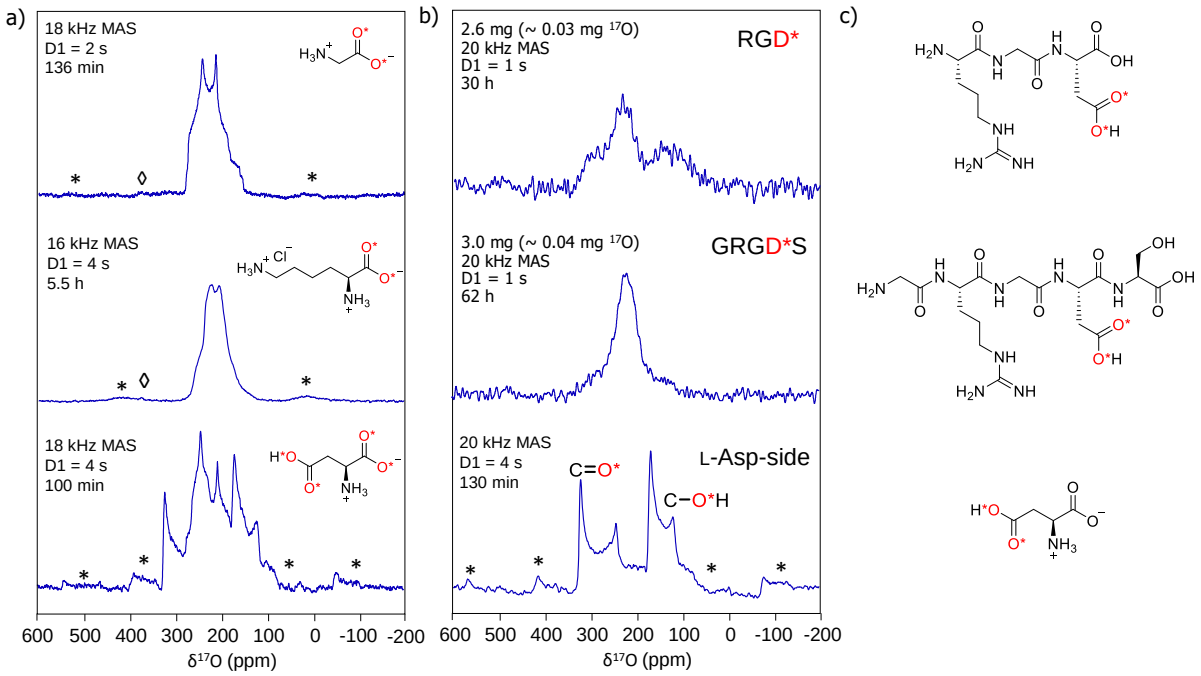


Figure 2: ^{17}O MAS NMR spectra recorded at 14.1 T with temperature regulated at 0 ºC of ^{17}O-labeled a) Gly, L-Lys·HCl (isolated as dihydrate phase) and L-Asp-both (labeled at both carboxylic groups). b) RGD* tri-peptide and GRGD*S penta-peptide labeled at Asp side-chain (isolated *via* lyophilisation, spectra recorded using DFS Hahn echo sequence, see ESI for acquisition parameters) compared to Asp labeled selectively only at the side-chain (L-Asp-side, crystalline phase). c) Chemical structures of ^{17}O-labeled RGD* tri-peptide, GRGD*S penta-peptide and L-Asp labeled at the side-chain. For Gly and L-

Asp-both, samples were physically mixed with silica before packing the rotor, to prevent preferential orientation of the crystallites (leading to lineshape distortions in the initial analyses performed). Symbol "*" corresponds to the spinning side-bands, symbol "◊" corresponds to the ^{17}O natural abundance signal of the zirconia rotor. Only one oxygen was labeled per carboxylic function using mechanochemistry, but globally, the label is distributed between both oxygen sites in the isolated products (*i.e.* C=O* and C-O*H, 50 : 50), which is why both carboxylic oxygen atoms are shown here as labeled in red. The acquisition conditions and parameters extracted from the fits can be found in the ESI.

In the solid state, glycine is known to exhibit several polymorphs under ambient conditions, and their relative stability and interconversions have been the subject of previous NMR studies.[59-60] Here, the recorded ^{17}O NMR spectrum shown in Figure 2a-top belongs to *β*-glycine (the least stable Gly polymorph), which was obtained as a result of work-up conditions used for its isolation. The ^{17}O NMR spectrum was tentatively fitted with two oxygen sites (as expected from its crystal structure),[61] and the extracted parameters were in line with the previously published data for this polymorph (see ESI).[62] It should be highlighted that our initial ^{17}O NMR analyses on this phase also revealed that careful attention needed to be paid to preferential orientation of the crystallites in the rotor (which could lead to lineshape distortions), and to the possible evolution of the *β*-form towards a more stable polymorph.

Regarding ^{17}O-enriched Lys·HCl, likewise two oxygen sites were expected based on its crystal structure.[63] However, as apparent from Figure 2a-middle, the recorded ^{17}O MAS NMR spectrum of L-Lys·HCl phase prepared here consists of overlapping resonances situated in the middle of the chemical shift range typical for carboxylic functions, and is lacking the sharp features often observed for second-order quadrupolar lineshapes of carboxylic groups in crystalline phases. This featureless appearance could be due to ongoing dynamics in proximity of the carboxylic function,[64-65] although additional investigations would be needed to confirm this, which is out of the scope of this publication. Interestingly, in a previously published study by Pike *et al.*,[30] the ^{17}O MAS NMR spectrum of a so-called "Lys·HCl" phase had been reported, which was composed of two separated oxygen signals positioned with isotropic shifts at 347 and 181 ppm (*i.e.* of C=O and C-OH, respectively). Considering the difference with the ^{17}O signals found here, and based on the results reported for other AA·HCl phases,[30, 58] we hypothesized that the isolated phase described by Pike *et al.* was in fact a lysine dihydrochloride salt (Lys·2HCl). Indeed, based on the synthetic experimental details provided,[30] their lysine phase had been labeled in ^{17}O using an acid-catalysed exchange. Yet, unfortunately, no information regarding pH adjustment *prior* to its isolation had been mentioned. To verify this hypothesis, we prepared Lys·2HCl *via* the re-optimized acid-catalysed exchange procedure

(described in previous section), and characterized it by ^{17}O MAS NMR (spectrum available in ESI). The extracted parameters of thus prepared phase were found to correspond to the data previously published by Pike *et al.*,[30] confirming that the compound they had studied was in fact Lys·2HCl.

In the ^{17}O spectrum of L-Asp labeled at both carboxylic functions, four oxygen sites were expected based on the crystal structure.[66] Here, four overlapping oxygen resonances were observed (Figure 2a-bottom). The quadrupolar parameters for all four sites were successfully extracted taking advantage of the possibility of performing site-selective labeling (by enriching only the side-chain (β) or main-chain (α) carboxylic function), combined with multiple-field NMR analyses (details in ESI). In the literature, only the ^{17}O NMR parameters of the α-carboxylic group had been previously published,[67] but they were found to differ from the parameters obtained here. This discrepancy might come from the fact that a different form of L-Asp had been studied (possibly a different polymorph or hydrate), or that a different temperature had been used for the analysis. Unfortunately, this information was not provided in the previously published work.[67]

In Figure 2b, the ^{17}O solid-state NMR spectra of RGD* and GRGD*S peptides labeled selectively at the side-chain of the aspartic acid unit are shown, together with the spectrum of L-Asp labeled likewise solely at the β-carboxylic function (recorded at 14.1 T). Compared to the crystalline L-Asp phase, the environment around oxygen atoms in peptides (obtained by lyophilisation) is distributed, resulting in broad overlapping oxygen resonances situated in between the two extremes (*i.e.* defined C=O and C-OH sites). Further high-resolution analyses were not performed at this stage. It is important to highlight that it is the first time that ^{17}O solid-state NMR spectra of RGD* and GRGD*S peptides were recorded, and the results are all the more impressive considering the very small amount of sample analysed (~ 3 mg of peptide/NaCl mixture), which corresponds to ~ 0.04 mg of ^{17}O. Altogether, owing to the high enrichment level achieved by VBM saponification, the results shown here represent a promising prerequisite for the peptides to be studied by ^{17}O NMR spectroscopy when incorporated in more complex biological systems. Such studies could then shed light, for example, on the functions of integrins in cell adhesion processes by providing a unique information regarding oxygen binding state, due to the high sensitivity of oxygen parameters to its local environment. Such studies will be part of future work.

CONCLUSION:

In this work, we have at first presented simple, fast and highly cost-efficient ^{17}O-labeling procedures for the enrichment of *free* amino acids, namely: Gly, L-Leu, L-Phe, L-Asp and L-Lys·HCl. Using mechanochemistry, we were able to produce the labeled molecules rapidly (30 min reaction time) under ambient conditions, consuming minimal amounts of costly ^{17}O-labeled water (only 2 - 3 equivalent excess). The ^{17}O-enriched amino acids were isolated pure in their zwitterionic forms, in medium to high yields (~ 60 - 85 %), with no loss of their optical purity (*ee* > 99 %), and with high ^{17}O-enrichment levels (up to ~ 40 %). The second labeling scheme proposed here is based on acid-catalysed oxygen labeling protocol (adapted from a 1938 study by Mears),[48] and was re-optimized to serve as a comparison point for the mechanochemical saponification. The first results of ^{18}O-enriched *free* amino acids showed enrichment levels comparable with those after saponification (using the same amount of water), after only 2 hours of reaction time, making this approach an attractive alternative for the ^{17}O-enrichment of free amino acids, which are available as HCl salts, as exemplified here for L-Lys·2HCl. However, the broad application of this approach may be hindered by the lower commercial availability of amino acid hydrochloride salts compared to their ester derivatives, which are used as substrates in the mechanochemical saponification enrichment scheme.

Subsequently, the ^{17}O-enriched Fmoc/*t*Bu *orthogonally protected* amino acids, used as precursors in SPPS, were prepared under mild conditions by combining the VBM saponification with subsequent Fmoc-protection in solution. The labeled products were isolated with high purity, in high synthetic yields (~ 75 - 85 %), and with high enrichment levels (~ 40%). Scaling-up the mechanochemical step enabled to produce up to 1 g/day quantities of ^{17}O-labeled protected amino acids cost-efficiently, consuming only ~ 18 μL of 90% ^{17}O-enriched water per 100 mg of isolated Fmoc-AA-O*H, which corresponds to the labeling cost of 50 € / 100 mg of the labeled product. Finally, the VBM saponification was applied as a site-selective enrichment strategy for labeling solely the side-chain carboxylic functions in RGD and GRGDS peptide chains, while reaching ~ 28% enrichment level.

The high-enrichment efficiency of the above proposed protocols also appears through the rapidity with which the ^{17}O MAS NMR spectra of ^{17}O-labeled amino-acids were recorded in good signal-to-noise ratio at 14.1 T, many of them being reported here for the first time. Moreover, thanks to the highly efficient mechanochemical labeling scheme, the ^{17}O MAS NMR spectra of ^{17}O-labeled RGD and GRGDS peptides were recorded at 14.1 T using only ~ 3 mg

of peptides. Such analyses represent an important prerequisite for the subsequent application of amino acids and peptides in ^{17}O NMR studies of more complex biological systems.

Most importantly, in contrast with the majority of previously published studies, great attention was given here to provide efficient and reproducible protocols including extensive experimental details. We believe that this will facilitate their further adoption by other researchers in view of expanding the knowledge regarding the structure and dynamics of peptides and proteins accessible *via* ^{17}O NMR studies.

ACKNOWLEDGEMENTS:

This project has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 772204; 2017 ERC-COG, MISOTOP project). A part of the NMR analyses shown in supporting information were performed at the National High Magnetic Field Laboratory (which is supported by the National Science Foundation Cooperative Agreement No. DMR-1644779 and the State of Florida), and on instruments belonging to the French IR-RMN network in Orleans and Lille, with financial support from the IR-RMN-THC FR 3050 CNRS. Dr. Julien Trébosc (IMEC-ISB-UCCS, Lille, France), Drs. Ivan Hung and Zhehong Gan (MagLab, Tallahassee, USA), and Drs. César Leroy and Philippe Gaveau (ICGM, Montpellier, France) are acknowledged for their assistance with performing some of the ^{17}O NMR experiments. Dr. Nicolas Vanthuyne (iSm2, Marseille, France) is acknowledged for performing chiral HPLC analyses. Manon Maurel is acknowledged for her assistance with the solid-phase peptide synthesis, and Léa Tavares for her work on optimizing the synthesis of ^{17}O/^{18}O-labeled Fmoc-Gly-O*H and scaling-up the production of ^{17}O-labeled Fmoc-L-Asp(O*t*Bu)-O*H.

AUTHOR CONTRIBUTION:

The project was conducted by JŠ, in close interaction with DL and TXM. JŠ performed the syntheses of labeled products, prepared all the figures, schemes and tables. RY took part in the syntheses of ^{17}O/^{18}O-labeled Gly and Fmoc-Gly-O*H. IG and JŠ performed the solid-state NMR analyses, processed data and fitted the spectra. PV and JŠ performed the solid-phase peptide synthesis. GC and AL performed LCMS, HRMS and solution NMR analyses of the final products. JŠ, DL and TXM wrote the initial draft of the manuscript. All authors discussed the results and contributed to the final preparation of the manuscript.

REFERENCES:

1. Langel, U.; Cravatt, B. F.; Graslund, A.; Von Heijne, N.; Zorko, M.; Land, T.; Niessen, S., *Introduction to peptides and proteins*. CRC press: 2009.
2. Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D., The Future of Peptide-based Drugs. *Chem. Biol. Drug. Des.* **2013,** *81* (1), 136-147.
3. Levin, A.; Hakala, T. A.; Schnaider, L.; Bernardes, G. J. L.; Gazit, E.; Knowles, T. P. J., Biomimetic peptide self-assembly for functional materials. *Nat. Rev. Chem.* **2020,** *4* (11), 615-634.
4. Reid, C. M.; Sutherland, A., Synthesis of Isotopically Labeled α-Amino Acids. In *Amino Acids, Peptides and Proteins in Organic Chemistry*, pp 473-494.
5. Tian, X.; Permentier, H. P.; Bischoff, R., Chemical isotope labeling for quantitative proteomics. *Mass Spectrom. Rev.* **2021**, 1-31.
6. Ning, L.; Lichun, Z.; Chengyan, H.; Zongwei, C., Advances in Technologies and Biological Applications of 18O Labeling Strategies in LC-MS Based Proteomics: An Updated Review. *Curr. Anal. Chem.* **2012,** *8* (1), 22-34.
7. Arkin, I. T., Isotope-edited IR spectroscopy for the study of membrane proteins. *Curr. Opin. Chem. Biol.* **2006,** *10* (5), 394-401.
8. Ghosh, A.; Ostrander, J. S.; Zanni, M. T., Watching Proteins Wiggle: Mapping Structures with Two-Dimensional Infrared Spectroscopy. *Chem. Rev.* **2017,** *117* (16), 10726-10759.
9. Atreya, H. S., *Isotope labeling in biomolecular NMR*. Springer Science & Business Media: 2012; Vol. 992.
10. Wu, G., Solid-State 17O NMR studies of organic and biological molecules: Recent advances and future directions. *Solid State Nucl. Magn. Reson.* **2016,** *73*, 1-14.
11. Torres, J.; Kukol, A.; Goodman, J. M.; Arkin, I. T., Site-specific examination of secondary structure and orientation determination in membrane proteins: The peptidic 13C=18O group as a novel infrared probe. *Biopolymers* **2001,** *59* (6), 396-401.
12. Wang, L.; Skinner, J. L., Thermally induced protein unfolding probed by isotope-edited IR spectroscopy. *J. Phys. Chem. B* **2012,** *116* (32), 9627-9634.
13. Middleton, C.; Woys, A.; Mukherjee, S.; Zanni, M., Residue-Specific Structural Kinetics of Proteins through the Union of Isotope Labeling, Mid-IR Pulse Shaping, and Coherent 2D IR Spectroscopy. *Methods* **2010,** *52*, 12-22.
14. Truong, H. N. D.; Lim, J.-M., Quantitative Proteomics Using 18O Labeling on Target Peptides and Unlabeled Standards. *Bull. Korean Chem. Soc.* **2021,** *42* (4), 633-640.
15. Muniyappan, S.; Lin, Y.; Lee, Y.-H.; Kim, J. H., 17O NMR Spectroscopy: A Novel Probe for Characterizing Protein Structure and Folding. *Biology* **2021,** *10* (6), 453.
16. Wu, G., 17O NMR studies of organic and biological molecules in aqueous solution and in the solid state. *Prog. Nucl. Magn. Reson. Spectrosc.* **2019,** *114-115*, 135-191.
17. Ohki, S.-y.; Kainosho, M., Stable isotope labeling methods for protein NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **2008,** *53* (4), 208-226.
18. Goldberga, I.; Li, R.; Duer, M. J., Collagen Structure–Function Relationships from Solid-State NMR Spectroscopy. *Accounts. Chem. Res.* **2018,** *51* (7), 1621-1629.
19. Kay, L. E., NMR studies of protein structure and dynamics. *J. Magn. Reson.* **2011,** *213* (2), 477-491.
20. Ladizhansky, V., Applications of solid-state NMR to membrane proteins. *Biochim. Biophys. Acta - Proteins Proteom.* **2017,** *1865* (11, Part B), 1577-1586.
21. Gerothanassis, I., Oxygen-17 NMR spectroscopy: Basic principles and applications (Part I). *Prog. Nucl. Magn. Reson. Spectrosc.* **2010,** *56*, 95-197.
22. Lemaître, V.; de Planque, M. R. R.; Howes, A. P.; Smith, M. E.; Dupree, R.; Watts, A., Solid-State 17O NMR as a Probe for Structural Studies of Proteins in Biomembranes. *J. Am. Chem. Soc.* **2004,** *126* (47), 15320-15321.
23. Paulino, J.; Yi, M.; Hung, I.; Gan, Z.; Wang, X.; Chekmenev, E. Y.; Zhou, H.-X.; Cross, T. A., Functional stability of water wire–carbonyl interactions in an ion channel. *Proc. Natl. Acad. Sci.* **2020,** *117* (22), 11908-11915.
24. Lin, B.; Hung, I.; Gan, Z.; Chien, P.-H.; Spencer, H. L.; Smith, S. P.; Wu, G., 17O NMR studies of yeast ubiquitin in aqueous solution and in the solid state. *Chem. Bio. Chem.* **2021,** *22* (5), 826.

25. Cambridge Isotope Laboratoires, Inc. https://www.isotope.com/index.cfm.
26. Theodorou, V.; Skobridis, K.; Alivertis, D.; Gerothanassis, I. P., Synthetic methodologies in organic chemistry involving incorporation of [17O] and [18O] isotopes. *J. Labelled Comp. Radiopharm.* **2014,** *57* (8), 481-508.
27. Murphy, R. C.; Clay, K. L., Synthesis and back exchange of 18O labeled amino acids for use as internal standards with mass spectrometry. *Biomed. Mass Spectrom.* **1979,** *6* (7), 309-314.
28. Steinschneider, A.; Burgar, M. I.; Buku, A.; Fiat, D., Labeling of mamino acids and peptides with isotopic oxygen as followed by 17O-NMR. *Int. J. Pept. Protein Res.* **1981,** *18* (3), 324-333.
29. Ponnusamy, E.; Jones, C. C.; Fiat, D., Synthesis of oxygen-18 isotope labeled amino acids and dipeptides and its effects on carbon-13 NMR. *J. Labelled Compd. Rad.* **1987,** *24* (7), 773-778.
30. Pike, K. J.; Lemaitre, V.; Kukol, A.; Anupõld, T.; Samoson, A.; Howes, A. P.; Watts, A.; Smith, M. E.; Dupree, R., Solid-State 17O NMR of Amino Acids. *J. Phys. Chem. B* **2004,** *108* (26), 9256-9263.
31. Marecek, J.; Song, B.; Brewer, S.; Belyea, J.; Dyer, R. B.; Raleigh, D. P., A Simple and Economical Method for the Production of 13C,18O-Labeled Fmoc-Amino Acids with High Levels of Enrichment: Applications to Isotope-Edited IR Studies of Proteins. *Org. Lett.* **2007,** *9* (24), 4935-4937.
32. Seyfried, M. S.; Lauber, B. S.; Luedtke, N. W., Multiple-Turnover Isotopic Labeling of Fmoc- and Boc-Protected Amino Acids with Oxygen Isotopes. *Org. Lett.* **2010,** *12* (1), 104-106.
33. Kazuhiko, Y.; Toshio, Y.; Miwako, A.; Hiroshi, H.; Naoki, Y.; Yasuhiro, K., Solid-state 17O NMR in Biological Solids: Efficient 17O Labeling of Amino Acids and Peptides. *Chem. Lett.* **2007,** *36* (1), 192-193.
34. Klein, B. A.; Tkachuk, D. G.; Terskikh, V. V.; Michaelis, V. K., Expanding the NMR toolkit for biological solids: oxygen-17 enriched Fmoc-amino acids. *New J. Chem.* **2021,** *45* (28), 12384-12398.
35. James, S. L.; Adams, C. J.; Bolm, C.; Braga, D.; Collier, P.; Friščić, T.; Grepioni, F.; Harris, K. D.; Hyett, G.; Jones, W., Mechanochemistry: opportunities for new and cleaner synthesis. *Chem. Soc. Rev.* **2012,** *41* (1), 413-447.
36. Andersen, J.; Mack, J., Mechanochemistry and organic synthesis: from mystical to practical. *Green Chem.* **2018,** *20* (7), 1435-1443.
37. Boldyreva, E., Mechanochemistry of inorganic and organic systems: what is similar, what is different? *Chem. Soc. Rev.* **2013,** *42* (18), 7719-7738.
38. Howard, Joseph L.; Cao, Q.; Browne, D. L., Mechanochemistry as an emerging tool for molecular synthesis: what can it offer? *Chem. Sci.* **2018,** *9* (12), 3080-3094.
39. Howard, J. L.; Sagatov, Y.; Repusseau, L.; Schotten, C.; Browne, D. L., Controlling reactivity through liquid assisted grinding: the curious case of mechanochemical fluorination. *Green Chem.* **2017,** *19* (12), 2798-2802.
40. Trask, A. V.; Shan, N.; Motherwell, W. D. S.; Jones, W.; Feng, S.; Tan, R. B. H.; Carpenter, K. J., Selective polymorph transformation via solvent-drop grinding. *Chem. Commun.* **2005,** (7), 880-882.
41. Friščić, T.; Childs, S. L.; Rizvi, S. A. A.; Jones, W., The role of solvent in mechanochemical and sonochemical cocrystal formation: a solubility-based approach for predicting cocrystallisation outcome. *Cryst. Eng. Comm.* **2009,** *11* (3), 418-426.
42. Métro, T.-X.; Gervais, C.; Martinez, A.; Bonhomme, C.; Laurencin, D., Unleashing the Potential of 17O NMR Spectroscopy Using Mechanochemistry. *Angew. Chem., Int. Ed.* **2017,** *56* (24), 6803-6807.
43. Špačková, J.; Fabra, C.; Mittelette, S.; Gaillard, E.; Chen, C.-H.; Cazals, G.; Lebrun, A.; Sene, S.; Berthomieu, D.; Chen, K.; Gan, Z.; Gervais, C.; Métro, T.-X.; Laurencin, D., Unveiling the Structure and Reactivity of Fatty-Acid Based (Nano)materials Thanks to Efficient and Scalable 17O and 18O-Isotopic Labeling Schemes. *J. Am. Chem. Soc.* **2020,** *142* (50), 21068-21081.
44. Špačková, J.; Fabra, C.; Cazals, G.; Hubert-Roux, M.; Schmitz-Afonso, I.; Goldberga, I.; Berthomieu, D.; Lebrun, A.; Métro, T.-X.; Laurencin, D., Cost-efficient and user-friendly 17O/18O labeling procedures of fatty acids using mechanochemistry. *Chem. Comm.* **2021,** *57* (55), 6812-6815.
45. Unprotected free glycine was found to be difficult to ionize under the conditions used for the HRMS analyses, and the results obtained were therefore burdened with large error. In order to provide more reliable results, the enrichment level of glycine was estimated from its Fmoc-derivative supposing no loss of enrichment during Fmoc-protection (as was confirmed for other amino acids in this work).
46. Gerothanassis, I. P.; Hunston, R. N.; Lauterwein, J., 17O NMR chemical shifts of the twenty protein amino acids in aqueous solution. *Magn. Reson. Chem.* **1985,** *23* (8), 659-665.

47. Valentine, B.; Amour, T. S.; Walter, R.; Fiat, D., pH dependence of oxygen-17 chemical shifts and linewidths of l-alanine and glycine. *J. Magn. Reson.* **1980,** *38* (3), 413-418.
48. Mears, W. H., The Oxygen Exchange Reaction of Glycine Hydrochloride and Water. *J. Chem. Phys.* **1938,** *6* (5), 295-295.
49. Up to ~ 1 g/day of Fmoc-AA-O*H could be produced by performing two reactions simultaneously using MM400 mixer-mill equipped with two jars. For the instrument design see: https://www.retsch.com/products/milling/ball-mills/mixer-mill-mm-400/function-features/.
50. Ha, M.; Nader, S.; Pawsey, S.; Struppe, J.; Monette, M.; Mansy, S. S.; Boekhoven, J.; Michaelis, V. K., Racing toward Fast and Effective 17O Isotopic Labeling and Nuclear Magnetic Resonance Spectroscopy of N-Formyl-MLF-OH and Associated Building Blocks. *J. Phys. Chem. B* **2021,** *125* (43), 11916-11926.
51. Hung, I.; Keeler, E. G.; Mao, W.; Gor'kov, P. L.; Griffin, R.G.; Gan, Z., Residue-specific high-resolution 17O solid-state NMR of peptides: multidimensional indirect 1H detection and magic-angle spinning. *ChemRxiv* **2022**.
52. Current price of the 90% ^{17}O-enriched water from CortecNet was taken into account for the labeling cost calculation (15/06/2022, 2750 €/1mL). https://cortecnet.com/stable-isotope-for-nmr/materials/17o-water.html.
53. Information regarding the cost of ^{17}O-labeled water used was obtained via private communication with the authors. .
54. Plow, E. F.; Haas, T. A.; Zhang, L.; Loftus, J.; Smith, J. W., Ligand Binding to Integrins*. *J. Biol. Chem.* **2000,** *275* (29), 21785-21788.
55. Sun, Y.; Kang, C.; Liu, F.; Zhou, Y.; Luo, L.; Qiao, H., RGD Peptide-Based Target Drug Delivery of Doxorubicin Nanomedicine. *Drug Develop. Res.* **2017,** *78* (6), 283-291.
56. Meyers, S. R.; Grinstaff, M. W., Biocompatible and Bioactive Surface Modifications for Prolonged In Vivo Efficacy. *Chem. Rev.* **2012,** *112* (3), 1615-1632.
57. Samarah, L. Z.; Khattar, R.; Tran, T. H.; Stopka, S. A.; Brantner, C. A.; Parlanti, P.; Veličković, D.; Shaw, J. B.; Agtuca, B. J.; Stacey, G.; Paša-Tolić, L.; Tolić, N.; Anderton, C. R.; Vertes, A., Single-Cell Metabolic Profiling: Metabolite Formulas from Isotopic Fine Structures in Heterogeneous Plant Cell Populations. *Anal. Chem.* **2020,** *92* (10), 7289-7298.
58. Yamada, K.; Shimizu, T.; Ohki, S.; Yamazaki, T., A solid-state 17O NMR study of L -phenylalanine and L -valine hydrochlorides. *Magn. Reson. Chem.* **2008,** *46* (3), 226-234.
59. Stievano, L.; Tielens, F.; Lopes, I.; Folliet, N.; Gervais, C.; Costa, D.; Lambert, J.-F., Density Functional Theory Modeling and Calculation of NMR Parameters: An ab Initio Study of the Polymorphs of Bulk Glycine. *Cryst. Growth. Des.* **2010,** *10* (8), 3657-3667.
60. Veinberg, S. L.; Friedl, Z. W.; Harris, K. J.; O'Dell, L. A.; Schurko, R. W., Ultra-wideline 14N solid-state NMR as a method for differentiating polymorphs: glycine as a case study. *Cryst. Eng. Comm* **2015,** *17* (28), 5225-5236.
61. Iitaka, Y., The crystal structure of β-glycine. *Acta Crystallogr.* **1960,** *13* (1), 35-45.
62. Kazuhiko, Y.; Tadashi, S.; Toshio, Y.; Akira, S., A Solid-state 17O NMR Study of β-Glycine: High Sensitivity of 17O NMR Parameters to Hydrogen-bonding Interactions. *Chem. Lett.* **2008,** *37* (4), 472-473.
63. Wright, D. A.; Marsh, R. E., The crystal structure of l-lysine monohydrochloride dihydrate. *Acta Crystallogr.* **1962,** *15* (1), 54-64.
64. Lu, J.; Hung, I.; Brinkmann, A.; Gan, Z.; Kong, X.; Wu, G., Solid-State 17O NMR Reveals Hydrogen-Bonding Energetics: Not All Low-Barrier Hydrogen Bonds Are Strong. *Angew. Chem., Int. Ed.* **2017,** *56* (22), 6166-6170.
65. Chen, C.-H.; Goldberga, I.; Gaveau, P.; Mittelette, S.; Špačková, J.; Mullen, C.; Petit, I.; Métro, T.-X.; Alonso, B.; Gervais, C.; Laurencin, D., Looking into the dynamics of molecular crystals of ibuprofen and terephthalic acid using 17O and 2H nuclear magnetic resonance analyses. *Magn. Reson. Chem.* **2021,** *59* (9-10), 975-990.
66. Derissen, J.; Endeman, H.; Peerdeman, A., The crystal and molecular structure of L-aspartic acid. *Acta Crystallogr. B-Stru.* **1968,** *24* (10), 1349-1354.
67. Yamada, K.; Shimizu, T.; Tansho, M.; Nemoto, T.; Asanuma, M.; Yoshida, M.; Yamazaki, T.; Hirota, H., Solid-state 17O NMR study of the electric-field-gradient and chemical shielding tensors in polycrystalline amino acids. *Magn. Reson. Chem.* **2007,** *45* (7), 547-556.