



DIPLOMATHESES

Design and Synthesis of Isotope-Labeled Amino Acids and Precursors for Protein NMR Characterization

Conducted in partial fulfillment of the requirements for the academic
degree of a Diplom-Ingenieur (Dipl.-Ing.)

under the supervision of

Mag. Dr. Roman Lichtenecker, Privatdoz.

Department of Organic Chemistry, University of Vienna

Mag-Lab GmbH, 1030 Vienna

Assoc. Prof. Dipl.-Ing. Dr.techn. Michael Schnürch

E163 Institute of Applied Synthetic Chemistry

submitted at the

TU WIEN

Faculty of Technical Chemistry

submitted by

Sarah Kratzwald, BSc

Der experimentelle Teil dieser Arbeit wurde in der Zeit von April 2022 bis Oktober 2022 im Mag-Lab GmbH in 1030 Wien unter Supervision von Mag. Dr. Roman Lichtenecker, Privatdoz. durchgeführt.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass die vorliegende Arbeit nach den anerkannten Grundsätzen für wissenschaftliche Abhandlungen von mir selbstständig erstellt wurde. Alle verwendeten Hilfsmittel, insbesondere die zugrunde gelegte Literatur, sind in dieser Arbeit genannt und aufgelistet. Die aus den Quellen wörtlich entnommenen Stellen, sind als solche kenntlich gemacht.

Das Thema dieser Arbeit wurde von mir bisher weder im In- noch Ausland einer Beurteilerin/einem Beurteiler zur Begutachtung in irgendeiner Form als Prüfungsfach vorgelegt. Diese Arbeit stimmt mit der von den Begutachterinnen/Begutachtern beurteilte Arbeit überein.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Acknowledgements

I would like to express my deep gratitude to my supervisor, Dr. Lichtenecker, Privatdoz., for his constant support, encouragement, and guidance throughout the duration of my thesis. His expertise in bioorganic chemistry and his willingness to share his knowledge have been instrumental in the success of my thesis. I am grateful for the countless hours he spent discussing my research with me and for his unwavering support and encouragement. I appreciate the time he took to review my work and provide feedback.

I am also deeply grateful to Associate Prof. Dr.techn. Schnürch at TU Wien for his supervision from the TU Wien and for providing valuable insights throughout my research. Without him, the success of this thesis would not be possible.

I would also like to extend my sincere thanks to Matus and Thomas for their outstanding lab support, and their technical expertise, and for being such great colleagues. They were always willing to help, even in the most challenging times, and working with them was a pleasure. Special thanks go to Kathi, Gerald, and Sven for the fun and good company during my time in Mag-Lab as well as in the lab, as in the office. They made not only the research experience a memorable one and it was a pleasure having their company. Furthermore, I would like to thank Univ.-Prof. Dr. Robert Konrat for his scientific advice throughout this time.

I am also grateful for the persistent support of my boyfriend Philipp, who kept me laughing during the most challenging times and kept me going. He was my rock, confidant, and biggest supporter throughout this journey. I couldn't have done this without his love, support, and encouragement.

I would like to thank my family for their love, support, and encouragement throughout my academic journey. They have always been there for me, providing emotional and financial support, and I am eternally grateful for their love and support.

Finally, I would like to express my gratitude to everyone who has supported me in any way throughout this journey. Your help and encouragement have been invaluable, and I am deeply grateful for your support. This thesis is a culmination of the effort of many people, and I am honored to have had the opportunity to work with such a talented and dedicated team. Thank you all for making this journey such an enjoyable and rewarding experience.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Abstract

NMR spectroscopy has developed into a versatile tool to investigate proteins' structure, dynamics, and interaction in solution at an atomic resolution.^[1] To increase the sensitivity and resolution of protein NMR, labeling methods to incorporate stable isotopes such as ^2H , ^{13}C and ^{15}N into proteins have been developed and are constantly improved further. Protein labeling can be achieved by supplementing the growth media of the host cell with either metabolic amino acid precursors or isotopically labeled amino acids.^[2] Mainly, *E. coli* is the expression system of choice for protein production. Alternatively, mammalian cells, insect cells or yeast can be applied to incorporate isotopically labeled amino acids into desired proteins.

In the first part of the thesis, a synthetic route for N- α -Fmoc-O-(bis-dimethylamino phosphono)-[3,5- $^{13}\text{C}_2$ -2,6- D_2]-L-tyrosine (Fmoc-pTyr) has been developed. The procedure is based on a previously published synthetic strategy for introducing an isolated ^{13}C -H spin system into aromatic ring systems.^[3] This concept was further expanded by a Negishi cross-coupling reaction^[4], and a subsequent phosphorylation strategy^[5,6]. The target Fmoc-pTyr building block can be applied in solid-phase peptide synthesis (SPPS) with a subsequent chemical peptide ligation and allows for NMR analysis of pTyr-containing interaction partners.

The second part of the thesis deals with the synthesis of two metabolic precursors for valine/leucine and isoleucine.^[7]

Finally, novel isotopologues of the imidazole pyruvate tautomer have been introduced as metabolic precursors for selective histidine labeling. The imidazole side chain of histidine is highly abundant on protein interaction surfaces and plays an essential role in various enzyme mechanisms, such as catalytic triads.^[8] Previously, we could show that His-transaminase is highly reversible and effectively accepts imidazole-pyruvic acid as a substrate. Imidazole-pyruvic acid is the first intermediate in the minor histidine degradation pathway in *E. coli*.^[9] Based on our previous results, we could now introduce an isotopologue of imidazole-pyruvic acid containing a ^{15}N - ^{13}C - ^{15}N side-chain isotope pattern with low-cost sources of carbon-13 (formaldehyde) and nitrogen-15 (ammonium chloride).

Synthesis of Fmoc-pTyr and imidazole-pyruvic acid isotopologues have been developed and optimized using corresponding unlabeled compounds. Both target compounds were synthesized with reasonable yields and purity. They can be used in SPPS (Fmoc-pTyr) or cell-based protein overexpression (imidazole pyruvic acid), allowing for peptide or protein investigation using biomolecular NMR applications.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Kurzfassung

Die NMR-Spektroskopie ist heute eine der wichtigsten Methoden zur Untersuchung von Struktur, Dynamik und Interaktion von Proteinen in Lösung.^[1] Um sowohl Sensitivität als auch die Signalauflösung von Protein-NMR Spektren zu erhöhen, wurden Methoden entwickelt, um stabile Isotope wie ^2H , ^{13}C und ^{15}N in die Zielproteine einzubringen. Die Proteinmarkierung kann durch Zugabe von metabolischen Aminosäurevorläufern oder den isotopenmarkierten Aminosäuren selbst zum Wachstumsmedium des exprimierenden Systems erreicht werden.^[2] Hauptsächlich wird *E. coli* als Expressionssystem zur Proteinproduktion verwendet. Alternativ werden auch Zellkulturen aus Säugetieren, Insekten oder Hefen als Expressionssysteme zur Herstellung isotopenmarkierter Proteine genutzt.

Im ersten Teil der Arbeit wurde eine synthetische Route für N- α -Fmoc-O-(bisdimethylaminophosphono)-[3,5- $^{13}\text{C}_2$ -2,6- D_2]-L-tyrosine (Fmoc-pTyr) entwickelt. Das Verfahren basiert auf einer zuvor veröffentlichten Strategie zur Einführung eines isolierten ^{13}C -H Spin-Systems in aromatischen Ringen.^[3] Diese Synthese wurde in der vorliegenden Arbeit um eine Negishi-Kreuzkupplungsreaktion^[4] und einer anschließenden Phosphorylierung erweitert^[5,6]. Damit können isotopenmarkierte Bausteine hergestellt werden, die zur Synthese von Phosphotyrosin Peptiden mittels Festphasenpeptidsynthese (SPPS) herangezogen werden können. Diese Peptide können dann im Anschluss durch NMR-spektroskopische Methoden Information über Protein-Protein Wechselwirkungen liefern.

Zusätzlich wurden zwei metabolische Aminosäure-Vorstufen für Valin und Leucin, als auch Isoleucin nach literaturbekannter Vorschrift synthetisiert.^[7]

Schließlich wurde im letzten Teil der Arbeit eine Synthese eines Imidazolpyruvat Tautomeres erarbeitet, das die selektive Isotopenmarkierung von Histidin erlaubt. Die Imidazol Seitenkette dieser Aminosäure weist herausragende strukturelle Eigenschaften auf. Die Seitenkette tritt in zwei tautomeren Formen auf und spielt eine wesentliche Rolle in verschiedenen Proteinmechanismen und Enzymaktivitäten.^[8] Es konnte in der Vergangenheit gezeigt werden, dass die Histidin-Transaminase im Histidin-Abbauweg reversibel ist, und somit Imidazole-Pyruvat als Substrat akzeptiert.^[9] Darauf basierend konnten wir davon ein neues Histidine Isotopolog synthetisieren, welches eine ^{15}N - ^{13}C - ^{15}N Seitenkette enthält.

Die Synthese der isotopenmarkierten Zielsubstanzen wurde unter Verwendung entsprechender nicht-isotopenmarkierter Verbindungen entwickelt und optimiert. Beide Verbindungen konnten markiert hergestellt werden.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Contents

Acknowledgements	v
Abstract	vii
Zusammenfassung	ix
1 Introduction	1
1.1 Structure Determination of Proteins	1
1.2 Protein NMR	2
1.2.1 TROSY	3
1.2.2 NOESY	5
1.2.3 Multidimensional Spectroscopy	6
1.3 Protein-Ligand Interactions	7
1.3.1 Chemical Shift Perturbations	7
1.3.2 Saturation Transfer Difference NMR Spectroscopy	7
1.4 Protein Labeling	8
1.4.1 Cell-Based Methods	8
1.4.2 Cell-Free Methods	8
1.5 Labeling Strategies	9
1.5.1 Uniform Labeling	9
1.5.2 Selective Labeling	9
1.5.3 Stereo-array Isotopic Labeling (SAIL) Strategy	14
1.6 Tyrosine	15
1.6.1 Tyrosine Phosphorylation as Posttranslational Modification . .	15
1.6.2 Synthesis of Fmoc-phosphoTyrosine Building Blocks	16
1.6.3 Synthesis of Isotopically Labeled Tyrosine and Tyrosine Pre- cursors	17
2 Aims	20
3 Results and Discussion	22
3.1 Fmoc-phosphoTyrosine Synthesis	22
3.1.1 <i>N</i> - α -Fmoc- <i>O</i> -(bis-dimethylaminophosphono)-[3,5- $^{13}\text{C}_2$ -2,6- D_2]- L-tyrosine 11	22
3.1.2 <i>N</i> - α -Fmoc- <i>O</i> -benzyl-L-phosphotyrosine 18	26
3.2 Aliphatic Precursor Synthesis	29
3.3 Histidine Precursor Synthesis	31
4 Conclusion and Outlook	35

5	Experimental Section	36
5.1	General Remarks	36
5.2	<i>N</i> - α -Fmoc- <i>O</i> -(bis-dimethylamino-phosphono)-L-tyrosine 11	37
5.2.1	(<i>R</i>)-2-Benzoyloxycarbonylamino-3-iodo-propionic acid benzyl ester 2	37
5.2.2	[2,6- ¹³ C ₂]4-Nitrophenol 4	38
5.2.3	[2,6- ¹³ C ₂]4-Aminophenol 5	39
5.2.4	[2,6- ¹³ C ₂]3,5-Dideuterio-4-aminophenol 6	40
5.2.5	[2,6- ¹³ C ₂]3,5-Dideuterio-4-iodophenol 7	41
5.2.6	<i>N</i> -Carbobenzoxy[3,5- ¹³ C ₂ -2,6-D ₂]-L-tyrosine benzyl ester 8	42
5.2.7	<i>N</i> -Carbobenzoxy[3,5- ¹³ C ₂ -2,6-D ₂]-L-tyrosine[P(O)(NMe ₂) ₂] benzyl ester 9	43
5.2.8	[3,5- ¹³ C ₂ -2,6-D ₂]-L-Tyrosine[P(O)(NMe ₂) ₂] 10	44
5.2.9	<i>N</i> -Fmoc[3,5- ¹³ C ₂ -2,6-D ₂]-L-tyrosine[P(O)(NMe ₂) ₂] 11	45
5.3	<i>N</i> - α -Fmoc- <i>O</i> -benzyl-L-phosphotyrosine 18	46
5.3.1	<i>N</i> -Fmoc-L-serine 13	46
5.3.2	<i>N</i> -Fmoc-L-serine methyl ester 14	47
5.3.3	<i>N</i> -Fmoc-3-iodo-L-alanine methyl ester 15	48
5.3.4	4-Iodophenol 7b	49
5.3.5	<i>N</i> -Fmoc-L-tyrosine methyl ester 16	50
5.3.6	<i>N</i> -Fmoc-L-tyrosine 17	51
5.3.7	<i>N</i> - α -Fmoc- <i>O</i> -benzyl-L-phosphotyrosine 18	52
5.4	Isoleucine Precursor 24	53
5.4.1	<i>tert</i> -Butyl 2-(hydroxymethyl)acrylate 20	53
5.4.2	<i>tert</i> -Butyl 2-(bromomethyl)acrylate 21	54
5.4.3	[4- ¹³ C ₂] <i>tert</i> -Butyl 2-methylenebutanoate 22	55
5.4.4	[4- ¹³ C ₂] <i>tert</i> -Butyl 2-ketobutanoate 23	56
5.4.5	[4- ¹³ C ₂ , 3-D ₂] 2-Ketobutanoic acid 24	57
5.5	Valine and Leucine Precursor 30	58
5.5.1	[4-CD ₃] <i>tert</i> -Butyl 2-methylenebutanoate 25	58
5.5.2	[4-CD ₃] <i>tert</i> -Butyl 2-ketobutanoate 26	59
5.5.3	[4-CD ₃] <i>tert</i> -Butyl 2-(2,2-dimethylhydrazono)butanoate 27	60
5.5.4	[4- ¹³ C; 4-CD ₃] <i>tert</i> -Butyl 2-(2,2-dimethylhydrazono)-3-methylbutanoate 28	61
5.5.5	[4- ¹³ C; 4-CD ₃] <i>tert</i> -Butyl 3-methyl-2-ketobutanoate 29	62
5.5.6	[4- ¹³ C; 4-CD ₃] Ketoisovaleric acid 30	63
5.6	Histidine Precursor	64
5.6.1	Ethyl 2-(diethoxyphosphoryl)-2-hydroxyacetate 32	64

5.6.2	Ethyl 2-((<i>tert</i> -butyldimethylsilyl)oxy)-2-(diethoxyphosphoryl)acetate 33	65
5.6.3	[1,3- ¹⁵ N, 2- ¹³ C](Imidazol-1 <i>H</i> -5-yl)methanol 34	66
5.6.4	1 <i>H</i> -Imidazole-5-carbaldehyde 35	67
5.6.5	<i>tert</i> -Butyl 4-formyl-1 <i>H</i> -imidazole-1-carboxylate 37a Route A .	68
5.6.6	<i>tert</i> -Butyl-[1,3- ¹⁵ N, 2- ¹³ C]imidazole-4-yl-methanol 36	69
5.6.7	<i>tert</i> -Butyl-[1,3- ¹⁵ N, 2- ¹³ C]imidazole-1-carboxylate 37 Route B	70
5.6.8	<i>tert</i> -Butyl-4-(2-((<i>tert</i> -butyldimethylsilyl)oxy)-3-ethoxy-3-oxoprop- 1-en-1-yl)-[1,3- ¹⁵ N, 2- ¹³ C]imidazole-1-carboxylate 38	71
5.6.9	[1,3- ¹⁵ N, 2- ¹³ C]4-(2-Carboxy-2-hydroxyvinyl)-1 <i>H</i> -imidazolium chloride 39	72
	Abbreviations	i
	List of Figures	iii
	List of Schemes	iv
	List of Tables	v
	References	vi
	NMR Spectra	xi



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

1 Introduction

1.1 Structure Determination of Proteins

Proteins are biologically relevant macromolecules containing at least one chain of amino acids. They hold a huge variety of functions within organisms. Their structure can be described using four levels, namely primary, secondary, tertiary, and quaternary structure. Hence, the primary structure solely describes the amino acid sequence. Whereas, the secondary structure is determined by hydrogen bonds within local structures. Well-known examples of secondary structures are α -helices, β -sheets, and turns. The tertiary structure is responsible for the basic functions of a protein. It is formed by the so-called hydrophobic core, which is stabilized by side chain interactions such as hydrogen bonds, Van-der-Waals- and hydrophobic interactions, as well as covalent binding such as disulfide bonds. Multiple tertiary structures by several protein molecules add up to quaternary structures leading to the formation of dimers, trimers, tetramers, et cetera.^[10]

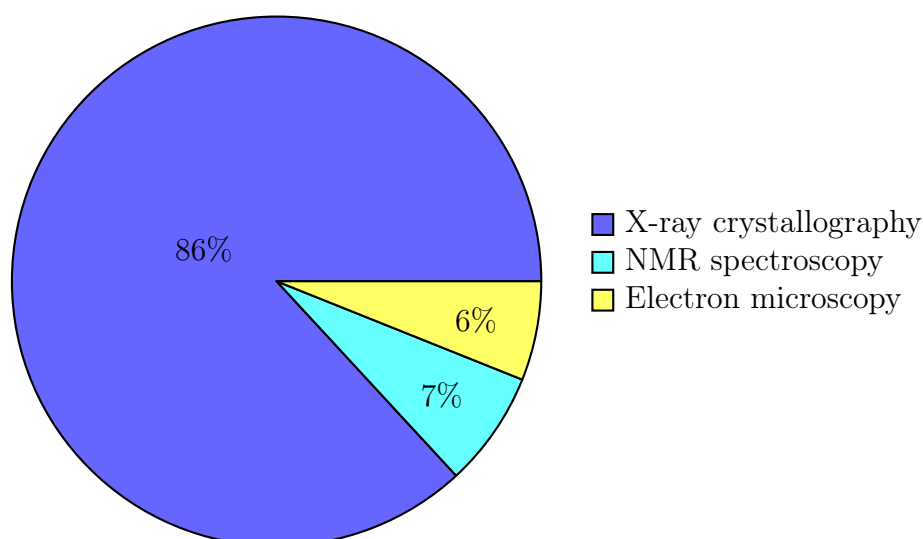


Figure 1: Distribution of experimental methods used for protein structure determination of proteins.^[11]

Knowledge about protein structure is key to understanding its function. Proteins were first able to be determined by their structure via X-ray crystallography in the 1950ies.^[12] Up to this date, there are 192095 biological macromolecules stored in the Protein Data Bank as of July 2022. A total of 86% of these structures have been solved by X-ray crystallography, 7% by Nuclear Magnetic Resonance (NMR) spectroscopy, and 6% by electron microscopy (EM) as depicted in Figure 1.^[11] Doubtlessly, X-ray crystallization is the dominating approach for molecular structure determination at an atomic resolution. Despite its dominance, X-ray crystallography comes with limitations, such as the need to form X-ray-grade crys-

tals and the chance of losing the protein's natural conformation.^[13] Recently, the cryo-EM technique has become a rising method to aim for near-atomic-resolution protein structures without the need for crystallization.^[14] NMR spectroscopy has become a complementary alternative to X-ray crystallization. It offers structural models of proteins without the need for their crystallization. However, even if the X-ray structure of a protein is available, a solution structure obtained via NMR may always be relevant. It provides protein structures almost at physiological conditions and thereby precious information about the proteins and their biological function can be gained.^[15,16] Nevertheless, apart from these positive attributes of protein NMR applications, there are also drawbacks. It comes with limitations such as line broadening as a result of more efficient transverse relaxation, unfavorable signal-to-noise ratios, and a great number of signals leading to complex spectra.^[17]

1.2 Protein NMR

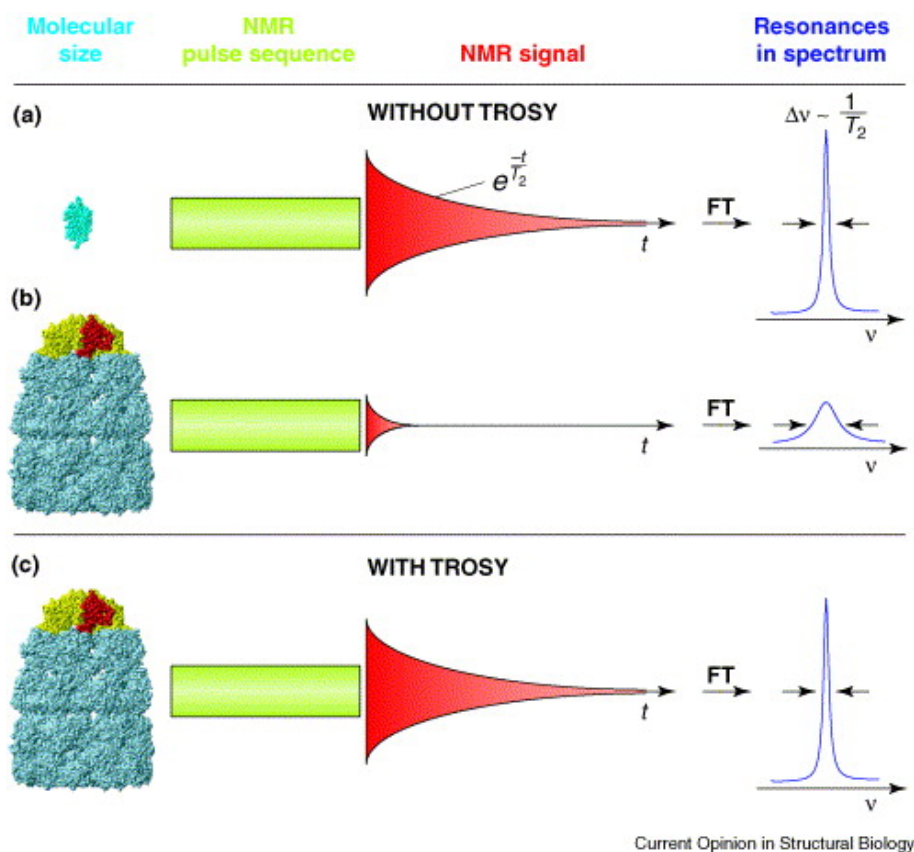
NMR spectroscopy was discovered in 1946 by Felix Bloch and Edward Purcell.^[18,19] The basic principle of NMR spectroscopy is based on the fact that nuclei with a spin of $1/2$ such as ^1H , ^{13}C or ^{15}N , can take on higher energy spin states within a magnetic field by radiation. The frequency at which the nucleus absorbs energy and therefore its NMR properties are determined by its chemical environment within a molecule.^[20]

NMR spectroscopy is not a stand-alone field within structural protein biology. Today, it is preferably used to gain information about kinetics and dynamics, interactions, and conformational changes of proteins.^[2] Protein NMR as an independent field of structural biology was established through the introduction of ^1H homonuclear two-dimensional NMR (2D-NMR) pioneered by the group of Karl Wüthrich and Richard R. Ernst from the ETH Zurich.^[2,21]

As mentioned in section 1.1 protein NMR does not only provide precious information, but it also comes with limitations. As the size of the protein increases, the width of the signal broadens and the resolution decreases.^[16,22] Moreover, NMR is not a sensitive technique regarding the sample amount which is required. Improvements have been achieved by the introduction of a pulse Fourier Transform NMR^[23] and the continuous development of higher field spectrometers which allow for a better resolution. Additionally, cryogenically cooled samples have led to an improved signal-to-noise ratio.^[24] In addition to all the methods mentioned, the incorporation of stable isotopes (^2H , ^{13}C and ^{15}N) into proteins is essential to overcome the limitations of the NMR analysis of large protein complexes.^[2,16,22,25]

A wide range of different NMR experiments with diverse pulse sequences are available today to elucidate protein structure, dynamics, and interaction.^[26]

1.2.1 TROSY



Current Opinion in Structural Biology

Figure 2: NMR spectroscopy with small and large molecules in solution. **A)** The NMR signal obtained from small molecules in solution relaxes slowly; it has a long transverse relaxation time (T_2). A large T_2 value translates into narrow line widths ($\Delta\nu$) in the NMR spectrum after Fourier transformation (FT) of the NMR signal. **B)** By contrast, for larger molecules, the decay of the NMR signal is faster (T_2 is smaller). This results in both in a weaker signal measured after the NMR pulse sequence and in broad lines in the spectra. **C)** Using TROSY, the transverse relaxation can be substantially reduced, which results in improved spectral resolution and improved sensitivity for large molecules; *taken from*^[27]

As mentioned, protein structure determination using NMR spectroscopy comes along with the issues of signal overlapping in the spectra due to a large number of residues. Furthermore, the relaxation time of large molecules is shorter. That in turn leads to line broadening and a loss of sensitivity (see Figure 2 (a) and (b)). Certain isotope patterns can reduce signal overlapping. The major source of relaxation in proteins is a large number of hydrogen atoms. One way to reduce the transverse relaxation is the replacement of ^1H with ^2H .^[27] The cause of this is that the gyromagnetic constant γ of ^2H is lower than the one of ^1H ($\gamma_{\text{H}}/\gamma_{\text{D}} \approx 6.5$).^[28]

The introduction of Transverse relaxation-optimized spectroscopy (TROSY) has allowed for further reduction of line widths. Thus, less line broadening and better sensitivity can be realized (see Figure 2 (c)).^[27] Hence, TROSY together with iso-

tope labeling techniques has allowed studying biomolecules up to 1000 kDa in size via NMR spectroscopy in solution.^[27]

In NMR spectroscopy, the signal decays with $e^{(-t/T_2)}$. After the Fourier Transformation of the signal, a spectrum with the corresponding resonance lines ready for analysis is obtained. T_2 is inversely proportional to the line width of the resonance lines. In other words, the shorter T_2 , the broader the line will be, whereas T_2 is dependent on the size of the molecule. The larger the molecule, the shorter T_2 will be.

The weaker NMR signal in large molecules is due to the fact that relaxation is taking place when data acquisition occurs as well as during the pulse sequence. When TROSY is applied, relaxation while the pulse sequence and data acquisition are reduced, resulting in less signal broadening in the NMR spectra.^[27]

One requirement of the TROSY effect is the presence of two different interfering relaxation mechanisms. The ^{15}N and ^1H underlie two interfering relaxation mechanisms, namely a dipole-dipole (DD) relaxation between them and the Chemical Shift Anisotropy (CSA) of the protons. Hence, in the spectra of a large protein, the ^{15}N signal consists of two lines with different line widths referring to the relaxation interference. Decoupling is usually applied in conventional NMR spectra of these two lines to obtain one. Though, this comes along with the averaging of the relaxation rate. In TROSY the faster relaxation resonance is eliminated and thereby the slower one is exclusively selected. The CSA increases with the strength of the magnetic field, whereas the DD coupling is independent of the magnetic field strength.

Hence, the optimal TROSY effect may be achieved by the proper magnetic field strength. Regarding the example of the ^{15}N amide proton the field strength lies about 23.T.^[27] Thus, with the development of TROSY, proteins in solution with <100 kDa can be studied via NMR spectroscopy.^[29]

1.2.2 NOESY

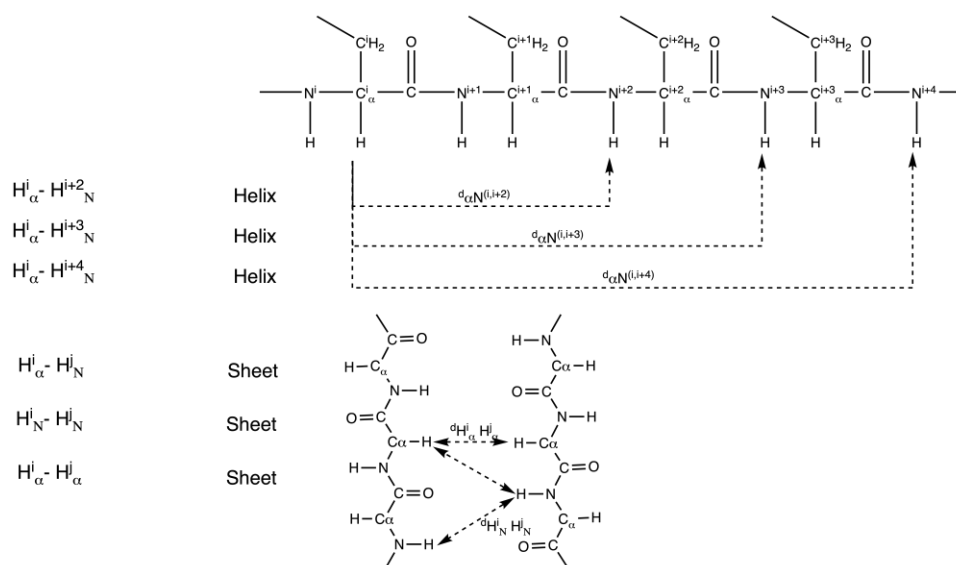


Figure 3: Arrows indicate NOE interactions that can be observed in polypeptide chains.

Nuclear Overhauser Effect Spectroscopy (NOESY) is a key experiment in protein NMR spectroscopy to acquire distance information for protein structure calculation.

The Nuclear Overhauser Effect (NOE) is based on dipolar interaction between protons through space, but also on scalar couplings. Thereby, information about distance and torsion angle constraints can be elucidated. The first 90° pulse leads to a population difference whereas magnetization is redirected into the plane. During the mixing delay, τ_m cross relaxation occurs, due to the difference from the z -magnetization to M_0 , whereas the population difference is changed and nearby nuclei undergo z magnetization as a result. The last 90° pulse then transfers these of the nearby nuclei into x' - y' magnetization and thereby the free induction decay (FID) gets induced.^[30]

The cross-peaks in a NOESY spectrum correspond to a strong NOE, meaning the two coupling entities are in close proximity, giving rise to through-space coupling. The corresponding signal intensity gives information about the proton distances ranging from the lower limit of 2 \AA corresponding to two hydrogen radii to the upper limit of 5 \AA . Usually, in a folded protein there are quite a few protons that have an internuclear distance of up to 5 \AA which may lead to crowded spectra.^[16] However, these long-range NOE signals in a NOESY spectrum give crucial information about the backbone fold of the protein.

1.2.3 Multidimensional Spectroscopy

Triple resonance NMR spectroscopy is a versatile tool to gain a deeper understanding of protein structure and dynamics. Hence, proteins with uniformly labeled carbon-13 and nitrogen-15 allow for the protein's backbone assignment.

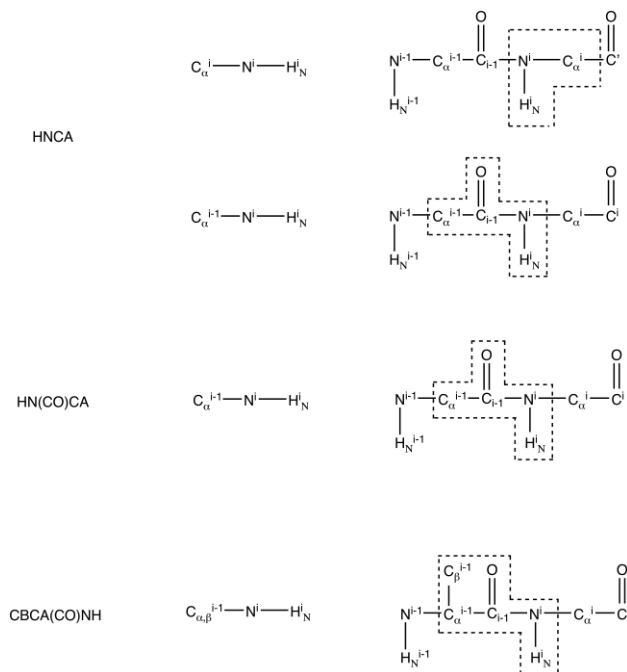


Figure 4: Magnetization transfer paths of the following 3D-NMR methods: HNCA, HN(CO)CA and CBCA(CO)NH; *adapted from*^[31]

Magnetization transfers of a few selected 3D-NMR are illustrated in Figure 4. The magnetization in an HNCA spectrum is passed from the proton (H_N^i) to the nitrogen-15 (N^i) and then from that to the α -carbon-13 ($C_{\alpha i}$), and then it retraces its path to the H_N^i for detection. However, as depicted in Figure 4, the nitrogen (N^i) is additionally coupled to the carbonyl carbon (C_{i-1}) and α -carbon (C_{α}^{i-1}) of the next residue. Thus, both magnetization transfer pathways can be seen in the spectrum, whereas the signal intensity of the directly coupled α -carbon ($C_{\alpha i}$) is substantial.^[31] Hence, an HNCA experiment displays the $^{15}\text{N}-C_{\alpha}$ coupling, which is rather small and lies within a range of 8-12 Hz. Furthermore, connectivity information can be gained via the weak $^2J_{NC\alpha}$ coupling between the ^{15}N and C_{α} of the preceding amino acids. The power of this approach was first shown in the case of calmodulin, which used to be a rather challenging protein regarding its high content of α -helices. Within an HNCA experiment, about half of Calmodulin's amino acids provided this connectivity information demonstrating its useful applications.^[32]

Other magnetization transfers are depicted in Figure 4. The CBCA(CO)NH experiment (bottom in Figure 4) displays a standard method in protein backbone

assignment.

Whereas an HN(CO)CA experiment (middle in Figure 4) is usually used in combination with an HNCA experiment for the protein backbone assignment.^[31]

1.3 Protein-Ligand Interactions

Protein NMR techniques can be crucial in drug discovery as it allows for exploring the interactions between proteins and ligands at atomic resolution.^[33–36]

1.3.1 Chemical Shift Perturbations

A ligand-protein interaction usually changes the position of the NMR signals derived from the nuclei involved. This so-called Chemical Shift Perturbations (CSP) depict the changes in a protein's chemical shifts upon the ligand's non-covalent binding. Thus, the protein spectrum (typically a heteronuclear single quantum correlation (HSQC) spectrum) is recorded without the addition of the ligand at first. Then the ligand is titrated against the protein. In the presence of protein-ligand interactions, specific crosspeaks will shift from their original position. This information can then be used to determine the location of the binding site but also allow the calculation of the equilibrium dissociation constant (K_D).^[36,37] Figure 5 demonstrates the application of CSPs for investigating protein-ligand interactions. Figure 5 shows the overlay of two ^1H - ^{13}C HSQC spectra of the BRD4-BD1 protein that has been selectively labeled with tryptophan. Upon binding of a suitable ligand, significant CSPs are observed.^[38]

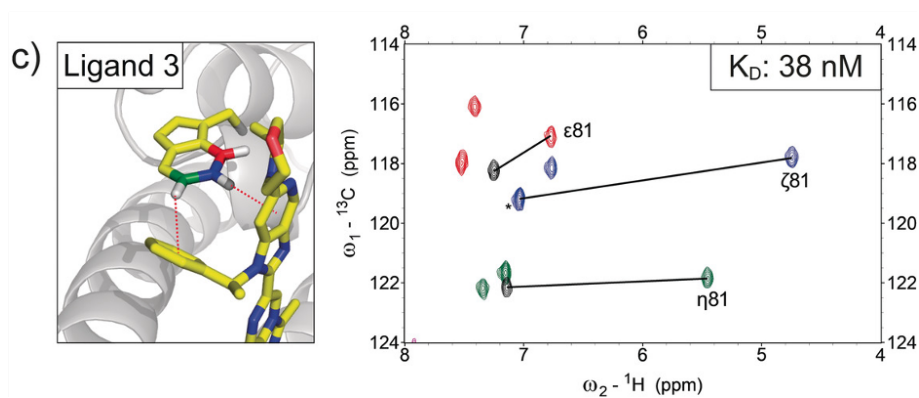


Figure 5: CSP for the BRD4-BD1 protein upon binding to a ligand; *taken from*^[38]

1.3.2 Saturation Transfer Difference NMR Spectroscopy

Another important tool to investigate protein-ligand interactions displays the Saturation Transfer Difference (STD) NMR. STD-NMR allows for the screening of ligands to proteins based on the Nuclear Overhauser Effect (NOE).^[36,39] In an STD-NMR experiment, certain protein nuclei are selectively irradiated. Small molecules (the

ligands) remain unaffected. Then, the saturation spreads through the protein's protons via spin diffusion. Through intermolecular NOEs, the saturation can be spread to any bound ligand. Thus, protons of the ligand that are near the binding side will undergo a larger saturation transfer. Hence, these protons will show more intense STD signals.^[36,39,40]

1.4 Protein Labeling

In NMR spectroscopy, there is a need for NMR active atoms, usually with a spin $l = 1/2$. Therefore, it is necessary to enrich samples with ^{13}C - and ^{15}N -atoms.^[41] On the other hand, labeling with ^2H ($l = 1$) plays a crucial role in protein NMR to improve the quality of the resulting spectra by reducing peak numbers and line widths.^[42]

1.4.1 Cell-Based Methods

Bacterial or yeast microorganisms are routinely used as host cells for protein overexpression. The cells are transformed with a plasmid containing the gene for the protein of interest, allowing for its transcription and translation at high levels. Mainly, *E. coli* is the host organism of choice since it comes with the advantages of easy handling and the availability of various commercial vectors.^[2,43] However, proteins often degrade or become toxic to the host cells. Furthermore, post-translational modifications like phosphorylation, methylation, glycosylation, etc. usually don't take place in procaryotic cells.^[2]

1.4.2 Cell-Free Methods

Cell-free (CF) expression methods have emerged as an alternative to cell-based methods for protein expression. Their productivity has risen over the past decades, in which a 1 ml reaction mixture can synthesize approximately 1 mg of protein. Viable hosts such as *E. coli* or yeast are no longer required in CF systems, which brings the advantage of eliminating the toxic effects of protein overexpression to the host cells. Furthermore, membrane proteins, which can be troublesome to express in living systems, can be expressed via CF methods.^[2] Along with transcription/translation systems extracted from host cells (*E. coli* or yeast), nucleotides, amino acids, and energy sources have to be supplied for the *in vitro* production of proteins.^[2,44] Labeling strategies for CF systems include the SAIL strategy, but also the incorporation of unnatural amino acids.^[2,43]

1.5 Labeling Strategies

1.5.1 Uniform Labeling

In uniform labeling, all nuclei of a specific atom (carbon, nitrogen or hydrogen) are replaced by their respective isotopes. Labeled substrates, usually ^{13}C -glucose and ^{15}N -ammonium salts, are provided in the minimal growth medium of the host organism allowing the conversion into isotope-labeled amino acids in the metabolic pathway.^[43,45] Nitrogen participates in the peptide bond formation in proteins. Thus, Nitrogen-15 labeling is essential for the analysis of the protein backbone.^[46] However, sensitivity can be lost through the dipolar ^1H - ^1H and heteronuclear (^1H - ^{13}C and ^1H - ^{15}N) relaxation pathways.^[43] Deuteration in uniform protein is achieved by simply supplementing minimal growth media with D_2O .^[42] The nucleus of ^2H has a large quadrupolar moment and its gyromagnetic ratio is about a sixth compared to ^1H .^[47] Hence, relaxation pathways leading to sensitivity loss are eliminated and informational content is reduced.^[43,48] However, media solely based on D_2O has been shown to hinder cell function and growth, thus resulting in low protein yields.^[49]

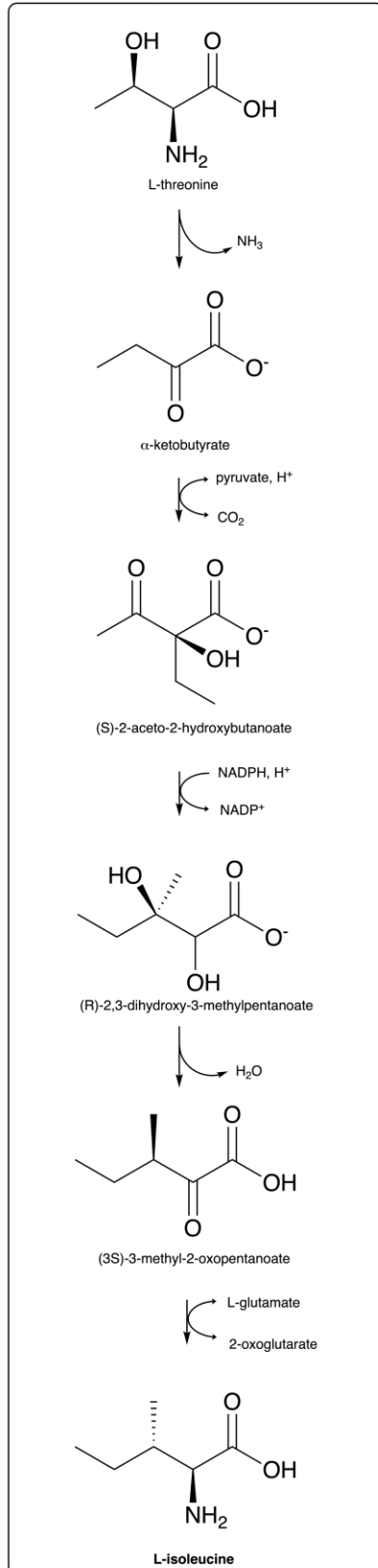
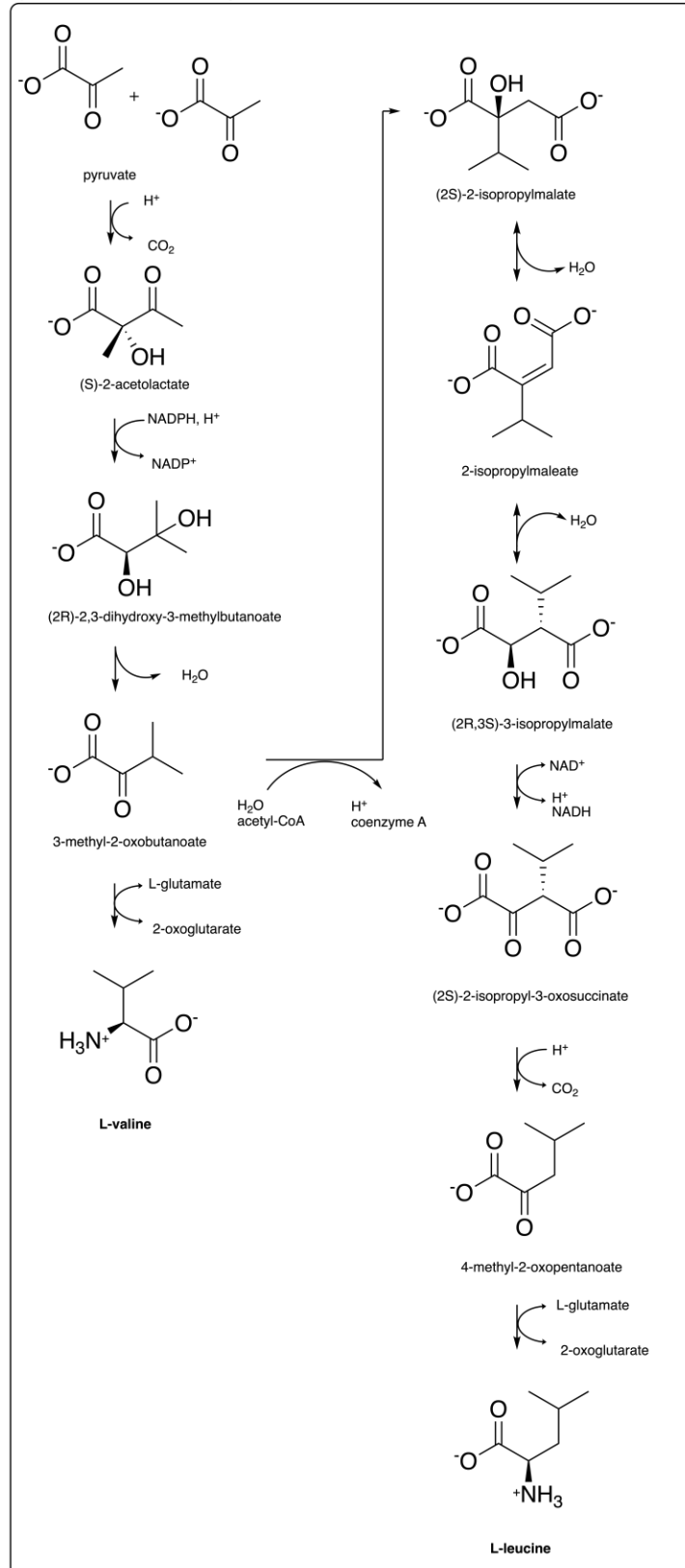
1.5.2 Selective Labeling

Selective isotope labeling describes the incorporation of isotopes into the protein at specific sites along the polypeptide chain. This approach takes place in cell-based labeling techniques, where the protein-expressing organisms (*E. coli*, but also yeast or insect cell lines) is grown in media, which is supplemented with a suitable isotope of the corresponding metabolic amino acid precursor. When the host organism takes up the metabolite, it is converted into the desired amino acid within its metabolic pathways.^[50-52] The host organism's metabolism must be well-defined and studied. The chances of cross-labeling are high when using early metabolic intermediates as isotopically labeled precursors. Thus, isotope patterns may not be selective. Therefore, auxotrophic strains or the addition of compounds inhibiting certain metabolic pathways have to be considered to avoid cross-labeling in this case. Alternatively, late metabolic precursors selective for a corresponding amino acid may be used.^[53]

Aliphatic Amino Acid Labeling

α -Ketoacids have been introduced as precursors for isoleucine, valine, and leucine^[7] and are highly selective in cell-based protein overexpression.^[54,55] Scheme 1 depicts the biosynthetic pathway of the amino acids isoleucine (Scheme 1 **A**) as well as valine and leucine (Scheme 1 **B**). The precursor chosen for isoleucine labeling is α -ketobutyrate. A synthetic procedure for synthesizing an isotopologue incorporated into the protein at a level of about 90% has been described.^[7] These high incorporation levels revealed that alternative biosynthetic pathways for isoleucine are irrelevant apart from the depicted one.^[56]

Using these α -ketoacids brings the advantage for organic synthesis that no stereocenter has to be introduced compared to the isotopologues of the actual amino acids. Furthermore, nitrogen-15 can be introduced by adding $^{15}\text{NH}_4\text{Cl}$ to the media. An isotopologue of α -ketoisovalerate is the precursor for valine and leucine since it is a metabolite in the biosynthetic pathway of both amino acids (see 3-methyl-2-oxobutanoate in Scheme 1). However, it has been shown that it is possible to exclusively isotopically label valine by adding unlabeled 2-ketoisocaproate (4-methyl-2-oxopentanoate). In this case, the metabolic pathway of labeled 3-methyl-2-oxobutanoate to leucine is blocked and this precursor will exclusively be metabolized to valine residues.^[55]

A) Isoleucine Biosynthesis**B) Valine and Leucine Biosynthesis**

Scheme 1: Biosynthesis of the aliphatic amino acids in *E. coli*: **A)** isoleucine biosynthesis; **B)** valine and leucine biosynthesis; adapted from^[57]

Apart from these mentioned aliphatic amino acid precursors, there has been

rising interest in introducing and developing isotope-labeled aromatic amino acids and corresponding precursors.^[58]

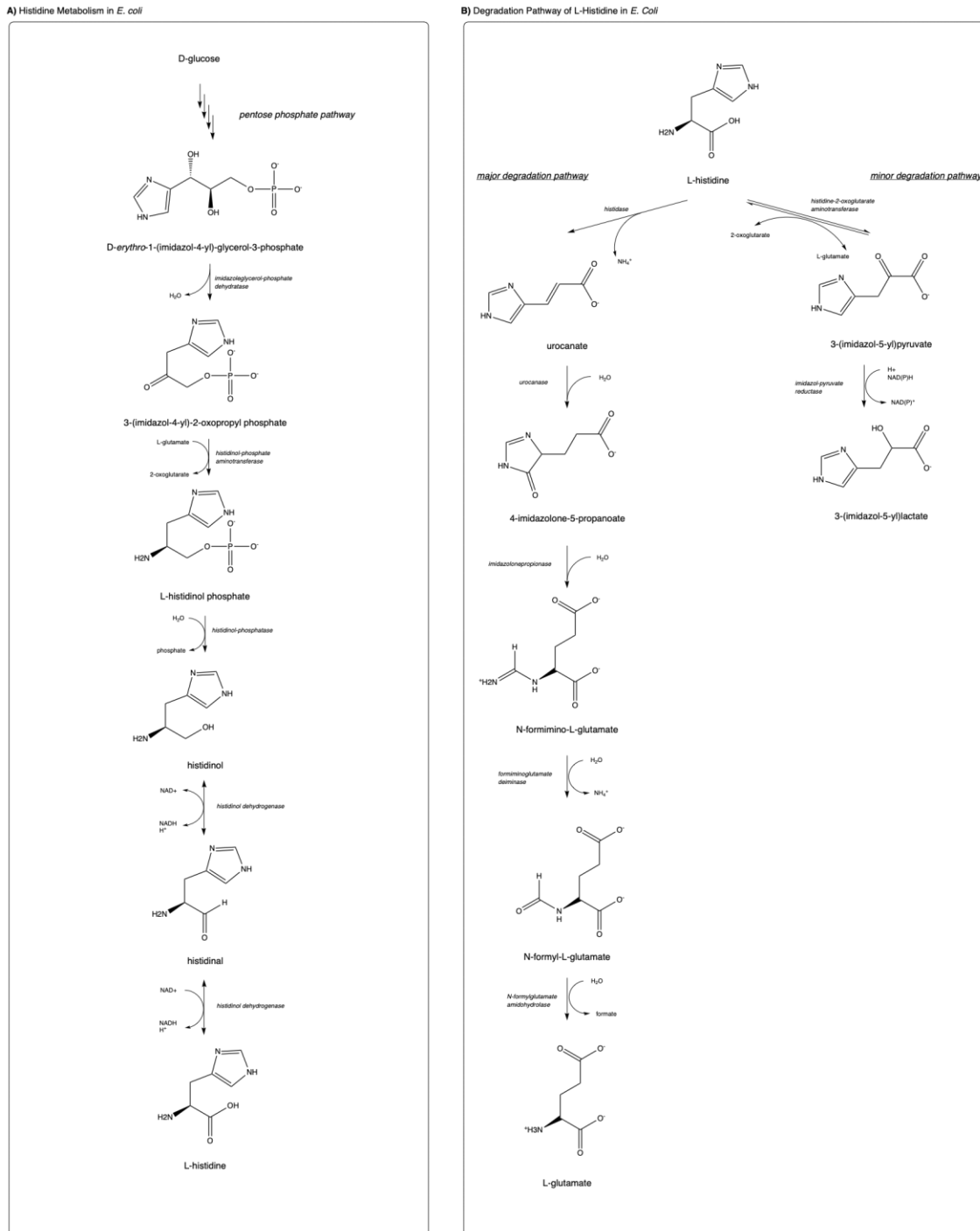
Histidine Labeling

Histidine is an α -amino-acid with an imidazole side chain, whereas it is showing outstanding properties due to its unique molecular structure.^[59,60] Its aromatic imidazole side chain with a $pK_a = 6.5$ in aqueous solution appears in two neutral tautomeric forms, as well as in a protonated form.^[61] Due to its enormous versatility, it plays many roles in molecular interactions. It can form cation- π interactions via its aromatic moiety with either metal cations or protonated forms of other amino acids such as lysine.^[62] Furthermore, histidine can function as a cation when protonated.^[63] With its aromatic properties of the imidazole moiety, histidine can take part in π - π interactions with other aromatic amino acids like tyrosine, phenylalanine, and tryptophan.^[64] The lone electron pair at the basic histidine nitrogen can act as a ligand to coordinate metals such as Zn^{2+} and Ca^{2+} .^[65] Additionally, the basic nitrogen can act as a hydrogen-bond acceptor, and the hydrogen atom on the imidazole moiety can act as a hydrogen-bond donor. Last but not least, histidine is also capable of forming hydrogen- π interactions.^[8,66]

The biosynthetic pathway of histidine in *E. coli* has been studied extensively. As depicted in Scheme 2, histidine is synthesized from phosphoribosyl pyrophosphate via the pentose phosphate pathway. Histidine is mainly degraded via deamination by the histidine deaminase to urocanate.^[57] In its minor degradation pathway, histidine is converted to 3-(imidazole-5-yl)pyruvic acid catalyzed by histidine 2-oxoglutarate transaminase as depicted in Scheme 2. Previously, it was shown that histidine transaminase is highly reversible and effectively accepts imidazole-pyruvic acid as a substrate. Therefore, a histidine precursor containing a carbon-13 on ϵ_1 suitable for cell-based expression was introduced. The synthesis was proven to be efficient and carbon-13 is introduced by a ^{13}C -formaldehyde solution as a relatively low-cost isotope source. Thus, the imidazole moiety is introduced within the first step as 4-hydroxymethyl-imidazole. Next, oxidation to the corresponding aldehyde using MnO_2 and subsequent Boc-protection were performed. That compound was then used in a Horner-Wadsworth-Emmons reaction yielding the TBDMS-protected imidazole-pyruvic acid, and the corresponding final precursor was obtained via an acidic deprotection.^[9]

Yet, not only a metabolic precursor for histidine overexpression in *E. coli*, but also a procedure for the introduction of $2'$ - ^{13}C -L-histidine has been described in literature. Thereby, carbon-13 is introduced at the first step of the reaction sequence via ^{13}C -thiocyanate.^[67]

In order to gain a deeper understanding of the overall role of histidine in protein functions and interactions, further isotopologues of the amino acids and their corresponding precursors are required.


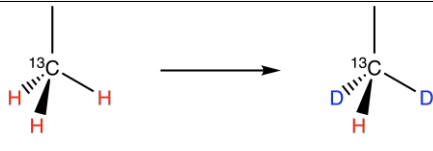
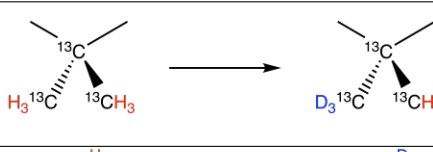
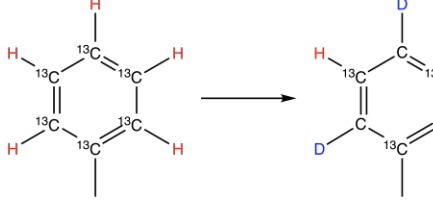


Scheme 2: Histidine metabolism in *E. coli*: **A)** showing the histidine metabolism from the pentose phosphate pathway; **B)** and the major and minor degradation pathway in *E. coli* with the reversible conversion of histidine to imidazole-pyruvic acid. adapted from^[9,57]

1.5.3 SAIL Strategy

Another labeling strategy can be achieved via a chemical or enzymatic synthesis of amino acids and a subsequent cell-free protein expression. The concept of the Stereo-array Isotopic Labeling (SAIL) aims for clearer NMR spectra without losing quality and information in NMR structure elucidation. The SAIL approach is based on certain isotope patterns, which are depicted in Table 1.

Table 1: SAIL Strategy

A)	^1H is stereo-selectively replaced by ^2H in methylene groups	
B)	two ^1H are replaced by two ^2H in each methyl group	
C)	two methyl groups are replaced by one $^{-13}\text{C}(^2\text{H})_3$ and one $^{-13}\text{C}^1\text{H}(^2\text{H})_2$ in prochiral methyl groups	
D)	alternating $^{13}\text{C}-^1\text{H}$ and $^{12}\text{C}-^2\text{H}$ moieties in six-membered aromatic ring systems	

Via these labeling patterns, on the one hand, through-bond connectivity information for backbone and side chain assignments remains maintained. On the other hand, signal-to-noise ratios are increased. Furthermore, coupling measurements are simplified, major sources for spin diffusion are eliminated, and the accuracy for inter-proton distance measurements is improved.^[68]

1.6 Tyrosine

L-Tyrosine or 4-hydroxyphenylalanine is a proteogenic, aromatic amino acid. Like other aromatic amino acids, it is an important part of forming the hydrophobic core of proteins, where aromatic-aromatic interactions are formed.^[69]

1.6.1 Tyrosine Phosphorylation as Posttranslational Modification

Posttranslational modifications (PTMs) describe the modification a protein undergoes after it has been translated. Examples of common PTMs are acetylations, phosphorylation, glycosylations, or methylations. These modifications are covalently bound to one or more amino acids of the protein, reversibly or irreversibly, and affect the structure and dynamics of a protein.^[70] Phosphorylation as PTM was first reported in 1906 by Phoebus Levene, who discovered the phosphate group in the protein vitellin.^[71] Another 20 years later, the first enzymatic protein phosphorylation was reported. Today, Posttranslational modification (PTM) is regarded as a key process to regulate enzyme activities and many other processes within the cell.^[72]

Phosphorylation as PTM can occur on serine, threonine, tyrosine, histidine, proline, arginine, aspartic acid, and cysteine. However, serine, threonine, and tyrosine are the amino acids most prone to undergo phosphorylation.^[70,73,74] The phosphate group as PTM is transferred by kinases from adenosine triphosphate to receptor residues. Vice versa, different phosphatases catalyze the removal of the phosphate group.^[70,75] A switch between phosphorylated and non-phosphorylated state may change the activity but also the accessibility of the protein.^[76] Usually, a protein's function change occurs either through binding to interaction domains or allostery. Therefore, protein phosphorylation is a key function of replication, transcription, apoptosis, and cell metabolism.^[70]

Introducing a phosphate with its two anionic groups into a protein changes the chemical properties and affects its structure and local environment. With the phosphate group, the possibility of forming hydrogen bonds or salt bridges arises, which can change a protein's activity or give rise to new binding sites.^[74,77] Compared to serine or threonine, the phosphate group on tyrosine is further apart from the peptide backbone. Additionally, tyrosine's phenolic ring offers further binding energy for phosphospecific binding domains such as π - or hydrophobic bond-ring interactions.^[75]

There are proteins containing phosphorylated Tyrosine (pTyr) binding modules such as Src homology 2 (SH2) and phosphotyrosine binding (PTB) domains, which can specifically recognize these pTyr motifs^[73], with SH2 domains being the most predominant ones in the human proteome.^[78,79] Their affinity for pTyr substrate has been studied and investigated extensively. It has been reported that about 50% of

the binding affinity of SH2 domains is directly attributed to pTyr.^[80]

SH2 domains are essential in a variety of signaling pathways, and therefore, mutations in the SH2 domain can be found in various human diseases.^[73,76,79]

In order to understand the function of different phosphorylation sites, the accessibility of phosphorylated building blocks is required.

1.6.2 Synthesis of Fmoc-phosphoTyrosine Building Blocks

To investigate the function of pTyr, synthetic routes to access phosphorylated proteins are required. Mostly, pTyr is introduced via Fmoc-based SPPS yielding a phospho peptide. The phosphate group can either be introduced globally or via single building blocks. By the introduction of building blocks, a phosphorylated N-protected amino acid is incorporated during the growth of the peptide chain.^[73,74]

Over the last decades, several building blocks to introduce phosphotyrosines via SPPS have been introduced. Fmoc-Tyr(PO₃H₂)-OH (Figure 6 (A)) itself is barely used in SPPS since the unprotected P-OH groups may cause synthetic problems like the formation of pyrophosphate between adjacent Tyr(PO₃H₂) residues.^[81] Thus, protected pTyr building blocks have been introduced. In the scope of this work, two of these building blocks Fmoc-Tyr(PO(OBzl)OH)-OH (Figure 6 (B)) and Fmoc-Tyr(PO(NMe₂)₂)-OH (Figure 6 (C)) will be described in more detail.

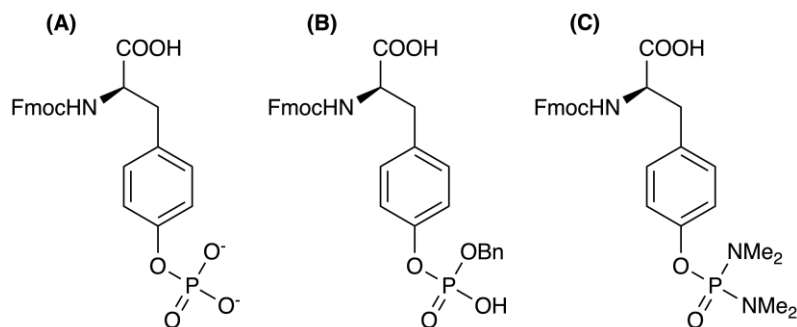
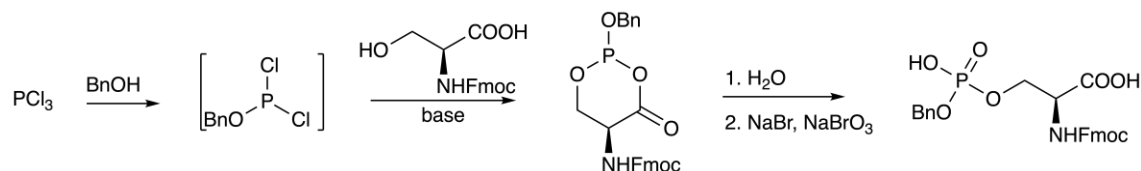


Figure 6: Fmoc-pTyr building blocks: **A)** Fmoc-Tyr(PO₃H₂)-OH, **B)** Fmoc-Tyr(PO(OBzl)OH)-OH, **C)** Fmoc-Tyr(PO(NMe₂)₂)-OH

Fmoc-Tyr(PO(OBzl)OH)-OH **B** has previously been synthesized in a multi-step procedure from Fmoc-Tyr via the N-Fmoc-O-dibenzylphosphono-Tyr-OH as intermediate.^[82]

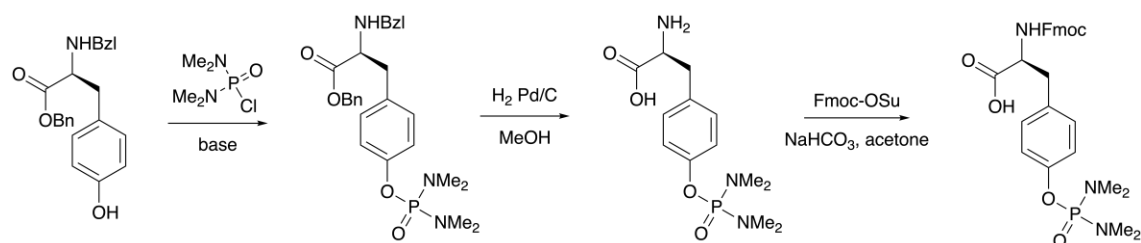
In 2012, Petrillo *et al.* reported on a one-pot synthesis of Fmoc-Tyr(PO(OBzl)OH)-OH. Treatment of PCl₃ in the presence of benzyl alcohol and a base led to the benzyl dichlorophosphite. Subsequent treatment with the amino acid in its anionic form allowed for the corresponding phosphite formation. In the case of serine or threonine, a cyclic phosphite as an intermediate was proposed to form. Oxidation

via bromine-related oxidants yielded the corresponding phosphate. However, the authors reported troublesome purification, whereas precipitation of the phosphorylated Fmoc-protected amino acid was not feasible due to impurities. Scheme 3 depicts the described one-pot synthesis of phosphorylated serine.^[6]



Scheme 3: One-pot synthesis of protected phosphorylated amino acids; *adapted from*^[6]

Tyr(PO(NMe₂)₂)-OH is a typical example of a phosphorodiamidate building block first introduced by Chao *et al.*^[5] Based on this procedure, another series of *N,N'*-dialkylamide-type phosphate protecting groups to introduce pTyr via SPPS have been introduced and optimized.^[83] Fmoc-phosphoTyr compound **C** is synthesized in three steps starting from the commercially available *Z*-Tyr-OBzl as depicted in Scheme 4.^[5] The hydrolytic lability of the P-N bond in these phosphorodiamidates under acidic conditions is the basis for this protection strategy. Hence, the P-N bond remains stable under alkaline conditions. However, the P-N bond hydrolyses to the corresponding phosphate compound in acidic conditions.^[84]



Scheme 4: Introduction of *N,N'*-dimethyldiaminophosphinoyl moiety to the tyrosine side-chain; *adapted from*^[5]

There is still huge interest today in implementing new pTyr-containing peptides to mediate atomistic details of phosphorylation sites.^[73]

1.6.3 Synthesis of Isotopically Labeled Tyrosine and Tyrosine Precursors

In contrast to selectively labeled aliphatic amino acid precursors, the labeling of aromatic ones has only developed recently.^[58] (4-Hydroxyphenyl)pyruvate is converted to tyrosine by the transaminase in the metabolic pathway of *E. coli*. Hence, due to the absence of a chiral center, (4-hydroxyphenyl)pyruvate marks an ideally

accessible compound for organic synthesis. Therefore, it has been widely applied in selective tyrosine labeling.^[3,58,85]

For optimal NMR spectra of the expressed protein, isolated ^{13}C - ^1H spin systems in the aromatic ring of Tyr are required for a magnetization transfer and a defined spin relaxation. By the introduction of alternating ^{13}C - ^{12}C - ^{13}C and ^2H - ^1H - ^2H spin systems in the aromatic side chain of Tyr, NMR signals show good resolution due to a significant decrease of scalar and dipolar couplings.^[86]

Previous reports on synthetic routes demonstrated the incorporation of these isolated spin systems from relatively cheap available ^{13}C -([1,3- $^{13}\text{C}_2$]acetone) and ^2H - (deuterium oxide) sources into 4-aminophenol (Figure 14) and subsequently into (4-hydroxyphenyl)pyruvate as a late metabolic precursor for tyrosine.^[3] Based on this synthesis, further developments toward the synthesis of the spin-isolated amino acid tyrosine have been made. Hence, via the iodination of 4-aminophenol in a Sandmeyer iodination, labeled 4-iodophenol has been synthesized. In a subsequent Negishi cross-coupling reaction, protected iodoalanine with the labeled 4-iodophenol yielded protected labeled Tyr in good yields.^[4]

Apart from the chemical aspects regarding the synthesis of isotope-labeled precursors, one also has to consider the effectiveness of its uptake by the overexpressing host organism. When supplementing *E. coli* with (4-hydroxyphenyl)pyruvate, 80-200 mg/L of growth media have been shown to lead to near quantitative uptake and incorporation of the precursor.^[58,85] However, it has been reported that lower amounts are required when supplementing the growth media with the corresponding amino acid itself.^[4]

Negishi cross-coupling reactions

As reported in literature, a Negishi cross-coupling reaction has been used to couple a protected iodo-alanine with the isotope containing iodophenol to yield a protected tyrosine compound.^[4] The Negishi coupling displays, amongst other substantial cross-coupling reactions, carbon-carbon bond-forming reactions.^[87,88] It was first reported in 1977^[89] and is based on Pd- or Ni-catalyzed coupling of organometals to a halide compound.^[87,88]

The proposed mechanism for the very reliable Negishi cross-coupling reaction is depicted in Figure 7.^[90]

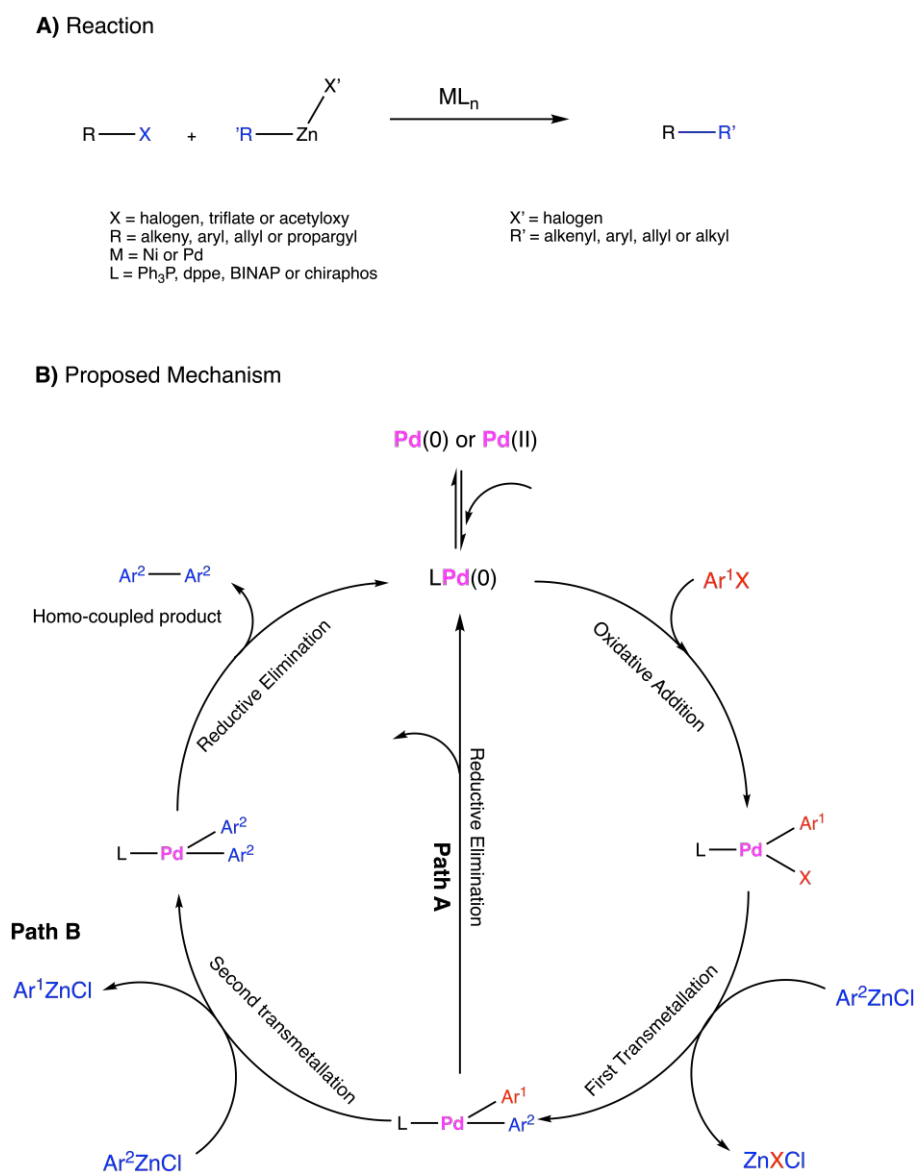


Figure 7: **A)** general reaction scheme of a Negishi cross-coupling reaction and **B)** proposed mechanism for the Negishi coupling; adapted from^[90]

2 Aims

In the first part, this thesis aimed to introduce a ^{13}C -H spin system into a building block suitable for SPSS. The synthetic strategy aiming for **18** as the target compound, is based on the introduction of the isotope pattern within 4-iodophenol, a Negishi cross-coupling reaction, and a final one-pot phosphorylation strategy. The first key steps of the second synthetic strategy aiming for the other Fmoc-pTyr building block **11** are similar to the one for **18**. However, phosphorylation is planned to be carried out in a three-step protocol. The resulting compound can then be applied in SPSS (and chemical ligation) protocols to generate pTyr-containing proteins. These protein samples will serve as important reporters to decipher the molecular details of phosphorylation cascades in healthy and pathogenic cells using biomolecular NMR.

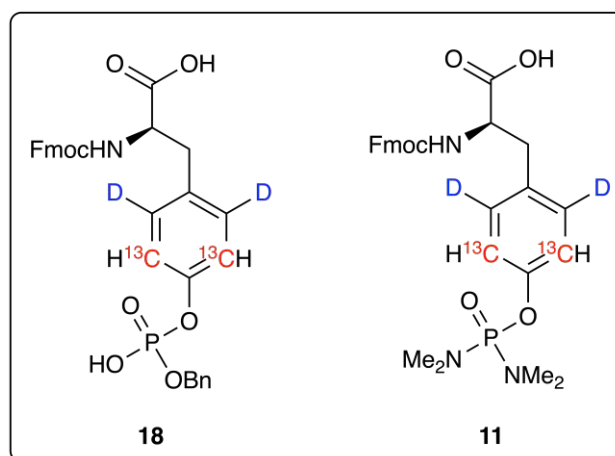


Figure 8: Fmoc-phosphoTyrosine building blocks **11** and **18** containing an isolated $^{12,13}\text{C}$ -H spin system

In the second part of this thesis, the aim was to synthesize the aliphatic amino acid precursors for isoleucine **24** and valine/leucine **30** according to known literature.^[7]

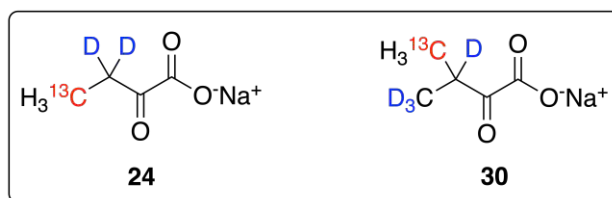


Figure 9: Aliphatic amino acid precursors for isoleucine **24** and valine and leucine **30**

In the third part, we aimed for a new isotopologue of imidazole-pyruvic acid **39** as a precursor for histidine. The precursor is intended to contain a ^{15}N - ^{13}C - ^{15}N side chain with low-cost sources of carbon-13 ((para-)formaldehyde) and nitrogen-15 (ammonium chloride). The precursor can be added to *E. coli* growth media. Thus, the resulting overexpressed protein is expected to give precious information about protonation levels of histidine within biomolecular NMR applications.^[91]

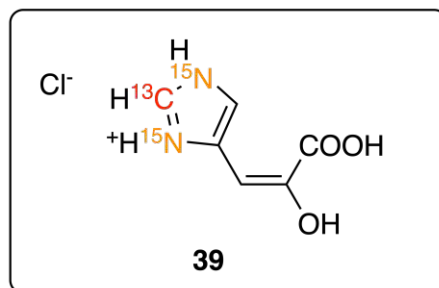


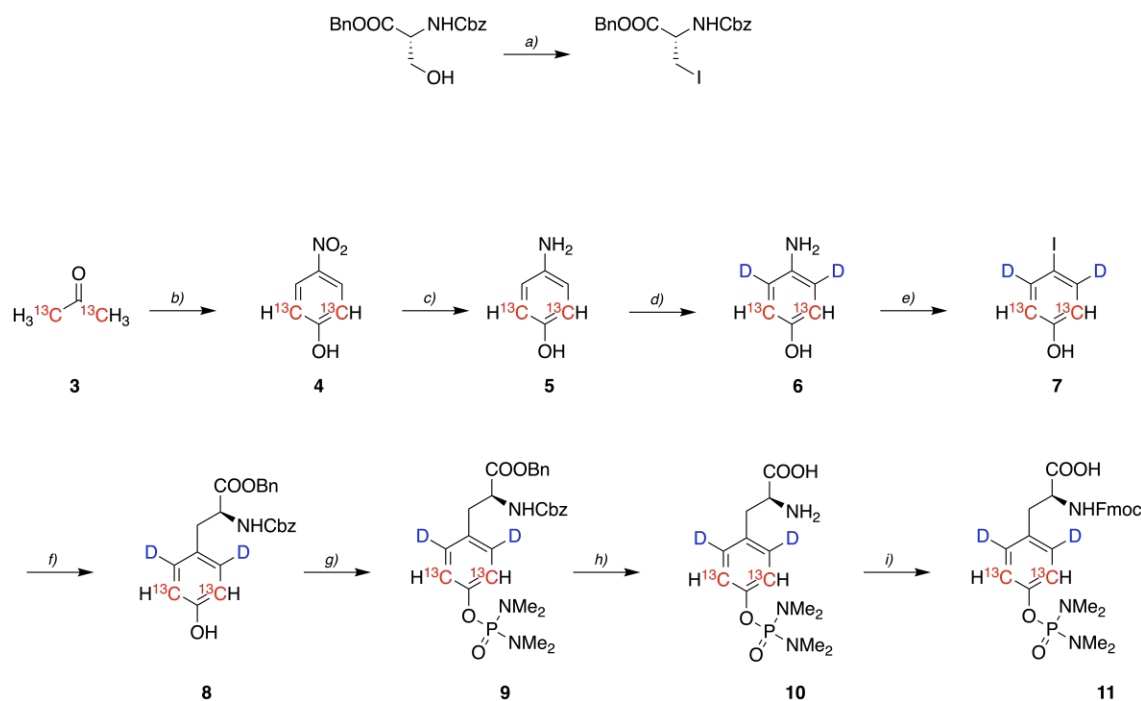
Figure 10: A new isotopologue of imidazole pyruvic acid **39** - a metabolic precursor for histidine

3 Results and Discussion

3.1 Fmoc-phosphoTyrosine Synthesis

To obtain a pTyr building block suitable for Fmoc-based SPPS, two different synthetic routes for two different Fmoc-pTyr compounds (**11** and **18**) were approached. Both synthetic sequences were optimized with the respective unlabeled compounds. However, purification of *N*- α -Fmoc-*O*-benzyl-L-phosphotyrosine **18** was not feasible within the scope of this work. Further optimization would be required, for example, purification via a preparative high performance liquid chromatography (pHPLC). Compound **11** could be synthesized.

3.1.1 *N*- α -Fmoc-*O*-(bis-dimethylaminophosphono)-[3,5- $^{13}\text{C}_2$ -2,6- D_2]-L-tyrosine **11**



Scheme 5: Synthesis of *N*- α -Fmoc-*O*-(bis-dimethylaminophosphono)-[3,5- $^{13}\text{C}_2$ -2,6- D_2]-L-tyrosine **11**: a) PPh_3 , Imidazole, I_2 , DCM, 0°C - RT, 87%; b) Nitromalon-aldehyde, H_2O , NaOH, 6 d, 4°C , 54%; c) H_2 balloon, Pd/C, MeOH, 2 h, RT, quant.; d) D_2O , HCl, Microwave 37 min, 180°C , quant.; e) 1) NaNO_2 , H_2SO_4 , DMSO, 1 h, $0 - 5^\circ\text{C}$, 2) NaI, overnight, RT, 44%; f) **2**, Zn, I_2 , $\text{Pd}_2(\text{dba})_3$, SPhos, DMF, overnight, 40°C , 81%; g) DBU, DMAP, bis-(dimethylamino)phosphorylchloride, DCM, 2 h, 0°C - RT, 90%; h) H_2 , Pd/C, MeOH, 2 h, 40°C , quant.; i) Fmoc-OSu, Acetone, NaHCO_3 , 4 h, 0°C - RT, 75%

N- α -Fmoc-*O*-(bis-dimethylaminophosphono)-[3,5- $^{13}\text{C}_2$ -2,6- D_2]-L-tyrosine **11** was obtained through a nine-step synthesis as illustrated in Scheme 5.

Starting from the commercially available protected L-serine **1**, the corresponding protected iodoalanine **2** was obtained via an Appel reaction.^[92] Purification of **2** via silica gel column chromatography was not feasible since the compound decomposed. However, most of the POPh_3 impurity could be removed from the reaction mixture by precipitation using diethyl ether, which was sufficient for the subsequent cross-coupling reaction.

Synthesis of compound **6** was performed similarly to a procedure reported in literature:^[3] Compound **4** was obtained by reacting isotope-labeled acetone with nitromalonaldehyde in basic aqueous conditions with moderate yields. Subsequent reduction of **4** to **5** was performed over Pd/C using a hydrogen balloon, giving quantitative yields. Studies on selective deuteration of 4-aminophenol induced by microwave have also been demonstrated in recent literature.^[3] Based on these results, parameters for deuteration of **5** to **6** were chosen. Figure 11 shows ^1H -NMR spectra of compounds **5** (top), **6** (middle) and **7** (bottom). The NMR spectra revealed a deuteration level of approximately 80% at positions 2 and 6.

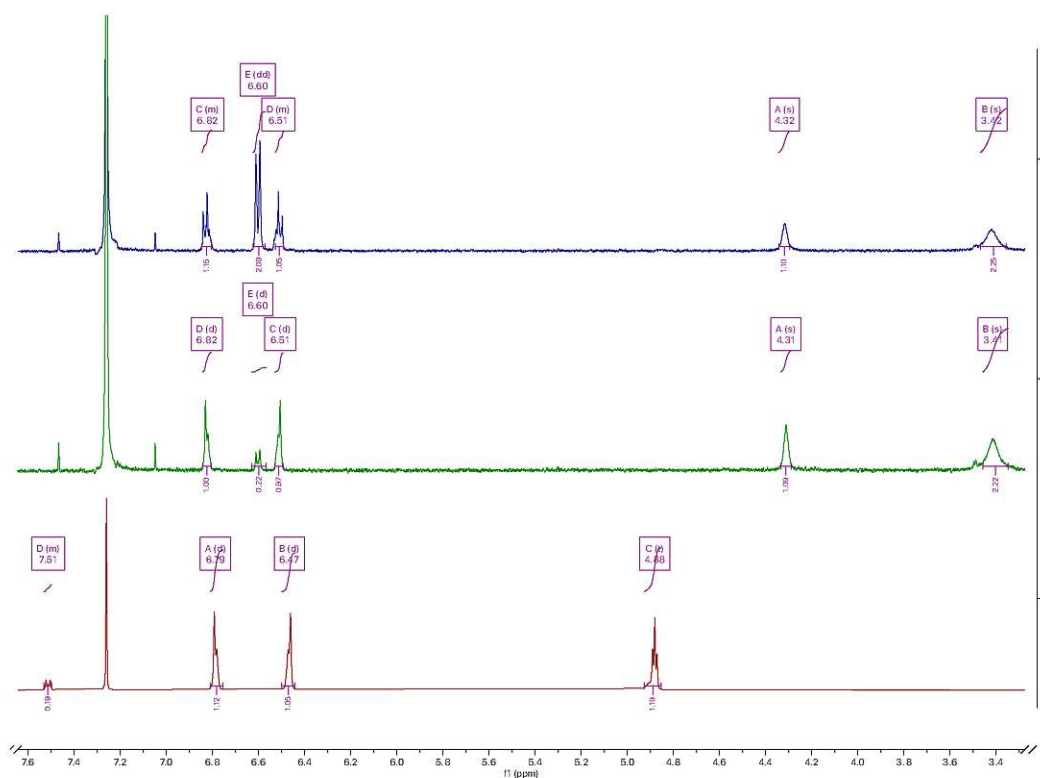


Figure 11: ^1H NMR spectra of compounds **5** (top), **6** (middle) and **7** (bottom) in CDCl_3 revealing deuteration grade of **6** of 80%. *Top*: Signal E corresponds to the protons on positions 2 and 6 of compound **5**. *Middle*: Signal E corresponds to the protons on positions 2 and 6 of compound **6**, revealing a deuteration level of 90% on positions 2 and 6. *Bottom*: Signal D corresponds to the protons on positions 2 and 6 of compound **7**, which reveals no change in deuteration level from **6** to **7**.

Compound **7** was prepared using a recently published Sandmeyer iodination procedure.^[4] Although the reaction occurs in quite acidic conditions (30% H₂SO₄), deuteration levels did not change significantly. As depicted in Figure 11 (bottom), about 10% each on positions 2 and 6 remain non-deuterated. Notably, yields for the unlabeled approaches for **7** were significantly higher (78% vs. 44%). A possible explanation for this would be the substrate's purity. In the unlabeled approach, commercially available 4-aminophenol was used. Whereas in the actual synthetic route, compounds **5** and **6** have partially undergone oxidation due to 4-aminophenol's light and air sensitivity^[93]. The color change from beige to dark purple, observed for **5** and **6**, indicated their oxidation.

Compound **8** was prepared by a Negishi cross-coupling reaction in good yields. Subsequent phosphorylation was planned according to a procedure by Chao et al.^[5] However, optimization and several modifications were implemented. The phosphorylation motif was introduced to **8** using bis-(dimethylamino)phosphorylchloride. NMR spectra after work-up showed small amounts of unreacted excess bis-(dimethylamino)phosphorylchloride or the corresponding hydrolyzed compound. However, this impurity did not hinder further synthesis steps, so no further purification was carried out. Deprotection from **9** to **10** was performed using 0.4 eq Pd/C in methanol at 40°C over 2 h using a H₂ balloon. As well as 0.1 - 0.2 eq Pd/C and/or hydrogenation at RT did not show any product formation on TLC control. Hydrogenation overnight, but also at 50°C with 0.1 - 0.2 eq Pd/C, led to the decomposition of the product resulting in a brownish oil. Thus, the hydrogenation reliably yielded the free amino acid by adding more equivalents of the catalyst. Concerning solvents, pure methanol was more efficient than a methanol/THF (3/1 v:v) mixture. Subsequent Fmoc-protection using dioxane/NaHCO₃ did not yield the desired product. Using acetone/NaHCO₃ instead and optimized work-up conditions yielded **11** as white fine crystals. Extracting **11** into the organic phase was challenging since the P-N bond is unstable under acidic conditions.^[5,46] Instead of HCl, citric acid was used to acidify the aqueous phase, decreasing the P-N bond hydrolysis. The compound was triturated with heptane to increase the purity of **11**. The structure of compound **11** is depicted in Figure 12 with the corresponding ¹H NMR spectra in Figure 12. Protons at carbon-13 in the aromatic ring of tyrosine (*E*, *F*) show a typical ¹J_{CH} coupling constant of 158.59 Hz. The ¹H NMR spectrum corresponds to the ¹H NMR spectra of the non-isotopically labeled compound from the original literature.^[5] Additionally, the experimental mass for **11** of 542.2301 (ESI-MS pos. mode) matches with the calculated mass of 542.2294.

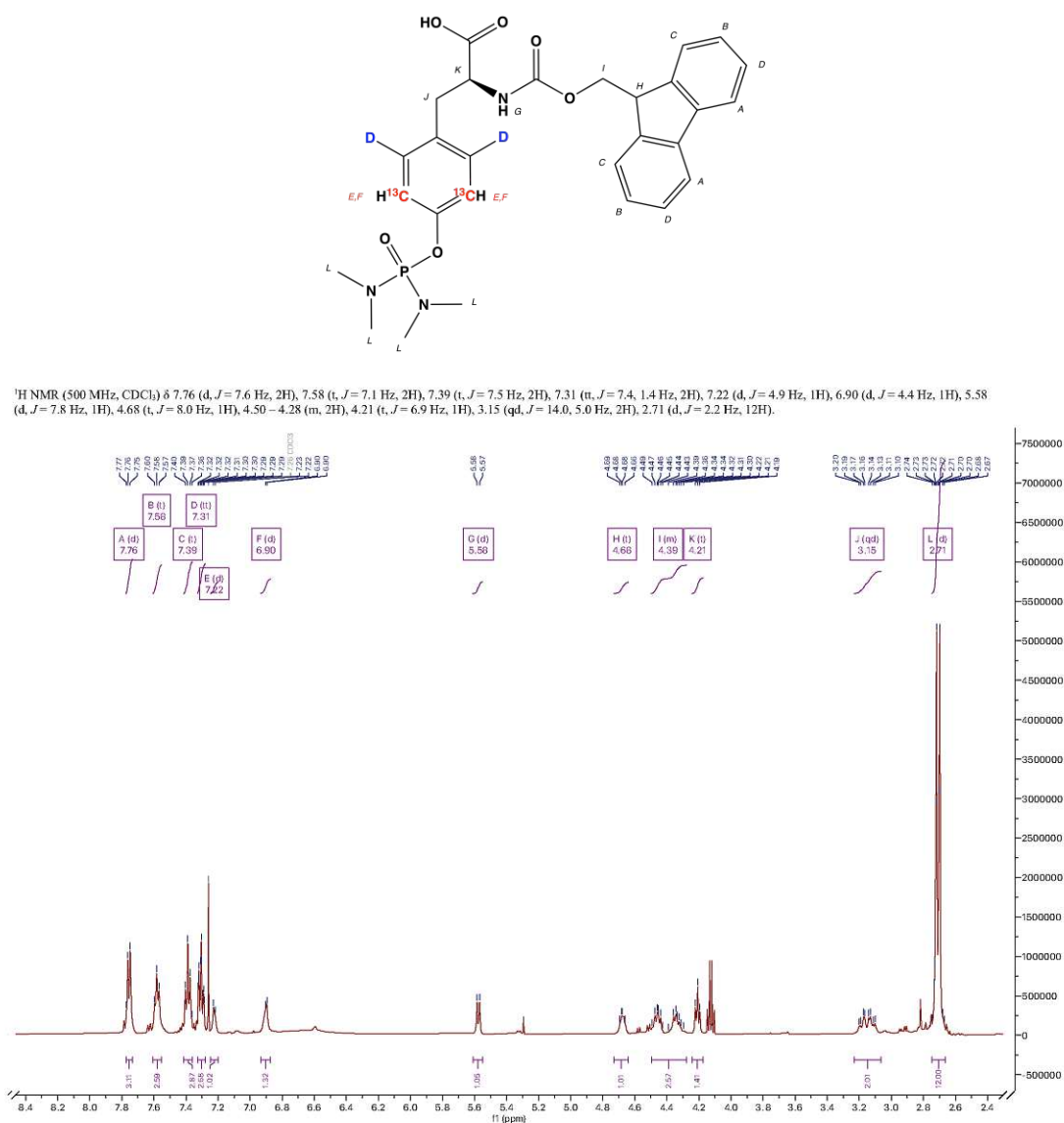
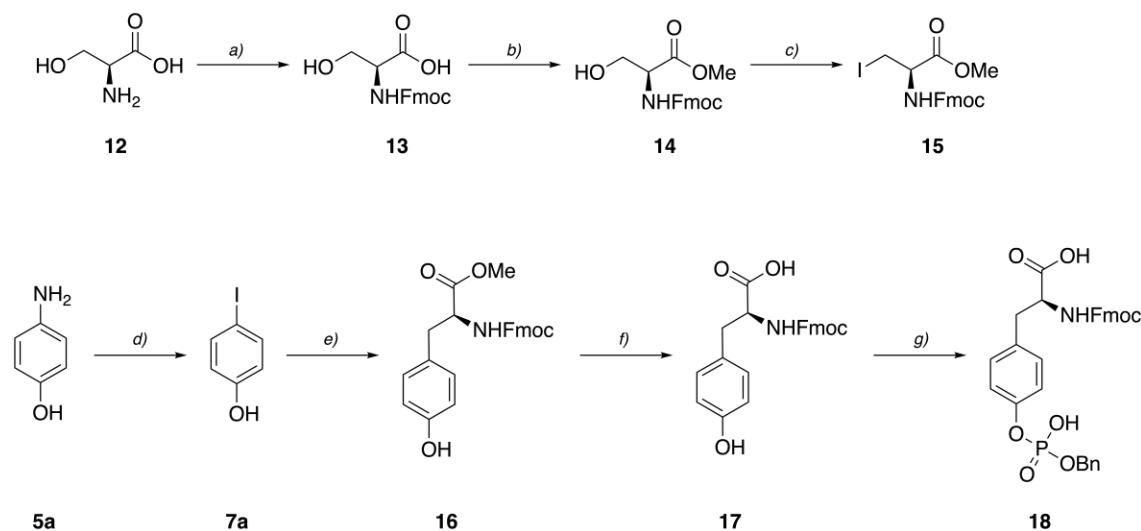


Figure 12: The structure of **11** with atom labels is shown (*top*, referring to the ¹H spectra on the *bottom*).

Overall, the synthesis of **11** was optimized and finally succeeded with an overall yield of 13%. Taking the huge costs of the carbon-13 labeled acetone into consideration, which is introduced within the first step of the synthesis sequence, further yield optimization would be advantageous. Especially the Sandmeyer iodination resulted in low yields of only 44%. Additional purification steps of the reaction's substrate would be required to improve its yield. However, Deuteration of **5** to **6** is performed in D₂O. Thereby, the air and light-sensitive aminophenol **5** and **6** partly undergo oxidation which most likely explains the low yield.

3.1.2 *N*- α -Fmoc-*O*-benzyl-L-phosphotyrosine **18**

Scheme 6: Synthesis of *N*- α -Fmoc-*O*-benzyl-L-phosphotyrosine **18**: a) Fmoc-Cl, Na₂CO₃, Dioxane, 2 h, 0°C - RT, 95%; b) H₂SO₄, MeOH, 3 h, reflux, 89%; c) PPh₃, Imidazole, I₂, DCM, 0°C - RT, 87%; d) 1) NaNO₂, H₂SO₄, DMSO, 1 h, 0 - 5°C, 2) NaI, overnight, RT, 78%; e) **15**, Zn, I₂, Pd₂(dba)₃, SPhos, DMF, overnight, 40°C, 82%; f) NaOH/CaCl₂, Isopropanol/H₂O, 7 h, RT, 90%, g) 1) PCl₃, BnOH, 2,6-Lutidine or Pyridine, THF, 90 min 0 - 5°C, 2) NaBr/NaBrO₃, H₂O; not purified from crude

Labeled *N*- α -Fmoc-*O*-benzyl-L-phosphotyrosine **18** as the second possible labeled phosphorylated Fmoc-Tyr compound, was planned to be synthesized with the same labeling pattern in the aromatic ring as for compound **11**. However, *N*- α -Fmoc-*O*-benzyl-L-phosphotyrosine **18** purification requires further optimization. Hence, within the scope of this thesis, the synthesis was not performed with corresponding labeled compounds.

In this synthetic strategy, the protected iodoalanine **15** was prepared from commercially available L-serine **12** in a three-step procedure with small adaptations to known literature in good yields.^[94] In the first step, L-serine was Fmoc protected using Fmoc-Cl in dioxane/Na₂CO₃. In subsequent acid esterification, the corresponding methyl ester **13** was formed, which could be used for the Negishi cross-coupling reaction. The protected iodoalanine **15** was obtained by an Appel iodination. Similarly, as described for **2**, purification of **15** using silica gel chromatography was not feasible since the compound decomposed on silica. Most of the POPh₃ was removed by precipitation from an ether solution. Hence, only small amounts of POPh₃ remained, and the partially purified **15** could be used for the subsequent Negishi cross-coupling reaction.

The 4-iodophenol **7a** was obtained from 4-aminophenol **5a** in a Sandmeyer iodination as described for the corresponding isotopologue **7**. However, yields were

significantly higher for the unlabeled synthesis, which is most certainly because 4-aminophenol **5a** from a commercial supplier is of higher purity than the respective isotopologues that had been exposed to light and air. Thus, the compound has undergone slight oxidation, which was observed by its change in color from beige to dark purple. The subsequent Negishi cross-coupling reaction yielded **16** in good yields. To obtain the free acid **17**, basic ester hydrolysis was performed according to known literature, where the Fmoc-group stability under basic conditions in the presence of CaCl_2 has been demonstrated.^[95] When carrying out the hydrolysis, on the one hand, no loss of the Fmoc-group was observed, and on the other hand, the ester was fully hydrolyzed to the free carboxylic acid.

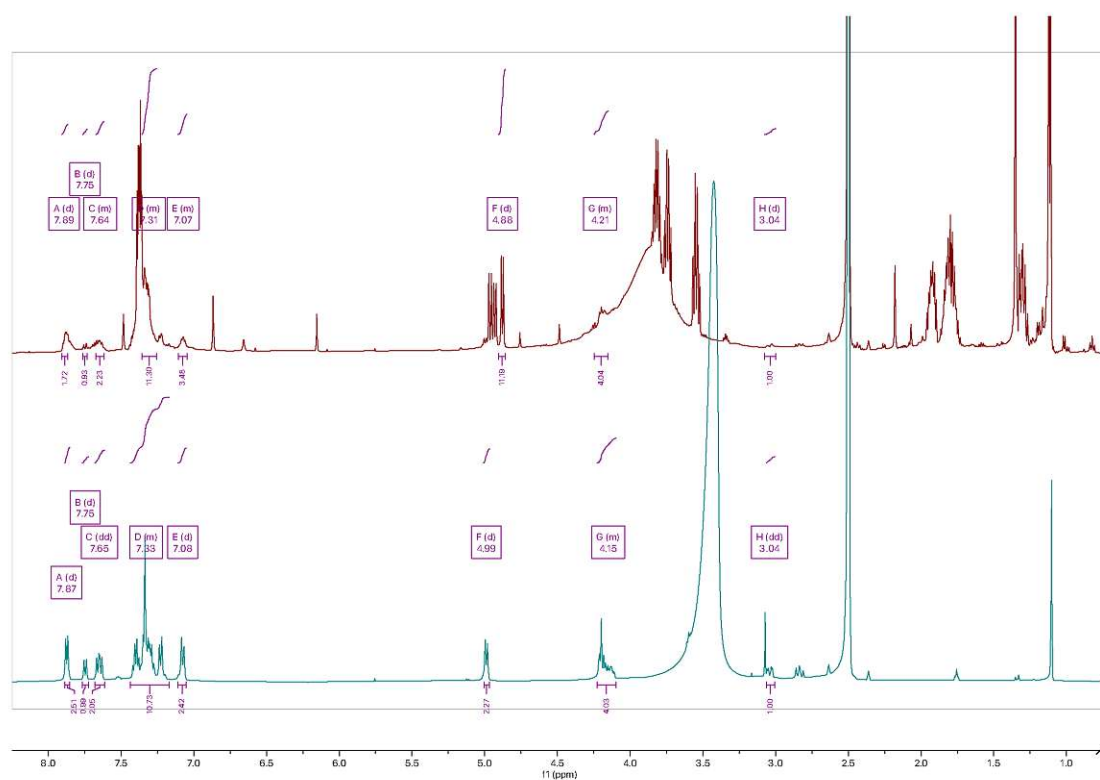


Figure 13: ^1H NMR spectra of **18** from the synthesized crude product (*top*) and the commercial sample (*bottom*)

Compound **18** was previously synthesized in a three-step procedure from tyrosine.^[82] More recent literature reported on a one-pot synthesis starting from Fmoc-protected amino acids with reasonable yields, which we aimed to reproduce for the isotope-labeled analog. **18** was synthesized according to that procedure.^[6] PCl_3 was treated with BnOH in either pyridine or 2,6-lutidine to form the benzyl dichloro phosphite. **17** was added in its anionic form to yield the phosphite intermediate. $\text{NaBrO}_3/\text{NaBr}$ was subsequently applied as an oxidation agent to the phosphate. Unfortunately, purification of **18** was not feasible via silica gel chromatography since the compound decomposed on silica. This was also tested and confirmed with a commercial sample

of compound **18**. Additionally, **18** was tried to be precipitated from several solvents (DCM, ethyl acetate, diethyl ether, heptane). Still, a pure product could either not be isolated or no precipitate formed. Furthermore, **18** seems unstable at temperatures above 40°C since a brown oil instead of an off-white solid was obtained when solvents were evaporated in vacuo above that temperature. Alternatively, an HPLC methodology could be applied to purify compound **18**. Figure 13 shows the ^1H NMR spectra of the crude product **18** (top) and the spectra of the commercial sample (bottom). According to the spectra of the crude, the synthesis of the desired compound **18** was accomplished. However, purification of **18** is not feasible within the scope of this thesis.

Figure 14 depicts the retrosynthetic design of a spin-isolated tyrosine, and the introduction of a phosphate moiety in tyrosine referring to compounds **11** and **18**.^[4]

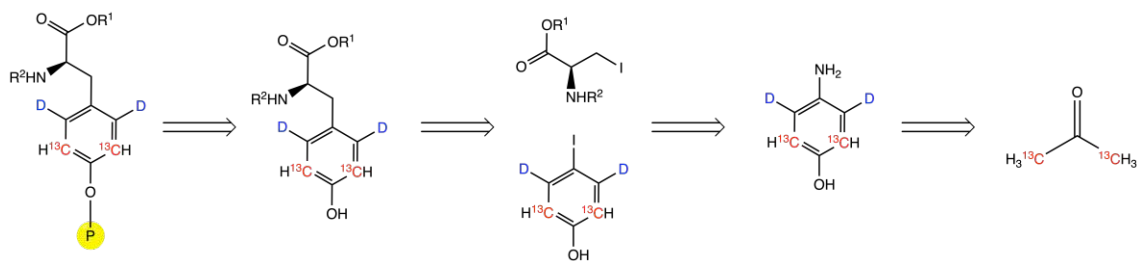
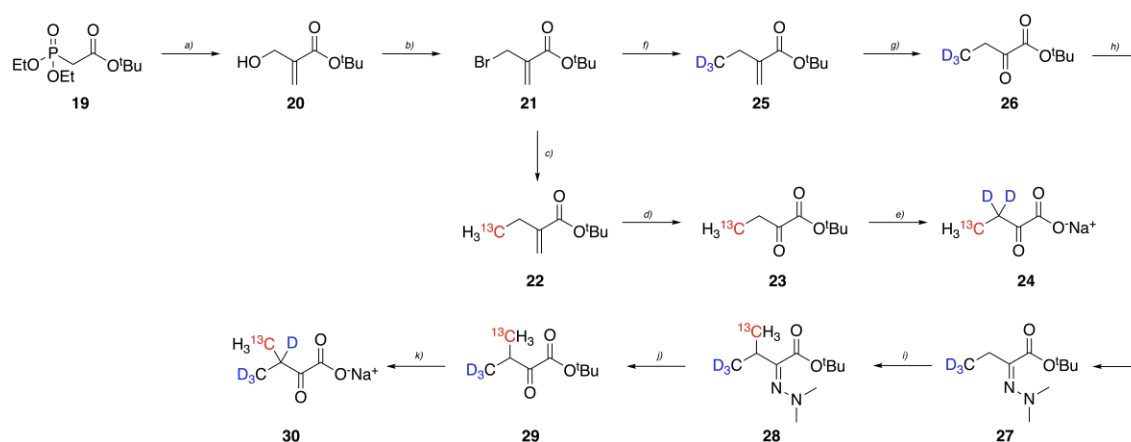


Figure 14: Retrosynthetic design of Fmoc-protected spin-isolated amino acids; adapted from^[4]

3.2 Aliphatic Precursor Synthesis

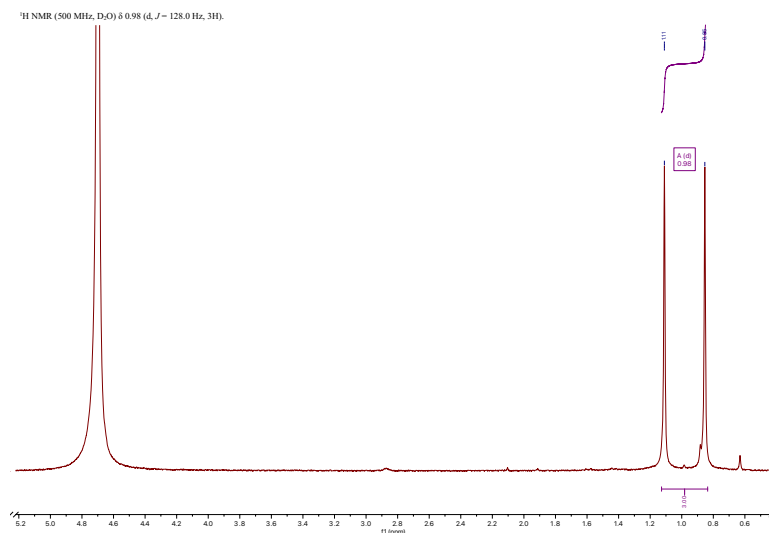
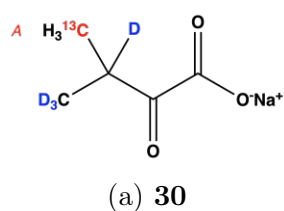


Scheme 7: Synthesis of the aliphatic amino acid precursors [4- $^{13}\text{C}_2$, 3- D_2] 2-ketobutanoic acid **24** and [4- ^{13}C ; 4- CD_3] ketoisovaleric acid **30**: a) formaldehyde, K_2CO_3 , H_2O , overnight, 40°C , 91%; b) PBr_3 , diethyl ether, -10°C - RT, 62%; c) Mg , $^{13}\text{CH}_3\text{I}$, diethyl ether, 80%; d) O_3 , DCM , PPh_3 , overnight, -78°C - RT, 69%; e) 1) $\text{HCl}(\text{g})$, DCM / diethyl ether, 1 h, 0°C - RT, 2) 0.5 M NaOD , D_2O , 2 h, RT, 92%; f) Mg , CD_3I , diethylether, 78%, g) = d), 58%; h) dimethylhydrazine, diethyl ether, overnight, RT, 78%; i) diisopropylamine, nBuli , $^{13}\text{CH}_3\text{I}$, THF , 2 h, -78°C - RT, 85%; j) 1 M HCl , THF , 1 h, RT, 72%; k) = e), then 6 M HCl , 2 h, reflux, 0.5 M NaOD , D_2O , 65%

The aliphatic amino acid precursors [4- $^{13}\text{C}_2$, 3- D_2] 2-ketobutanoic acid **24** and [4- ^{13}C ; 4- CD_3] ketoisovaleric acid **30** were prepared according to known literature starting from *tert*-butyl-diethylphosphorylacetate **19** as depicted in Scheme 7.^[7] For compound **24** an overall yield of 29% could be achieved. For compound **30**, which is synthesized in a longer route, an overall yield of only 8% was obtained. The highest potential of optimization is given within the hydrolysis sequences from **28** to **29** to **30**. The deprotections are carried out sequentially since that allows for the purification of **29** by silica gel chromatography. However, a small percentage of **28** might react straight to **30**, which was probably lost during the work-up.

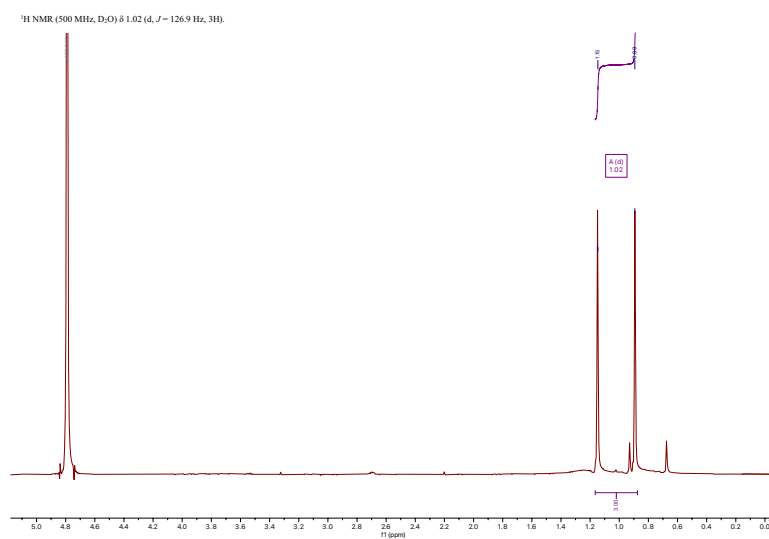
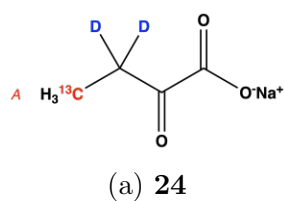
Nevertheless, complete hydrolysis from **29** of the *tert*-butyl group to **30** was not achieved by the usage of gaseous HCl . Hence, the partially hydrolyzed product was refluxed in 6 M HCl for 2 h resulting in complete hydrolysis to **30**.

Figure 15b and Figure 16b show the ^1H NMR spectra of **30** (see Figure 15a) and **24** (see Figure 16a). Thereby, it is visualized that the deuteration level on the β -carbon is at 100%. The coupling constants of the protons attached to carbon-13 show a value of about 128 Hz - a typical value for a $^1J_{\text{CH}}$ coupling and match to literature data.^[7]



(b) ¹H NMR spectra of **30** showing the protons *A* with the coupling constant of 128.0 Hz

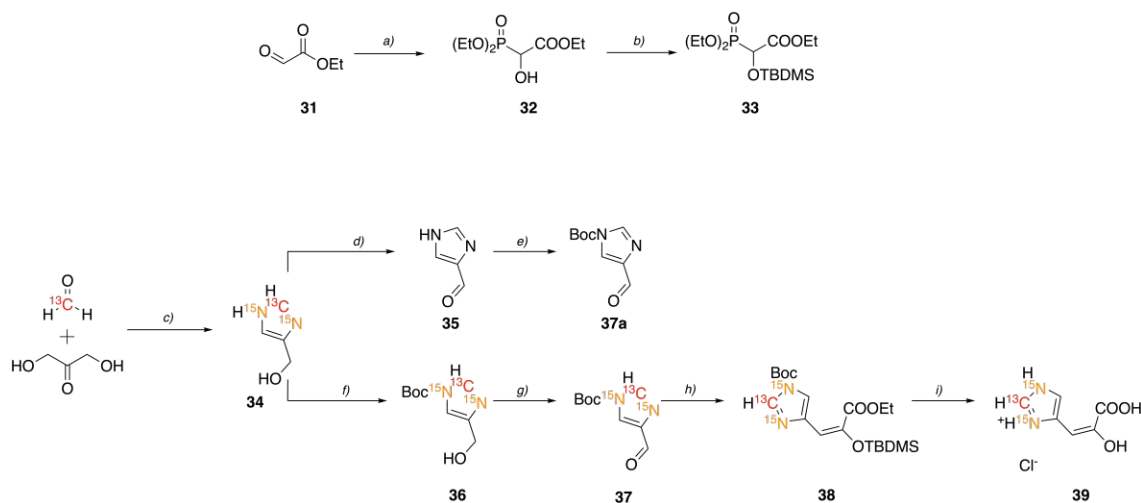
Figure 15: [4-¹³C; 4-CD₃] Ketoisovaleric acid **30**



(b) ¹H NMR spectra of **24** showing the protons *A* with the coupling constant of 128.2 Hz

Figure 16: [4-¹³C₂, 3-D₂] 2-Ketobutanoic acid **24**

3.3 Histidine Precursor Synthesis



Scheme 8: Synthesis of [1,3- ^{15}N , 2- ^{13}C]4-(2-carboxy-2-hydroxyvinyl)-1*H*-imidazolium chloride **39**: a) pTsOH, diethyl phosphonate, toluene, reflux, overnight, 65%; b) imidazole, TBDMSCl, DCM, 0°C - RT, overnight, 63%; c) $^{15}\text{NH}_4\text{Cl}$, $\text{CuCO}_3 \cdot \text{Cu}(\text{OH})_2$, MW 100°C, 1 h, then thioacetamide, H_2O , 50°C, 2 h, 62%; d) MnO_2 , MeOH, reflux, overnight, 80%; e) DMAP, $(\text{Boc})_2\text{O}$, ACN, 0°C - RT, overnight, 23%; f) $(\text{Boc})_2\text{O}$, Et_3N , DMF, THF, 0°C - RT, overnight, 74%; g) MnO_2 , MeOH, RT, 48 h, 84%; h) LiHMDS, **33**, THF, -78°C - RT, overnight, 58%; i) 6 M HCl, RT, overnight, quant.

When introducing nitrogen-15 into the imidazole ring, $^{15}\text{NH}_4\text{Cl}$ is the nitrogen source of choice. This compound has the advantage of being commercially available at lower costs compared to $^{15}\text{NH}_3$. Ammonia in the presence of Cu^{II} -ions was used in the original procedure for the synthesis of (imidazole-1*H*-5-yl)methanol (the unlabeled analog of **36**). There, aqueous ammonia was used as the nitrogen source in big excess.^[96] However, regarding the huge costs of isotope-enriched chemicals, we were looking for a strategy to incorporate nitrogen-15 as cost-efficient as possible.

The introduction of carbon-13 into the imidazole moiety was published recently.^[9] Based on this previous procedure, we aimed to synthesize **39** as depicted in Scheme 8. First, introducing nitrogen into the imidazole moiety was investigated and optimized. Table 2 lists the conditions that were applied to synthesize (imidazole-1*H*-5-yl)methanol, the unlabeled analogue of **36**. The best yields for (imidazole-1*H*-5-yl)methanol could be obtained using 4 equivalents of NH_4Cl and setting the pH to 10-12 with 1 M NaOH. A microwave reaction was chosen over a regular reflux reaction with the motivation to keep the formed ammonia inside the reaction vessel. Interestingly, adding two equivalents of NH_4Cl was insufficient to synthesize the desired compound. However, when the equivalents of NH_4Cl were increased to four, the (imidazole-1*H*-5-yl)methanol could be obtained in good yields. Adding more ammonium chloride (6 eq.) did not significantly change yields. Thus, the reaction

conditions for number 9 listed in Table 2 were chosen as the optimal approach using labeled compounds.

Table 2: Synthesis of (imidazole-1*H*-5-yl-)methanol: NaOH: +) NaOH was added up to a pH 12, –) no NaOH was added; work-up: *i*) precipitation of CuS using thioacetamide for 2 h at 50°C, *ii*) precipitation of CuS using thioacetamide for 2 h at 50°C with the addition of 1 M HCl

Nr.	Eq. NH ₄ Cl	NaOH	Type	Time	Work-up	Yield [%]
1	2	–	reflux	4 h	i	-
1	2	+	reflux	4 h	i	-
3	2	–	microwave 100°C	1 h	i	-
4	2	+	microwave 100°C	1 h	i	-
5	2	+	microwave 100°C	1 h	ii	-
6	4	+	microwave 100°C	1 h	i	75
7	4	+	microwave 100°C	1 h	ii	76
8	6	+	microwave 100°C	1 h	ii	70
9	4	+	microwave 100°C	1.5 h	i	82

As claimed before, the synthesis was optimized with corresponding non-labeled reagents. Unfortunately, when using 20% ¹³C-Formaldehyde solution, the reaction did not succeed as expected. Additionally, resources of 20% ¹³C-Formaldehyde solution were limited at our facilities. Thus, as an alternative carbon-13 source, ¹³C-Paraformaldehyde was used, which was stirred in 2 ml of 0.5 M NaOH overnight in the microwave vessel. The subsequent microwave reaction yielded compound **34** in good yields, which was by 10% lower than for the nonlabeled reaction. Figure 17 shows the ¹H NMR spectrum of **34** and reveals a carbon-13 content on position 2 of the imidazole motif of approximately 75%. The ¹J_{CH} coupling constant of 209 Hz aligns with a value from the literature and confirms the carbon-13 incorporation into the compound.^[9] The ²J_{NH} coupling constant is in the range of 6-7 Hz and can be nicely seen in Figure 17.

¹H NMR (500 MHz, DMSO) δ 8.03 (dtd, *J* = 209.5, 9.4, 1.2 Hz, 0.75H), 7.82 (td, *J* = 9.3, 1.2 Hz, 0.25H), 6.99 (dp, *J* = 7.6, 2.3 Hz, 1H), 4.39 (d, *J* = 2.0 Hz, 2H).

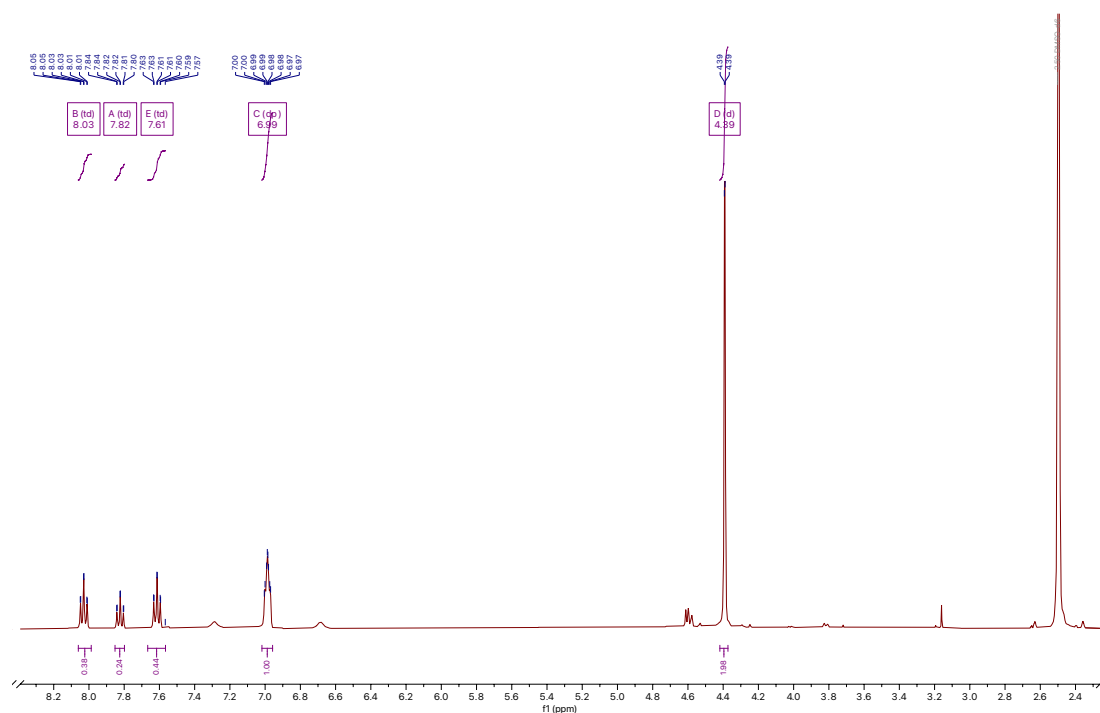


Figure 17: ¹H NMR spectrum of **34** picturing the incorporation of nitrogen-15 and carbon-13 by the depicted couplings.

Compound **37** has previously been synthesized via the oxidation of **34** to the corresponding aldehyde **35** using MnO₂ and subsequent Boc-protection, whereas purification was performed via silica gel chromatography.^[9] However, this procedure could not be reproduced in good yields. Purification using silica gel chromatography always resulted in yields around 20%. It was assumed that compound **37** is unstable on silica, which was also demonstrated on 2D TLC. Thus, other strategies to yield **37** were investigated. Compound **34** underwent Boc-protection yielding the two isomers of **36** for the unlabeled analogs. Boc-protection for the isotope-labeled approach interestingly only yielded one isomer. Besides extraction and washing of the organic phases with brine, no workup was performed since **36** partially underwent Boc-deprotection on silica. Therefore we chose to proceed with the crude product without further purification. **37** was obtained by oxidizing **36** to the corresponding aldehyde using MnO₂. Stirring the reaction mixture for 16 h overnight was insufficient, and only 50% of the alcohol was oxidized to the aldehyde. Therefore, the reaction was allowed to react for 48 h. Higher temperatures were not used due to the uncertain stability of the Boc group. However, the purification of compound **37** via silica gel chromatography has been described in literature before^[9,97] it was observed to be unstable on silica within this scope. Some minor steps like precipitation of the (side-)product were tested to purify the compounds, but finally, the crude **37**

3 RESULTS AND DISCUSSION

was used without further purification in the subsequent Horner-Wadsworth-Emmons (HWE) reaction to yield **38** as the *Z*- and *E*-isomer, as described previously.^[9] The final histidine precursor, the imidazole pyruvic acid **39** could be obtained by acid hydrolysis of the Boc- and TBDMS-protection-groups as well as the ester using 6 M HCl. Figure 18 shows the ¹H NMR spectrum of the imidazole pyruvic acid in good purity. Hereby, the ¹J_{NH} coupling of 90 Hz can be nicely seen for signal *D* in Figure 18. Furthermore, the NMR spectra analysis revealed a carbon-13 content at the ε₁ position in **39** of 70%. In summary, the desired compound **39** was successfully synthesized based on relatively cheap isotope sources for carbon-13 and nitrogen-15. However, the synthetic route requires some further optimization in order to allow for higher overall yields.

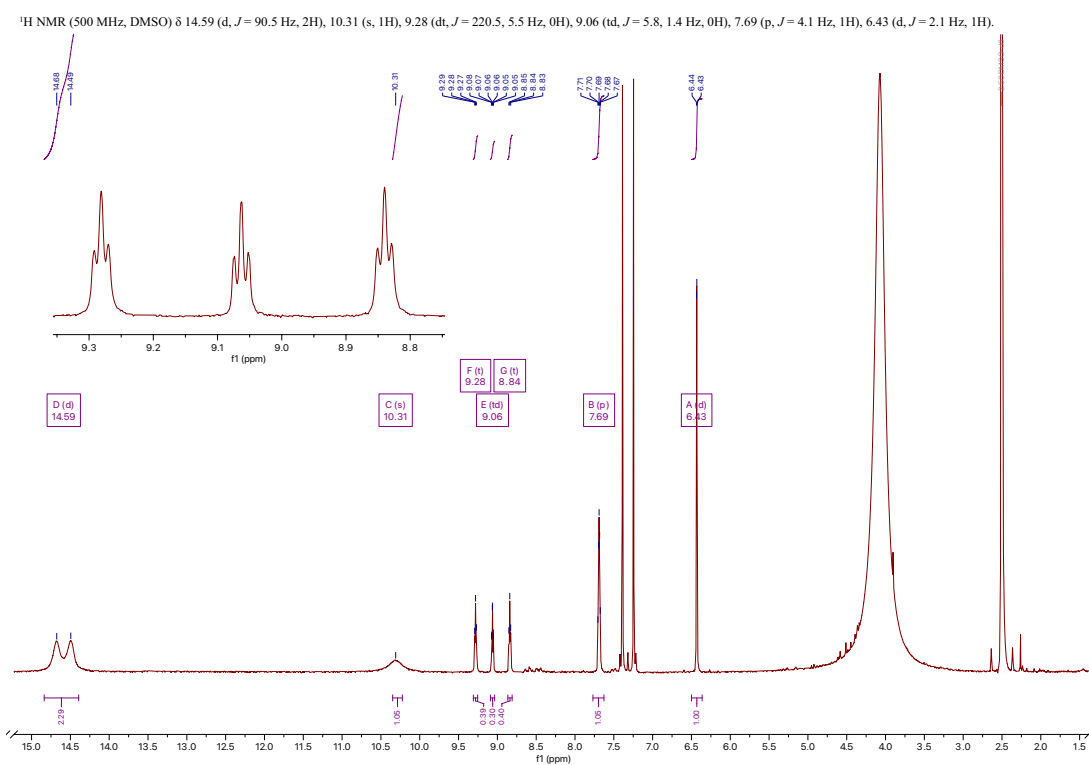


Figure 18: ¹H NMR spectrum showing the Enol-product of **34** with a carbon-13 content of 70%

4 Conclusion and Outlook

Regarding the Fmoc-pTyr building blocks **11** and **18**, the synthesis route for the latter one could not be fully optimized. However, **11** could be successfully synthesized and the protocol was optimized to allow for an overall yield of 13% with good purity. This building block now allows for the incorporation of an isotopically labeled pTyr into a small peptide. Thereby, new information about phosphorylated protein domains like the SH2 domain will be gained by using biomolecular NMR spectroscopy.

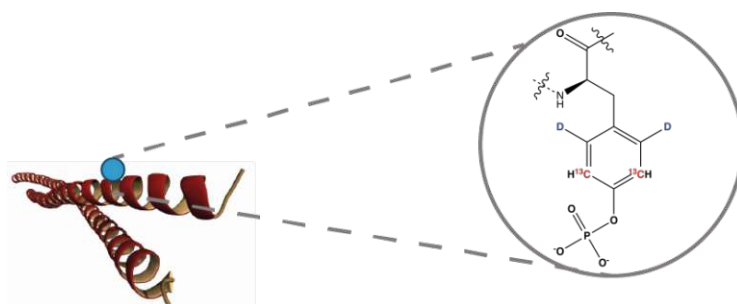


Figure 19: Isotopically labeled pTyr moiety in protein

As previously reported, the aliphatic amino acid precursors **24** and **30** were synthesized.^[7] The yielded compounds have been incorporated into several desired target proteins.

A protocol for a new isotopologue of imidazole pyruvic acid has been established. However, further optimization is still necessary since good yields in each step of the synthetic routes have not always been reproducible. Nevertheless, the synthesis also allows for other labeling patterns, such as nitrogen-15 as the only enriched isotope within the imidazole moiety.

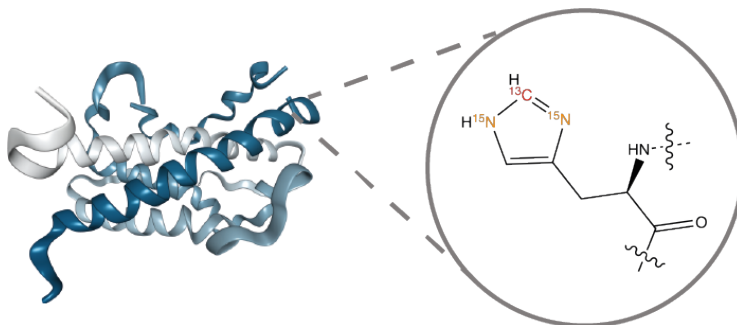


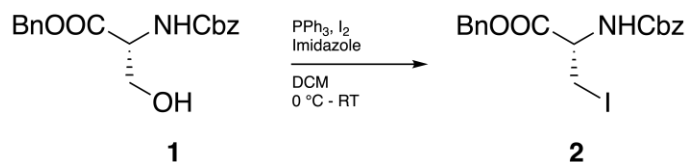
Figure 20: Protein with selectively labeled histidine

5 Experimental Section

5.1 General Remarks

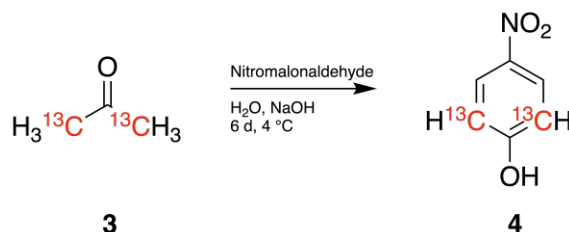
Unless otherwise stated, all reagents were purchased from commercial suppliers and used without further purification. Reaction progress was examined via thin layer chromatography (TLC) using pre-coated TLC sheets ALUGRAM[®] Xtra SIL G/UV₂₅₄. Compound peaks were visualized by UV-light irradiation (wavelength = 254 nm) and/or aqueous KMnO₄ solution, Mo-Ce(SO₄)₂ complex solution (48 g (NH₄)₆Mo₇O₂₄·4 H₂O and 2 g Ce(SO₄)₂ in 1 L 10% H₂SO₄) or Ninhydrin solution in Ethanol as dyeing reagents and subsequent heating via a heat gun. Column chromatography was performed manually using silica gel 60 from Merck (0.040 – 0.0633 mm) as the stationary phase. NaOD solutions were prepared by dissolving Sodium in D₂O. NMR spectra were recorded on a Bruker Avance NEO 500 or Bruker Avance NEO 600. NMR spectra were analyzed on MestReNova Version 14.3.1-31739. NMR solvent signals were calibrated to 7.26 ppm (CDCl₃), 4.79 ppm (D₂O) and 2.50 ppm (**6d**-DMSO). Mass spectra were measured on a Bruker maXis UHR-TOF spectrometer with electrospray ionization. Samples were lyophilized by freezing them in liquid N₂ and subsequently drying them at an Alpha 1–2 LDplus freeze dryer. Microwave reactions were performed on a BIOTAGE Initiator+ SW Version 5.1.0 build 12126. Ozonolysis reactions were performed with the usage of ozone ANSEROS generator Type SEP-100.

All synthesis steps were developed and optimized by using corresponding non-isotope enriched reagents. Isotope-enriched samples were purchased from Merck or CortecNet.

5.2 *N*- α -Fmoc-*O*-(bis-dimethylamino-phosphono)-*L*-tyrosine 115.2.1 (R)-2-Benzyloxycarbonylamino-3-iodo-propionic acid benzyl ester
2

Compound **2** was prepared following a literature procedure with slight modifications.^[92] PPh₃ (1.22 g, 4.6 mmol, 1.5 eq) and imidazole (1.18 g, 4.6 mmol, 1.5 eq) were dissolved in dry DCM and cooled in an ice bath. Iodine (320 mg, 4.6 mmol, 1.5 eq) was added in three portions sequentially. Stirring in the cooling bath was performed for 10 min, followed by stirring for 15 min at RT. Then, Z-Ser-OBzl **1** (1.00 g, 3.1 mmol, 1.0 eq) in DCM was added dropwise to the cooled reaction mixture. The reaction was stirred at RT for 3 h. Then, the reaction mixture was treated with 10% NaHSO₃ solution until the reaction mixture changed from an orange to a clear color. Extraction was performed thrice with DCM. The combined organic phases were dried over MgSO₄ and solvents were removed under reduced pressure until a slightly yellow oil was obtained. Product purification via silica gel chromatography was not feasible since the compound decomposed. To remove POPh₃, diethyl ether was added to the oil containing the product, and the flask was cooled in the fridge overnight to allow most of POPh₃ to precipitate. The next day, the supernatant containing the product was decanted, and the solvent was removed under reduced pressure to yield 1.19 g (87%) of **2** as white crystals.

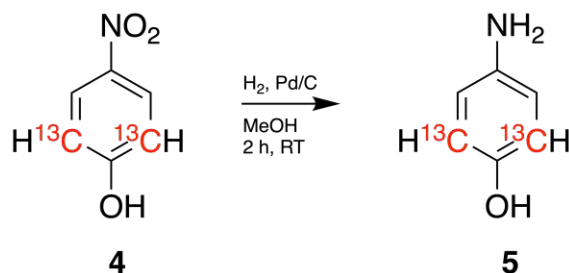
¹H NMR (500 MHz, CDCl₃) δ 7.74 – 7.61 (m, 10H), 5.68 (d, J = 7.5 Hz, 1H), 5.25 – 5.08 (m, 4H), 4.61 (dt, J = 7.6, 3.9 Hz, 1H), 3.59 (qd, J = 10.4, 3.8 Hz, 2H);

5.2.2 [2,6-¹³C₂]4-Nitrophenol 4

Synthesis of compound **4** was performed as described in literature.^[3] Nitromalonaldehyde (6.5 g, 55.5 mmol) was dissolved in 400 ml H₂O and cooled on an ice-water bath. [1,3-¹³C₂]acetone (2.0 g, 33.3 mmol) was added in one portion. Then, an aqueous solution of NaOH (8.80 g in 40 ml) was added dropwise to the reaction mixture using a dropping funnel. After addition, the flask was tightly closed and stirred for 6 days at 4°C. Then, the brown solution was dropwise treated with 50 ml of 6 M HCl, and the reaction mixture was filtered. The dark precipitate was taken up in 50 ml 6 M HCl and heated to 100°C for 13 min. After the filtration, the two filtrates were combined and extracted 10 times with diethyl ether. The combined organic phases were dried over MgSO₄, and evaporation of the solvent yielded the crude product as a dark yellow solid. Purification of the crude product via silica gel chromatography using heptane/ethyl acetate (3:2 v/v) yielded 2.49 g (53%) of **4** as bright yellow solid.

Rf-value 0.6 (heptane/ethyl acetate 3:2)

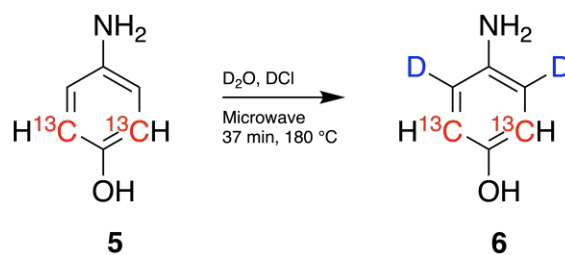
¹H NMR (600 MHz, D₂O) δ 8.12 (d, J = 7.9 Hz, 2H), 6.92 (dd, J = 169.6, 6.5 Hz, 2H).

5.2.3 [2,6-¹³C₂]4-Aminophenol **5**

Synthesis of compound **5** was performed as described in literature.^[3] [2,6-¹³C₂]4-nitrophenol **4** (500 mg, 3.54 mmol) and 10% Pd/C (46 mg, 0.43 mmol) were placed into a flask, and the flask was put under argon atmosphere. Dry methanol (10 ml) was added, and the mixture was flushed with H₂ using a H₂ balloon and was set under H₂ atmosphere. After 2.5 h, TLC showed completion of the reaction, and the catalyst was quickly removed via filtration over celite. Quick evaporation of the solvent under reduced pressure yielded 395 mg (100%) of **5** as greyish solid.

Rf-value 0.1 (heptane/ethyl acetate 3:2)

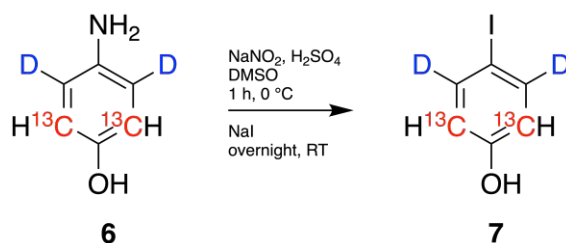
¹H NMR (500 MHz, CDCl₃) δ 6.83 (dd, J = 171.6, 8.5 Hz, 2H), 6.60 (dd, J = 8.0, 1.4 Hz, 2H), 4.32 (s, 1H), 3.42 (s, 2H).

5.2.4 [2,6-¹³C₂]3,5-Dideuterio-4-aminophenol **6**

Synthesis of compound **6** was performed as described in literature.^[3] In a microwave vessel [2,6-¹³C₂]4-aminophenol **5** (390 mg, 3.51 mmol) was placed into 5 ml of D₂O and 45 μ l concentrated HCl was added to the microwave vial. The microwave reaction was performed for 37 min at 180°C. After evaporation of the solvent under reduced pressure, methanol was added and the solvents were again removed under reduced pressure to yield 398 mg (100%) of [2,6-¹³C₂]3,5-dideuterio-4-aminophenol **6** as a grey solid. ¹H NMR spectroscopy showed a deuteration level of position 3 and 5 of 80%.

¹H NMR (600 MHz, D₂O) δ 6.84 (dd, $J = 159.0, 5.2$ Hz, 2H).

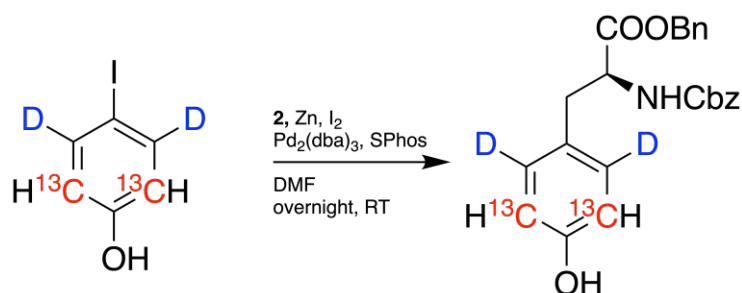
¹H NMR (500 MHz, CDCl₃) δ 6.82 (dd, $J = 153.8, 5.6$ Hz, 2H), 6.60 (d, $J = 8.4$ Hz, 0.2H), 4.31 (s, 1H), 3.41 (s, 2H).

5.2.5 [2,6-¹³C₂]3,5-Dideuterio-4-iodophenol **7**

Synthesis of compound **7** was slightly modified from literature.^[4] In a 3-neck round bottom flask, [2,6-¹³C₂]3,5-dideuterio-4-aminophenol **6** (380 mg, 3.36 mmol) was dissolved in 10 ml DMSO. A low temperature thermometer was applied, and the reaction mixture was cooled to -5 °C using an ice/NaCl cooling mixture. An aqueous 30% H₂SO₄ solution was added dropwise. When the internal temperature reached 0 °C, NaNO₂ (365 mg, 5.3 mmol) in 2 ml of H₂O was added dropwise so that the internal temperature remained below 5 °C. After stirring for 1 h at -5 - 0 °C on the cooling mixture, NaI (1.6 g, 10.7 mmol) in 2 ml of H₂O was added dropwise over 20 min. When the addition was complete, stirring of the reaction mixture was continued at RT for 1 h. Afterwards, NaI (1.6 g, 10.7 mmol) in 2 ml of H₂O was added again dropwise and the reaction mixture was stirred over night at RT. Then, extraction was performed four times using ethyl acetate. The combined organic phases were treated with 30 ml of 10% NaHCO₃ solution and stirred for 20 - 30 minutes at RT. Afterwards, the organic phase was separated, washed with brine and dried over MgSO₄. Evaporation of the solvents under reduced pressure yielded the crude product. Purification via silica gel chromatography using heptane/ethyl acetate (7/3 v:v) yielded 325 mg (43.3%) of [2,6-¹³C₂]3,5-dideuterio-4-iodophenol **7** as an off-white solid.

Rf-value 0.7 (heptane/ethyl acetate 7:3)

¹H NMR (500 MHz, CDCl₃) δ 7.51 (d, J = 9.7 Hz, 0.19H), 6.79 (dd, J = 165.4, 5.1 Hz, 2H), 4.88 (t, J = 4.7 Hz, 1H, OH).

5.2.6 *N*-Carbobenzoxy[3,5-¹³C₂-2,6-D₂]-L-tyrosine benzyl ester **8**

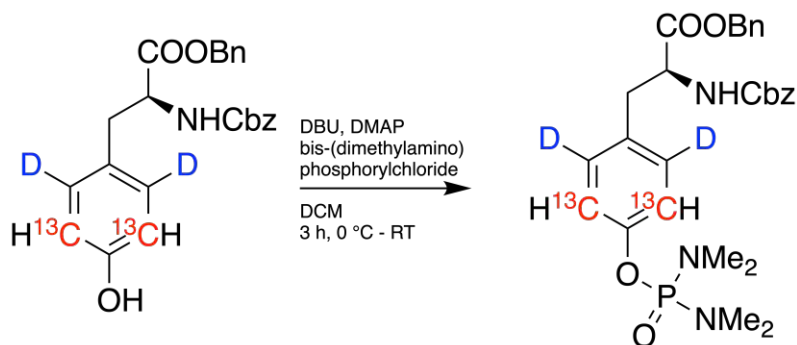
Synthesis of compound **8** was modified from literature.^[4] Zn (94 mg, 1.43 mmol) was placed into a two-neck flask, put under argon atmosphere and suspended in 2 ml absolute DMF. I₂ (13 mg, 0.10 mmol) was added and the reaction mixture was stirred for 5 min. (R)-Benzyl 2-(((benzyloxy)carbonyl)amino)-3-iodopropanoate **7** (225 mg, 0.57 mmol) dissolved in 2 ml absolute DMF was added dropwise over 5 min. Then, I₂ (13 mg, 0.10 mmol) was again added and the reaction mixture was allowed to stir for 30 min. Afterwards, Pd₂(dba)₃ (11 mg, 0.01 mmol), SPhos (9 mg, 0.02 mmol) and [2,6-¹³C₂]3,5-dideuterio-4-iodophenol (104 mg, 0.46 mmol) was added subsequently. The reaction mixture was stirred at 40°C overnight. Then, the reaction mixture was allowed to cool to RT, 20 ml of brine was added and the mixture was extracted 4 times with ethyl acetate. The combined organic phases were washed with brine, dried over MgSO₄, and the solvent was evaporated to yield the crude product. The product was purified via silica gel chromatography using heptane/ethyl acetate (3:2 v/v) as eluent, yielding 153 mg (81%) of *N*-carbobenzoxy[3,5-¹³C₂-2,6-D₂]-L-Tyrosine benzyl ester **8** as an off-white solid.

Rf-value 0.2 (heptane/ethyl acetate 3:2)

ESI-MS (pos. mode): calculated for C₂₄H₂₄NO₅ [M+H]⁺ 406.1649; found 406.149

¹H NMR (500 MHz, CDCl₃) δ 7.40 – 7.29 (m, 10H), 6.80 (dd, J = 157.2, 5.0 Hz, 2H), 5.21 (d, J = 8.3 Hz, 1H), 5.17 (d, J = 12.1 Hz, 1H), 5.12 (s, 1H), 5.09 (d, J = 5.6 Hz, 2H), 4.88 (brs, 1H), 4.66 (dt, J = 8.3, 5.8 Hz, 1H), 3.08 – 2.99 (m, 2H).

5.2.7 *N*-Carbobenzoxy[3,5-¹³C₂-2,6-D₂]-L-tyrosine[P(O)(NMe₂)₂] benzyl ester **9**

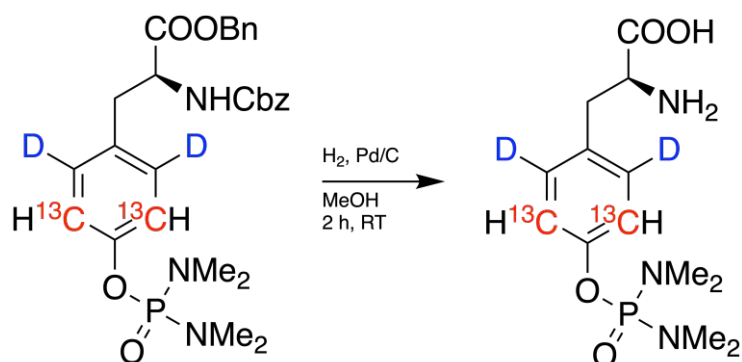


Synthesis of compound **9** was modified from literature.^[5] *N*-Carbobenzoxy[3,5-¹³C₂-2,6-D₂]-L-tyrosine benzyl ester **8** (150 mg, 0.30 mmol) was put under argon atmosphere and dissolved in absolute DCM. While cooling in an ice bath diazabicycloundecene (DBU) (1.5 eq), 4-dimethylaminopyridine (DMAP) (2.0 eq) and bis-(dimethylamino)phosphoryl chloride (2.0 eq) were added subsequently. The reaction mixture was stirred in the ice bath for 1 h. Afterwards, the ice bath was removed and the reaction mixture was stirred for another 90 min at RT. The reaction mixture was washed subsequently with water, 10% citric acid, saturated NaHCO₃ solution, and brine. Drying of the organic phase over MgSO₄ and evaporation of the solvents yielded the product as white crystals. A quick silica gel chromatography using DCM/methanol/formic acid (90:9:1) as eluent yielded 135 mg (90%) of *N*-carbobenzoxy[3,5-¹³C₂-2,6-D₂]-L-tyrosine[P(O)(NMe₂)₂] benzyl ester **9** as a white solid.

Rf-value 0.7 (DCM/methanol/formic acid 90:9:1)

ESI-MS (pos. mode) calculated for C₂₈H₃₅N₃O₆P [M+H]⁺ 540.2258; found 540.2258

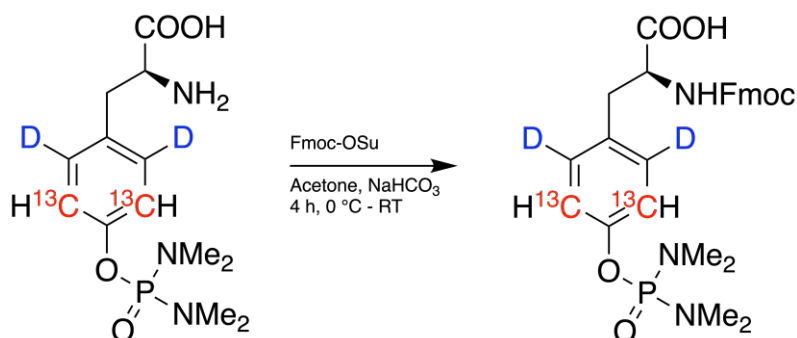
¹H NMR (500 MHz, CDCl₃) of **9** δ 7.35 (ddq, J = 16.1, 8.7, 4.6 Hz, 10H), 7.19 (d, J = 166.9 Hz, 2H), 6.93 (d, J = 8.1 Hz, 0H), 6.86 (s, 1H), 5.23 – 5.14 (m, 2H), 5.09 (s, 2H), 4.72 – 4.62 (m, 1H), 3.13 – 3.01 (m, 2H), 2.70 (d, J = 10.1 Hz, 12H).

5.2.8 [3,5-¹³C₂-2,6-D₂]-L-Tyrosine[P(O)(NMe₂)₂] **10**

Synthesis of compound **10** was modified from literature.^[5] Compound **9** (125 mg, 0.25 mmol) together with 10% Pd/C (0.2 eq) was put under argon atmosphere and dissolved in 10 ml dry Methanol. The flask was flushed with H₂ via a balloon and the reaction was stirred under H₂ atmosphere at 40°C for 2 h. When reaction completion was confirmed on TLC, Pd/C was removed via filtering over celite, and evaporation of the solvents gave 79 mg (100%) [3,5-¹³C₂-2,6-D₂]-L-Tyrosine[P(O)(NMe₂)₂] **10** as white crystals.

Rf-value 0.1 (ethyl acetate)

¹H NMR (500 MHz, D₂O) δ 6.96 (dd, J = 167.1, 4.7 Hz, 2H), 4.01 (dd, J = 7.9, 5.3 Hz, 1H), 3.29 – 3.08 (m, 2H), 2.69 (d, J = 10.4 Hz, 12H).

5.2.9 *N*-Fmoc[3,5-¹³C₂-2,6-D₂]-L-tyrosine[P(O)(NMe₂)₂] **11**

Synthesis of compound **11** was modified from literature.^[5] [3,5-¹³C₂-2,6-D₂]-L-Tyrosine [P(O)(NMe₂)₂] **10** (70 mg, 0.22 mmol) was dissolved in 5 ml 10% NaHCO₃ solution and cooled on an ice bath. Fmoc-OSu (78 mg, 0.23 mmol) was dissolved in acetone and added dropwise to the reaction mixture. After stirring for one hour on ice, stirring was continued for another 3 h at RT. Afterwards, acetone was removed under reduced pressure and the remaining aqueous phase was washed twice with diethylether. Then, the aqueous phase was acidified with 10% citric acid to a pH of 2 and extracted 5 times with ethyl acetate. Drying over MgSO₄ and subsequent removal of the solvents gave 115 mg (97%) of *N*-Fmoc[3,5-¹³C₂-2,6-D₂]-L-tyrosine[P(O)(NMe₂)₂] **11** as an off-white solid.

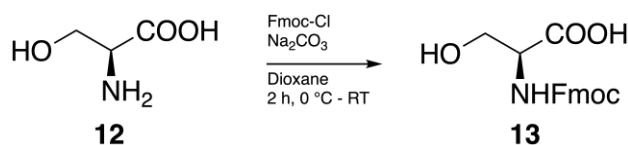
¹H NMR (500 MHz, CDCl₃) δ 7.78 (d, J = 7.6 Hz, 2H), 7.61 (t, J = 7.1 Hz, 2H), 7.41 (t, J = 7.5 Hz, 2H), 7.33 (t, J = 7.5 Hz, 2H), 7.08 (dd, J = 161.7, 4.6 Hz, 2H), 5.60 (d, J = 7.8 Hz, 1H), 4.70 (t, J = 8.0 Hz, 1H), 4.53 – 4.31 (m, 2H), 4.23 (t, J = 6.9 Hz, 2H), 3.18 (qd, J = 14.0, 5.0 Hz, 2H), 2.73 (dd, J = 10.2, 2.2 Hz, 12H).

¹³C NMR (500 MHz, CDCl₃) δ 120.1039.

ESI-MS (pos. mode): calculated for C₂₆¹³C₂H₃₁D₂N₃O₆P [M+H]⁺ 542.2294; found 542.2301.

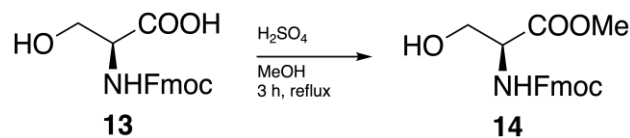
5.3 *N*- α -Fmoc-*O*-benzyl-L-phosphotyrosine 18

5.3.1 *N*-Fmoc-L-serine 13



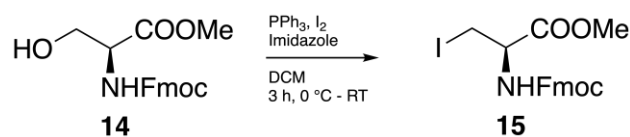
Synthesis of compound **13** was performed as described in literature.^[94] L-serine **12** (1.03 g 9.8 mmol) was dissolved in 30 ml of saturated Na_2CO_3 solution and cooled on an ice bath. 9-Fluorenylmethoxycarbonyl chloride (Fmoc-Cl) (2.45 g, 9.5 mmol) was dissolved in 30 ml 1,4-dioxane and dropwise added to the L-serine solution. Afterwards, the ice bath was removed and stirring was continued for 1.5 h at RT. Then, the mixture was washed with ethyl acetate to remove side products. The pH of the aqueous phase was set to 2-3 by addition of 1 M HCl and concentrated HCl and extracted thrice with ethyl acetate. The combined organic phases were washed with brine and dried over CaCl_2 . Evaporation of the solvents under reduced pressure yielded 2.91 g of *N*-Fmoc-L-serine **13** as a white solid (95% yield).

$^1\text{H NMR}$ (500 MHz, DMSO) δ 7.90 (d, $J = 7.5$ Hz, 2H), 7.75 (dd, $J = 7.6, 4.1$ Hz, 2H), 7.43 (t, $J = 7.7$ Hz, 2H), 7.34 (t, $J = 7.5$ Hz, 2H), 4.33 – 4.27 (m, 2H), 4.24 (t, $J = 7.2$ Hz, 1H), 4.07 (dt, $J = 8.8, 5.1$ Hz, 1H), 3.69 (d, $J = 5.0$ Hz, 2H), 3.18 (s, 1H).

5.3.2 *N*-Fmoc-L-serine methyl ester **14**

Synthesis of compound **14** was performed as described in literature.^[94] *N*-Fmoc-L-serine **13** (2.90 g, 8.85 mmol) was suspended in 30 ml methanol and 1 ml of concentrated H₂SO₄ was added. The reaction mixture was refluxed for 4 h. After cooling down to RT the pH of the mixture was set to 8 with saturated Na₂CO₃ solution and extracted thrice with ethyl acetate. The combined organic phases were washed with brine and dried over MgSO₄. The solvent was evaporated under reduced pressure to yield 2.68 g of *N*-Fmoc-L-serine methyl ester **14** as a white solid (89% yield).

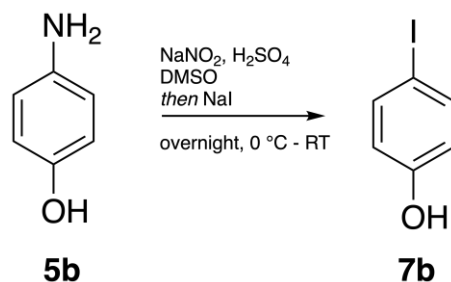
¹H NMR (500 MHz, DMSO) δ 7.87 (d, J = 3.6 Hz, 2H), 7.86 (d, J = 3.5 Hz, 2H), 7.73 (t, J = 6.6 Hz, 2H), 7.41 – 7.38 (m, 2H), 4.35 – 4.25 (m, 2H), 4.22 (t, J = 6.9 Hz, 1H), 3.72 – 3.64 (m, 2H), 3.62 (s, 3H)

5.3.3 *N*-Fmoc-3-iodo-L-alanine methyl ester **15**

Synthesis of compound **14** was modified from literature.^[94] Under an argon atmosphere, triphenylphosphine (2.16 g, 8.23 mmol) and imidazole (0.69 g, 10.15 mmol) were dissolved in dry DCM. The solution was cooled in an ice bath. Then, iodine (2.17 g, 8.58 mmol) was added. Then, the ice bath was removed, and the reaction mixture was stirred for 20 min at RT, whereas the mixture's yellow to brown color was observed. Then, the mixture was cooled on an ice bath again and *N*-Fmoc-L-serine methyl ester **14** (2.00 g, 5.80 mmol) in DCM was added dropwise. After 2.5 h of stirring on the ice bath, most of the solvent was removed under reduced pressure. Diethylether ether was added to precipitate the triphenylphosphine oxide and the mixture was filtered over celite and rinsed with diethylether. Subsequent evaporation of the solvents yielded 3.23 g of the crude product *N*-Fmoc-3-iodo-L-alanine methyl ester **15** as a pale yellow solid (87% yield).

Further purification using column chromatography on silica gel was not feasible since the desired compound gradually decomposed on silica gel. Hence, diethyl ether was added to allow POPh_3 to precipitate overnight at 4°C overnight, followed by subsequent removal of POPh_3 via filtration.

$^1\text{H NMR}$ (500 MHz, DMSO) δ 7.91 (d, $J = 7.5$ Hz, 2H), 7.74 (d, $J = 7.5$ Hz, 2H), 7.43 (t, $J = 7.4$ Hz, 2H), 7.34 (t, $J = 7.4$ Hz, 2H), 4.42 – 4.29 (m, 3H), 4.26 (t, $J = 7.0$ Hz, 1H), 3.67 (s, 3H), 3.57 – 3.33 (m, 2H).

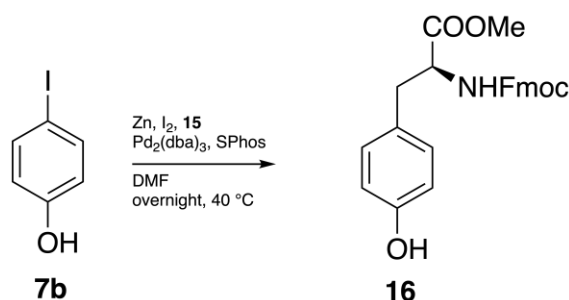
5.3.4 4-Iodophenol **7b**

Compound **7b** was synthesized as described for compound **7**. Aminophenol **5b** (9.17 mmol, 1.0 g) was dissolved in 10 ml DMSO and cooled via an ice-salt bath cooling mix to -5°C . Then 10 ml of 30% aqueous H_2SO_4 were added. When the internal temperature of the reaction reached 0°C , NaNO_2 (13.75 mmol, 948 mg) in 2 ml H_2O was added dropwise, whereas the internal temperature remained $< 5^\circ\text{C}$. After stirring for 1 h at this temperature, a solution of NaI (27.51 mmol, 4.12 g) in 2 ml H_2O was added dropwise. Stirring was continued for 1 h at RT. Then, a second portion of NaI (27.51 mmol, 4.12 g) was added, and the reaction was allowed to stir overnight at RT.

Then, the reaction was treated with 20 ml ethyl acetate and 20 ml of H_2O . The phases were allowed to separate, and the aqueous phase was extracted thrice with ethyl acetate. The combined organic phases were treated with 10% aqueous NaHSO_3 and stirred at RT for 30 min. The aqueous phase was separated and the organic phase was washed with saturated NaCl, and dried over MgSO_4 . The solvents were removed under reduced pressure to yield the crude product. Silica gel chromatography using heptane/ethyl acetate (7:3 v/v) yielded 1.57 g 4-Iodophenol **7b** as an off-white solid (78% yield).

Rf-value 0.7 (heptane/ethyl acetate 7:3)

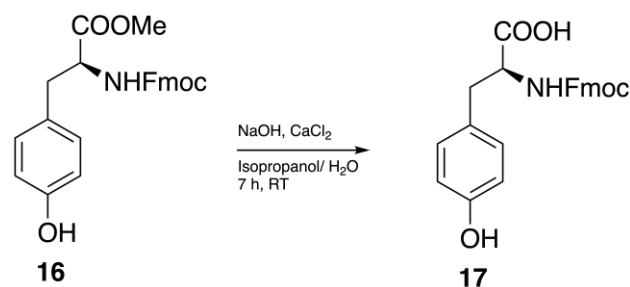
$^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.54 – 7.49 (m, 2H), 6.66 – 6.60 (m, 2H), 5.19 (br s, 1H).

5.3.5 *N*-Fmoc-L-tyrosine methyl ester **16**

Synthesis of compound **16** was modified from literature.^[4] Zn (2.77 mmol, 180 mg) was placed into a 3-neck flask, evacuated, and set under argon atmosphere. It was suspended in 3 ml of dry dimethylformamide (DMF) and I₂ (0.09 mmol, 24 mg) was added. *N*-Fmoc-3-iodo-L-alanine methyl ester **15** (1.6 g, 3.54 mmol) was dissolved in 3 ml abs. DMF was added dropwise to the suspension. Afterwards, I₂ (0.09 mmol, 24 mg) was added and the reaction mixture was stirred for 30 min at RT. Then, Pd₂(dba)₃ (0.02 mmol, 22 mg), S-Phos (0.04 mmol, 18 mg) and 4-iodophenol **7b** (0.94 mmol, 206 mg) were added sequentially. The reaction mixture was stirred overnight at 40°C. Then, it was cooled down to RT and 20 ml of ethyl acetate were added to the reaction mixture. The mixture was extracted with sat. NH₄Cl solution. The aqueous phase was extracted thrice with ethyl acetate. The combined organic phases were washed with sat. NaCl solution and dried over MgSO₄. Evaporation of the solvents under reduced pressure yielded the crude product. Silica gel chromatography using ethyl acetate/heptane (1:1 v/v) yielded 320 mg of *N*-Fmoc-L-tyrosine methyl ester **16** (82% yield).

R_f-value 0.3 (heptane/ethyl acetate 1:1)

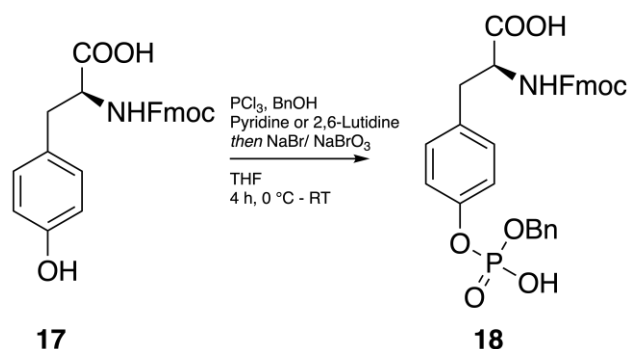
¹H NMR (500 MHz, DMSO) δ 7.90 (d, J = 7.5 Hz, 2H), 7.86 (d, J = 7.5 Hz, 2H), 7.43 (t, 2H), 7.36 (td, J = 7.4, 1.2 Hz, 2H), 7.06 – 6.93 (m, 2H), 6.66 (dd, J = 8.3, 3.9 Hz, 2H), 6.29 (s, 1H), 4.37 (d, J = 7.0 Hz, 2H), 3.57 (s, 3H), 2.96 – 2.64 (m, 2H).

5.3.6 *N*-Fmoc-L-tyrosine **17**

Synthesis of compound **17** was modified from literature.^[95] *N*-Fmoc-L-tyrosine methyl ester **16** (0.69 mmol, 290 mg) was suspended in a 0.8 M CaCl₂ isopropanol/H₂O (7:3 v/v) mixture. NaOH (0.83 mmol, 34 mg) was added to the reaction mixture, and the reaction mixture was stirred for 7 h. When reaction completion was confirmed by TLC, the mixture was neutralized with 1 M HCl and the solvents were evaporated to yield the crude *N*-Fmoc-L-tyrosine **17**. Silica gel chromatography using DCM/methanol/ACN (90:9:1 v/v/v) yielded 250 mg of *N*-Fmoc-L-tyrosine **17** (90% yield) as an off-white solid.

Rf-value 0.7 (DCM/methanol/ACN 90:9:1)

¹H NMR (500 MHz, DMSO) δ 9.39 (s, 1H), 7.87 (d, *J* = 7.5 Hz, 2H), 7.62 (dd, *J* = 12.2, 7.5 Hz, 2H), 7.43 – 7.35 (m, 2H), 7.30 (dt, *J* = 10.2, 7.3 Hz, 2H), 6.93 (d, *J* = 8.0 Hz, 1H), 6.71 (d, *J* = 8.1 Hz, 1H), 6.64 (d, *J* = 8.2 Hz, 2H), 6.27 (s, 1H), 4.22 – 3.87 (m, 2H), 2.97 (d, *J* = 14.9 Hz, 1H), 2.73 (dd, *J* = 13.7, 8.5 Hz, 1H).

5.3.7 *N*- α -Fmoc-*O*-benzyl-L-phosphotyrosine **18**

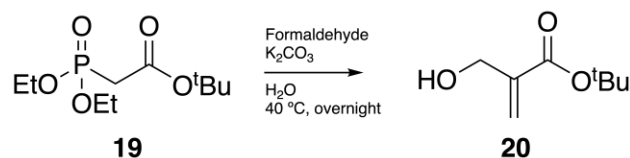
Synthesis of compound **18** was modified from literature.^[6] A 3-neck flask with a thermometer and a septum was set under argon atmosphere. PCl_3 (0.67 mmol, 92 mg, 1.3 eq) was dissolved in dry THF and cooled to 0°C . BnOH (0.78 mmol, 84 mg, 1.5 eq) was added to the reaction mixture, whereas the internal temperature remained at 0°C by cooling with an ice-NaCl cooling mixture. After 5 min, pyridine (1.59 mmol, 125 mg, 3.0 eq) was added, and a white slurry was formed. Fmoc-tyrosine **17** (0.52 mmol, 200 mg, 1.0 eq) was suspended in 3 ml dry THF and pyridine (0.52 mmol, 41 mg, 1.0 eq) was added. Then, this was added dropwise to the reaction mixture, whereas the internal temperature remained at 0°C . After 90 min of stirring in an ice bath, 1 ml of H_2O was added dropwise, maintaining the internal temperature $< 5^\circ\text{C}$. Then, NaBr (1.25 mmol, 130 mg, 2.4 eq) was added. When the NaBr was completely dissolved, 250 μl of 20 wt% NaBrO_3 solution was added. stirring was continued in the ice bath for 15 min. Then the ice bath was removed and the reaction mixture was stirred for 3 h at RT. To the yellow reaction mixture, a 10wt% $\text{Na}_2\text{S}_2\text{O}_5$ solution was added dropwise until the yellow reaction mixture turned clear. The reaction mixture was extracted thrice with 2-MeTHF. The combined organic phases were washed twice with sat. NaCl solution. Drying over MgSO_4 and subsequent evaporation of the solvents yielded the solvate containing the crude product. Purification via column chromatography using silica gel was not feasible due to the decomposition of compound **18** over silica. Purification by precipitation in heptane, ethyl acetate, diethyl ether, acetone, or DCM was not successful.

$^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.88 (d, $J = 7.7$ Hz, 2H), 7.64 (d, $J = 7.9$ Hz, 2H), 7.44 – 7.20 (m, 11H), 6.98 (d, $J = 8.0$ Hz, 2H), 6.61 (d, $J = 7.9$ Hz, 2H), 4.74 – 4.64 (m, 2H), 4.19 (ddt, $J = 30.9, 17.5, 7.3$ Hz, 4H), 2.94 (dd, $J = 13.9, 4.6$ Hz, 1H).

5.4 Isoleucine Precursor 24

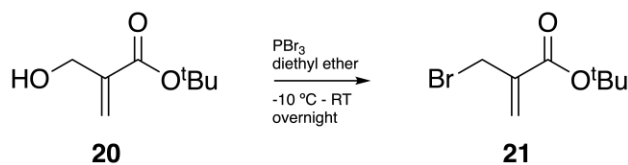
Compound **24** was synthesized as described in literature.^[7]

5.4.1 *tert*-Butyl 2-(hydroxymethyl)acrylate **20**



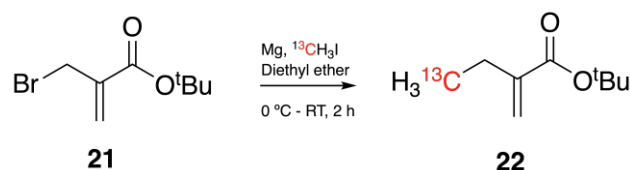
Paraformaldehyde (PFA) (9.78 g, 41 mmol) was dissolved in 50 ml H₂O, 1.05 ml of 1 M H₃PO₄ was added and the reaction mixture was refluxed to yield a formaldehyde solution. After the reaction was cooled to room temperature (RT), *tert*-butyl 2-(diethoxyphosphoryl)acetate **19** (20 g, 73 mmol) and K₂CO₃ (17.46 g, 12.64 mmol) in H₂O was added. The reaction mixture was stirred at 40°C for 18 h. Afterwards, diethylether and saturated NaCl solution were added. The aqueous phase was extracted twice with diethylether. The combined organic phases were washed with saturated NaCl solution and finally dried over MgSO₄. The crude product was purified by bulb to bulb distillation (10 mbar, up to 157°C), yielding 11.39 g *tert*-butyl 2-(hydroxymethyl)acrylate **20** (91% yield).

¹H NMR (500 MHz, CDCl₃) δ 6.15 (s, 1H), 5.74 (s, 1H), 4.29 (d, J = 4.3 Hz, 2H), 2.36 (d, J = 6.3 Hz, 1H), 1.51 (s, 9H).

5.4.2 *tert*-Butyl 2-(bromomethyl)acrylate **21**

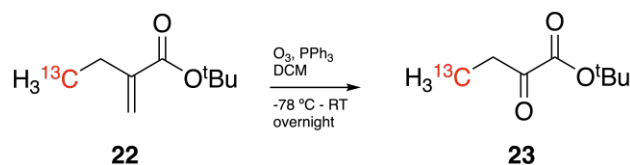
tert-Butyl 2-(hydroxy)acrylate **20** (11.39 g, 71 mmol) was dissolved in 20 ml of absolute diethylether and set under argon atmosphere. Then, the reaction mixture was cooled to -10°C using dry ice in acetone and PBr₃ (12.6 g, 46.5 mmol) was added via syringe. Afterwards, the reaction mixture was stirred for 3 h at RT followed by cooling to -10°C and quenching by the addition of 10 ml H₂O. Then, the aqueous phase was extracted three times with diethylether. The combined organic phases were washed with saturated NaCl solution and dried over MgSO₄. Evaporation of diethylether gave the crude product as a yellow liquid. Bulb to bulb distillation (15 mbar, up to 157°C) yielded 9.69 g *tert*-butyl 2-(bromomethyl)acrylate **21** (62% yield).

¹H NMR (500 MHz, CDCl₃) δ 6.23 (s, 1H), 5.86 (s, 1H), 4.14 (s, 2H), 1.52 (s, 9H).

5.4.3 [4-¹³C₂] *tert*-Butyl 2-methylenebutanoate **22**

Magnesium (108 mg, 4.45 mmol) was placed into a three-necked flask, evacuated on the vacuum line and set under argon atmosphere. Dry diethylether (10 ml) was added, followed by the dropwise addition of ¹³C-Iodomethan (5.4 mmol) until no more Mg was visible. The reaction mixture was heated to 40°C and stirred for 30 min until no more magnesium was visible anymore. Afterwards, the reaction mixture was cooled in an ice bath and *tert*-butyl 2-(bromomethyl)acrylate **21** (0.9 g, 4.05 mmol) was added dropwise. After 15 min the reaction mixture was allowed to warm to RT and stirring was continued for 2 h. Finally, quenching of the reaction was done by the addition of ice cubes and saturated NH₄Cl solution. Afterwards, the aqueous phase was extracted with diethylether four times. The combined organic phases were washed with saturated NaHCO₃ solution and H₂O and then dried over MgSO₄. Evaporation of the solvent under reduced pressure yielded 483 mg [4-¹³C₂] *tert*-butyl 2-methylenebutanoate **22** (80% yield).

¹H NMR (600 MHz, CDCl₃) δ 6.03 (s, 1H), 5.44 (s, 1H), 2.34 – 2.23 (m, 2H), 1.49 (s, 9H), 1.05 (dt, J = 126.4, 7.4 Hz, 3H).

5.4.4 [4-¹³C₂] *tert*-Butyl 2-ketobutanoate **23**

[4-¹³C₂] *tert*-Butyl 2-methylenebutanoate **22** (438 mg, 2.77 mmol) was dissolved in absolute dichloromethane (DCM) and cooled to -78°C using an acetone dry-ice cooling solution. Ozone was bubbled through the solution (100 ml/h, 70% ozone) for 10 min until the reaction mixture appeared blue. Afterwards, air was bubbled through to remove any ozone excess. Triphenylphosphine (931 mg, 3.55 mmol) was added to quench the reaction. After 30 min of stirring, the cooling bath was removed and the reaction mixture was allowed to be stirred at RT overnight. The next day, solvents were evaporated and the product was isolated using bulb-to-bulb distillation (7 mbar, up to 150°C), which yielded 306 mg [4-¹³C₂] *tert*-butyl 2-ketobutanoate **23** as a yellow oil (69% yield).

¹H NMR (500 MHz, CDCl₃) δ 2.84 – 2.75 (m, 2H), 1.54 (s, 9H), 0.97 (dt, J = 134.5, 7.2 Hz, 3H).

5.4.5 [4-¹³C₂, 3-D₂] 2-Ketobutanoic acid **24**

[4-¹³C₂] *tert*-Butyl 2-ketobutanoate **23** (300 mg, 1.87 mmol) was dissolved in 10 ml of absolute diethylether and 10 ml of DCM. Gaseous HCl was bubbled through the reaction mix which was stirred at 0 °C using an ice-water bath. Thereby concentrated HCl was added to concentrated H₂SO₄ in a Schlenk flask using a dropping funnel. After 15 min of gaseous HCl treatment, the flask was tightly closed and stirred for 90 min at RT. TLC showed the hydrolysis to be completed. The solvent was then evaporated at 450 mbar. Further purification was done via bulb-to-bulb distillation (30 mbar, up to 135 °C). The oil was then dissolved in D₂O and the pH was set to 14 using 0.5 M NaOD. The solution was then stirred for 2 h under argon atmosphere at RT. Afterwards, the pH of the reaction mix was set to 1 by the addition of 6 M DCl and 5 times extracted with diethylether. The diethylether was removed at 500 mbar and the pH was set to 7 by the addition of 0.5 M NaOD and 5 M DCl. Lyophilization yielded 180 mg [4-¹³C₂, 3-D₂] 2-Ketobutanoic acid **24** as a white solid (92% yield).

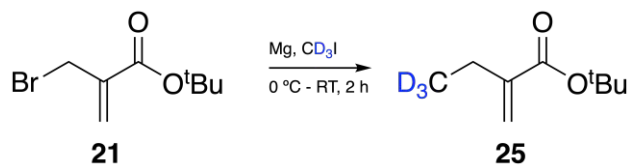
¹H NMR of the free acid before deuteration (500 MHz, D₂O) δ 4.70 (s, 1H), 2.49 – 2.47 (m, 2H), 0.92 (dt, J = 124.2, 1.7 Hz, 3H).

¹H NMR (500 MHz, D₂O) δ 0.90 (d, J = 128.3 Hz, 1H).

5.5 Valine and Leucine Precursor 30

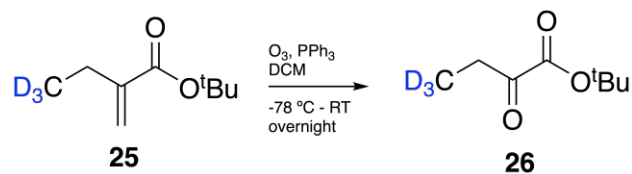
Compound **30** was synthesized as described in literature.^[7]

5.5.1 [4-CD₃] *tert*-Butyl 2-methylenebutanoate **25**



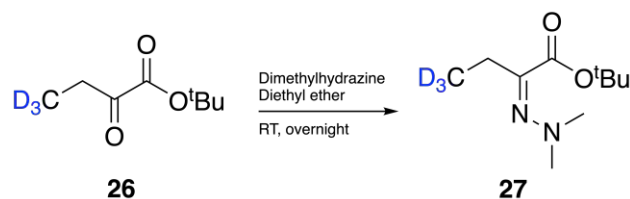
The synthesis was performed according to the preparation of [4-¹³C₂] *tert*-butyl 2-methylene butanoate **22**, using *tert*-butyl 2-(bromomethyl)acrylate **21** (7.11 mmol, 1.58 g), Mg (7.82 mmol, 191 mg) and CD₃-Iodomethan (1.14 g, 7.86 mmol). The reaction yielded 923 mg [4-CD₃] *tert*-butyl 2-methylenebutanoate **25** as a clear oil (78% yield).

¹H NMR (600 MHz, CDCl₃) δ 6.03 (s, 1H), 5.44 (s, 1H), 2.28 (dd, J = 12.7, 7.0 Hz, 2H), 1.49 (s, 9H).

5.5.2 [4-CD₃] *tert*-Butyl 2-ketobutanoate **26**

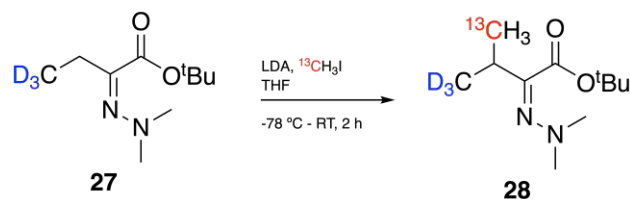
The synthesis was performed according to the synthesis of [4-¹³C₂] *tert*-butyl 2-ketobutanoate **23**, yielding 542 mg [4-CD₃]-*tert*-butyl 2-ketobutanoate **26** as a clear liquid (58% yield).

¹H NMR (500 MHz, CDCl₃) δ 2.77 (s, 2H), 1.53 (s, 9H).

5.5.3 [4-CD₃] *tert*-Butyl 2-(2,2-dimethylhydrazono)butanoate **27**

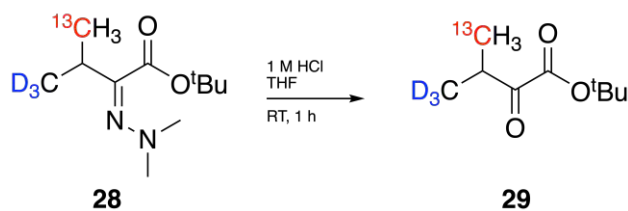
[4-CD₃] *tert*-butyl 2-ketobutanoate **26** (3.29 mmol, 530 mg) was dissolved in 2 ml of dry diethylether under argon atmosphere. Afterwards, dimethylhydrazine (300 μ l, 3.93 mmol) was added dropwise via a syringe. Stirring was continued overnight. Then, the reaction mixture was diluted with 20 ml of diethylether, washed with H₂O, and dried over MgSO₄. Evaporation of the solvent under reduced pressure yielded 330 mg [4-CD₃] *tert*-butyl 2-(2,2-dimethylhydrazono)butanoate **27** as a yellow liquid (48% yield).

¹H NMR (500 MHz, CDCl₃) δ 2.77 and 2.50 (s, 6H), 2.53 and 2.35 (s, 2H), 1.52 and 1.51 (s, 9H).

5.5.4 [4-¹³C; 4-CD₃] *tert*-Butyl 2-(2,2-dimethylhydrazono)-3-methylbutanoate **28**

Under argon atmosphere, 10 ml of dry tetrahydrofuran (THF) was mixed with diisopropylamine (320 μ l, 2.28 mmol) and cooled to -78°C using acetone in dry ice. Carefully, 1 ml of *n*-butyllithium (nBuLi) (2.5 M in hexane) was added dropwise by syringe. Stirring was continued for 30 min at 0°C using an ice-water bath, followed by cooling to -78°C. Then, [4-CD₃] *tert*-butyl 2-(2,2-dimethylhydrazono)butanoate **27** (330 mg, 1.63 mmol) was added dropwise, whereas the reaction mixture changed color from yellow to brown, indicating enolate formation. Stirring was performed for 1 h, followed by the addition of [¹³C]-iodomethane (2.43 mmol, 155 μ l). The reaction mixture was stirred for 2 h, and then quenched by the addition of H₂O, whereas the reaction mixture changed its color to bright yellow. Afterwards, the reaction was allowed to warm to RT, the solvents were evaporated under reduced pressure. The residue was taken up in ethyl acetate and washed with H₂O until the aqueous phases appeared clear. The organic phase was dried over MgSO₄ and the solvent was removed under reduced pressure, yielding 303 mg of [4-¹³C; 4-CD₃] *tert*-butyl 2-(2,2-dimethylhydrazono)-3-methylbutanoate **28** (85% yield).

¹H NMR (500 MHz, CDCl₃) δ 2.63 – 2.58 (m, 1H), 2.46 (s, 6H), 1.51 (s, 9H), 1.01 (dd, *J* = 125.3, 6.9 Hz, 1H).

5.5.5 [4-¹³C; 4-CD₃] *tert*-Butyl 3-methyl-2-ketobutanoate **29**

[4-¹³C; 4-CD₃] *tert*-Butyl 2-(2,2-dimethylhydrazono)-3-methylbutanoate **28** (303 mg, 1.38 mmol) was dissolved in THF and 5 ml 1 M HCl was added. The reaction mixture was stirred for 2 h at RT and afterwards diluted with diethyl ether. The aqueous phase was extracted thrice with diethylether, the combined organic phases were washed twice with saturated NaCl solution and dried over MgSO₄. Evaporation of the solvents yielded 176 mg [4-¹³C; 4-CD₃] *tert*-butyl 3-methyl-2-ketobutanoate **29** (72% yield).

¹H NMR (500 MHz, CDCl₃) δ 3.18 – 3.10 (m, 1H), 1.53 (s, 9H), 1.26 and 0.99 (two d, J = 130.4, 7.0 Hz, 3H).

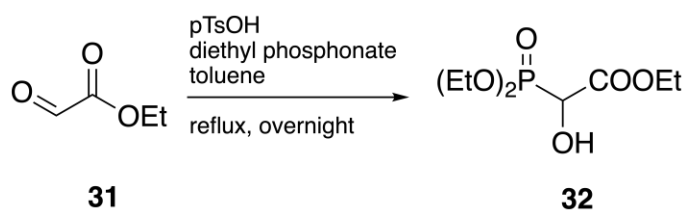
5.5.6 [4-¹³C; 4-CD₃] Ketoisovaleric acid **30**

Compound **30** (176 mg, 1 mmol) was prepared as described for **24**. However, ¹H NMR revealed the remaining *tert*-butyl ester after lyophilization. Therefore, the mixture of product and substrate was refluxed in 6 M HCl for 2 h and subsequently extracted 6 times with diethyl ether. Further workup was performed as described for **24** yielding 103 mg of [4-¹³C; 4-CD₃] Ketoisovaleric acid **30** (65% yield).

¹H NMR (500 MHz, D₂O) δ 1.21 (d, J = 127.8 Hz, 3H).

5.6 Histidine Precursor

5.6.1 Ethyl 2-(diethoxyphosphoryl)-2-hydroxyacetate **32**

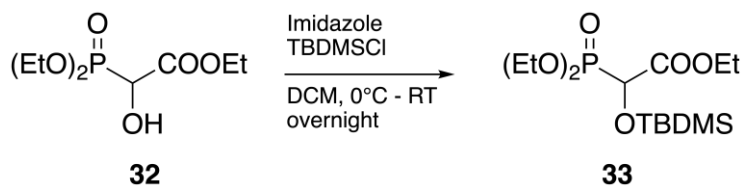


Compound **32** was synthesized according to a procedure from the literature.^[9] Ethyl glyoxylate **31** (1.00 g, 9.8 mmol) was added to 10 ml toluene. Then, p-toluene sulfonic acid (16 mg) and diethyl phosphonate (1.62 g, 11.8 mmol) were added. The reaction mixture was refluxed overnight at 135°C. The next day, after cooling down, the solvent was removed under reduced pressure, yielding the crude product. Subsequent silica gel chromatography using DCM/methanol (79:1 v/v) yielded 1.54 g of the purified product **32** as a colorless oil (65% yield).

Rf-value 0.4 (DCM/methanol 79:1)

¹H NMR (500 MHz, CDCl₃) δ 4.53 (d, J = 15.7, 1H), 4.37 – 4.27 (m, 2H), 4.26 – 4.16 (m, 4H), 1.37 – 1.29 (m, 9H).

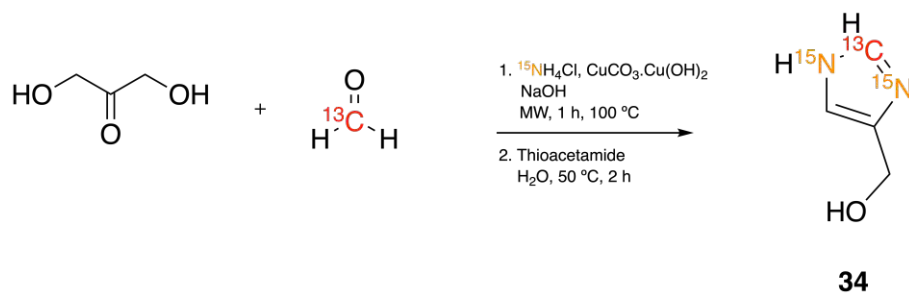
5.6.2 Ethyl 2-((*tert*-butyldimethylsilyl)oxy)-2-(diethoxyphosphoryl)acetate **33**



Compound **33** was synthesized according to a procedure from the literature.^[9] Ethyl 2-(diethoxyphosphoryl)-2-hydroxyacetate **32** (0.90 g, 3.8 mmol) was dissolved in 20 ml DCM and cooled to 0°C using an ice-water bath. Imidazole (0.49 g, 7.2 mmol) and TBDMSCl (840 mg, 5.58 mmol) were added subsequently. The reaction mixture was stirred and allowed to warm to RT overnight. The next day, the reaction was quenched with saturated NH₄Cl solution and the aqueous phase was extracted 5 times with DCM. The combined organic phases were dried over MgSO₄ and subsequent evaporation of the solvents yielded the crude product. Silica gel chromatography using heptane/ethyl acetate (1:1 v/v) yielded 0.84 g of the purified product **33** as a colorless oil (63% yield).

Rf-value 0.4 (heptane/ethyl acetate 1:1)

¹H NMR (500 MHz, CDCl₃) δ 4.59 (d, J = 18.0 Hz, 1H), 4.32 – 4.10 (m, 6H), 1.37 – 1.28 (m, 9H), 0.92 (s, 9H), 0.11 (s, 3H), 0.10 (s, 3H).

5.6.3 [1,3-¹⁵N, 2-¹³C](Imidazol-1*H*-5-yl)methanol **34**

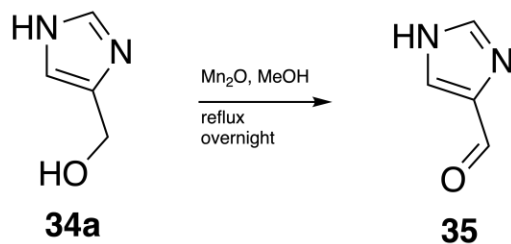
CuCO₃·(OH)₂ (1.44 g, 6.5 mmol), NH₄Cl (1.39 g, 26 mmol) and 1,3-dihydroxyacetone dimer (1.16 g, 6.45 mmol) were added to a microwave vessel (V = 10.0 – 20.0 ml). Formaldehyde solution (37% wt, 0.55 ml) and 6 ml H₂O were added. The pH was set to 12 by the addition of 1 M NaOH and the microwave vessel was set under argon atmosphere. A microwave reaction was performed for 1 h at 100°C. After reaction completion, the microwave vessel was stored at 4°C overnight. Then, the brown copper complex was filtered off and thoroughly washed with ice-cold H₂O. The precipitate was added with thioacetamide (360 mg, 8.4 mmol) in 20 ml of H₂O. The reaction mixture was stirred at 50°C for 2 h. Then, the resulting CuS was filtered off from the hot reaction mixture and washed with H₂O. The filtrate was evaporated under reduced pressure, yielding a brown solid crude product. Subsequent silica gel chromatography using methanol/ethyl acetate/NH₄OH (28%) (35:65:5 v/v/v) yielded 480 mg of 1(3)H-Imidazol-4-yl-methanol (76% yield) as a yellow solid.

Synthesis for the labeled compound was performed as described using ¹⁵NH₄Cl as nitrogen source and a 20% wt ¹³C-Formaldehyde solution, however, no product **34** could be obtained.

Alternatively, 200 mg of ¹³C-paraformaldehyde were stirred overnight at RT in 4 ml of 0.5 M NaOH solution. Then, it proceeded as described yielding 390 mg of compound **34** as an off-white solid (62% yield.) ¹H NMR revealed a carbon-13 content on the epsilon position of 60%.

Rf-value 0.6 (methanol/ethyl acetate/NH₄OH (28%) (35:65:5))

¹H NMR (500 MHz, D₂O) δ 8.30 (t, J = 6.5 Hz, 0.4H), 8.08 (dt, J = 216.2, 6.9 Hz, 0.6H), 7.29 (s, 1H), 4.64 (d, J = 2.3 Hz, 2H).

5.6.4 *1H*-Imidazole-5-carbaldehyde **35**

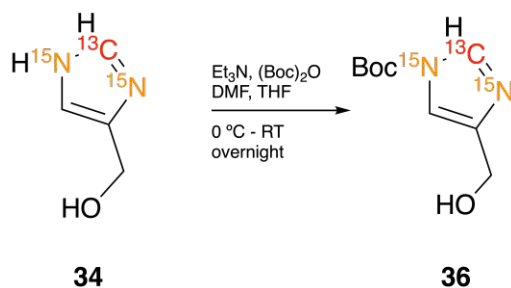
Synthesis of compound **35** was modified from literature.^[9] *1(3)H*-Imidazol-4-ylmethanol (200 mg, 2 mmol) was dissolved in 10 ml dry methanol and treated with 5 eq. MnO₂. The reaction mixture was refluxed overnight. Thereafter, MnO₂ was filtered off over celite. The solvent was removed under reduced pressure to yield the crude product. Column chromatography using ethyl acetate/methanol (4:1 v/v) yielded *1H*-imidazole-5-carbaldehyde (80% yield).

¹H NMR (500 MHz, DMSO) δ 9.74 (s, 1H), 7.99 (s, 1H), 7.92 (s, 1H).

5.6.5 *tert*-Butyl 4-formyl-1*H*-imidazole-1-carboxylate **37a** Route A

Synthesis of compound **37a** was modified from literature.^[9] Imidazole-carbaldehyde **35** (100mg, 1 mmol) was dissolved in 3 ml dry ACN and a catalytic amount of DMAP (15 mg, 0.10 mmol) was added. Then, $(\text{Boc})_2\text{O}$ (280 μl , 1.34 mmol) in 3 ml dry ACN was added dropwise. The reaction mixture was stirred overnight at RT. Evaporation of the solvents the next day yielded dark yellow solid. Purification was performed by silica gel chromatography using heptane/ethylacetate (1:1 v/v) and yielded 40 mg of **37** as white crystals (23% yield). Further investigation revealed that the compound was not stable on silica.

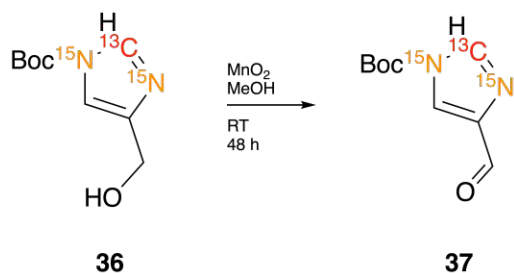
$^1\text{H NMR}$ (500 MHz, CDCl_3) δ 9.91 (s, 1H), 8.12 (s, 1H), 8.01 (s, 1H), 1.63 (s, 9H).

5.6.6 *tert*-Butyl-[1,3-¹⁵N, 2-¹³C]imidazole-4-yl-methanol **36**

Compound **34** (200 mg, 2.0 mmol) was put under argon atmosphere and dissolved in 4 ml dry DMF. Triethylamine (252 mg, 2.5 mmol, 0.35 ml) in DMF was added dropwise while the reaction mixture was cooled in an ice bath. Then, Boc₂O (545 mg, 2.5 mmol) in 2 ml dry THF was added dropwise. The reaction mixture was allowed to warm to RT and stirred overnight.

Water was added and the aqueous phase was extracted with ethyl acetate several times. The combined organic phases were washed with brine and H₂O, dried over MgSO₄ and the solvents were removed under reduced pressure to yield 298 mg of **36** as white to yellow crystals (74% yield).

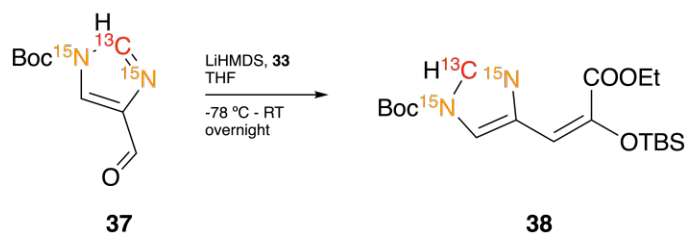
¹H NMR (500 MHz, CDCl₃) δ 8.20 (dt, J = 213.4, 8.4 Hz, 0.75H), 8.00 (t, J = 8.4 Hz, 0.25H), 7.07 (s, 1H), 4.80 (s, 2H), 1.61 (s, 9H).

5.6.7 *tert*-Butyl-[1,3-¹⁵N, 2-¹³C]imidazole-1-carboxylate **37** Route B

Compound **36** (270 mg, 1.38 mmol) was dissolved in methanol and MnO₂ (602 mg, 6.91 mmol) were added. The reaction mixture was allowed to stir for 48 h. The next day MnO₂ was filtered off over celite and washed with methanol. The solvent from the filtrate was evaporated and white to yellow crystals remained. Purification via silica gel chromatography could not be performed, since the compound remained unstable on silica. The side product was partially removed by its precipitation using acetone, which yielded 230 mg of **37** as off-white crystals (85% yield).

¹H NMR (500 MHz, CDCl₃) δ 9.94 (s, 1H), 8.35 (dt, J = 217.5, 9.3 Hz, 0.75H), 8.12 (t, J = 9.8 Hz, 0.25H), 8.03 (d, J = 3.0 Hz, 1H), 1.65 (s, 9H).

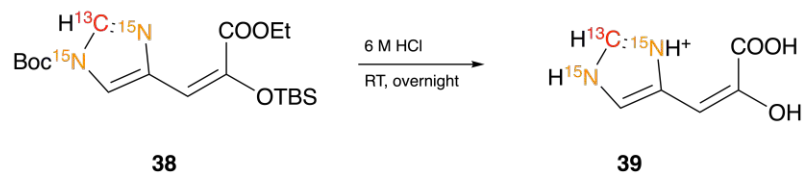
5.6.8 *tert*-Butyl-4-(2-((*tert*-butyldimethylsilyl)oxy)-3-ethoxy-3-oxoprop-1-en-1-yl)-[1,3-¹⁵N, 2-¹³C]-imidazole-1-carboxylate **38**



Synthesis of **38** was modified from literature.^[58] Compound **33** (0.75 mmol, 272 mg) was placed under argon atmosphere and dissolved in dry THF. The mixture was cooled to -78°C using dry ice in acetone. LiHMDS solution (1 M in THF, 0.75 mmol, 750 μl) was added dropwise and the reaction mixture was stirred for 30 min at -78°C . Compound **37** (0.6 mmol, 120 mg) was dissolved in dry THF and added dropwise to the reaction mixture. Stirring was continued overnight, whereas warming to RT was allowed. The reaction was quenched with saturated NH_4Cl the next day. The aqueous phase was extracted six times with DCM. The combined organic phases were dried over MgSO_4 , and the solvent was removed under reduced pressure yielding brown oil as a crude product. Purification was performed by silica gel chromatography using heptane/ethyl acetate (4:1 v/v), yielding 85 mg of **38** as a clear oil (38% yield).

Rf-value 0.6 (heptane/ethyl acetate 4:1)

¹H NMR of the *Z*-isomer of **38** (500 MHz, CDCl_3) δ 8.23 (dddd, $J = 220.5, 11.6, 7.4, 1.3$ Hz, 0.6H), 8.08 (dd, $J = 4.5, 2.6$ Hz, 1H), 8.01 (ddd, $J = 11.5, 7.4, 1.3$ Hz, 0.4H), 6.39 (d, $J = 2.4$ Hz, 1H), 4.28 (q, $J = 7.2$ Hz, 2H), 1.62 (s, 9H), 1.35 (t, $J = 7.2$ Hz, 3H), 0.97 (s, 9H), 0.19 (s, 6H).

5.6.9 [1,3-¹⁵N, 2-¹³C]4-(2-Carboxy-2-hydroxyvinyl)-1*H*-imidazolium chloride **39**

Synthesis of compound **39** was performed as described in literature.^[9] Compound **38** (80 mg, 0.21 mmol) was dissolved in 6 M HCl (10 ml) and stirred overnight at RT. Evaporation of the solvents yielded 40 mg of 4-(2-carboxy-2-hydroxyvinyl)-1*H*-imidazolium chloride as a beige solid (100% yield).

¹H NMR (500 MHz, DMSO) δ 14.59 (d, $J = 90.5$ Hz, 2H), 10.31 (s, 1H), 9.28 (dt, $J = 220.5, 5.5$ Hz, 0.7H), 9.06 (td, $J = 5.8, 1.4$ Hz, 0.3H), 7.69 (p, $J = 4.1$ Hz, 1H), 6.43 (d, $J = 2.1$ Hz, 1H).

Abbreviations

2D-NMR two-dimensional NMR

3D-NMR three-dimensional NMR

Bn Benzyl group

Boc *tert*-Butyloxycarbonyl protecting group

Cbz Carboxybenzyl group

CF Cell-free

COSY 2D correlated spectroscopy

CSD Chemical Shift Anisotropy

CSP Chemical Shift Perturbations

DBU diazabicycloundecene

DCM dichlormethane

DD dipole-dipole

DMAP 4-dimethylaminopyridine

DMF dimethylformamide

ds double-stranded

E. coli Escherichia coli

EM electron microscopy

FID free induction decay

FMOC-Cl 9-Fluorenylmethoxycarbonyl chloride

HWE Horner-Wadsworth-Emmons

HSQC heteronuclear single quantum correlation

K_D equilibrium dissociation constant

LB lysogeny broth

ABBREVIATIONS

LDA lithiumdiisopropylamin

MeOH methanol

nBuLi n-butyllithium

NMR Nuclear Magnetic Resonance

NOE Nuclear Overhauser Effect

NOESY Nuclear Overhauser Effect Spectroscopy

PFA Paraformaldehyde

PLCG Phospholipase C γ

pHPLC preparative high performance liquid chromatography

PTB phosphotyrosine binding

PTM Posttranslational modification

PTMs Posttranslational modifications

pTyr phosphorylated Tyrosine

RT room temperature

SAIL Stereo-array Isotopic Labeling

SAR structure-activity relationships

SH2 Src homology 2

SPPS solid-phase peptide synthesis

STD Saturation Transfer Difference

T₂ transverse relaxation time

THF tetrahydrofuran

TLC thin layer chromatography

TROSY Transverse relaxation-optimized spectroscopy

List of Figures

1	Distribution of experimental methods used for protein structure determination of proteins. ^[11]	1
2	NMR spectroscopy with small and large molecules in solution. A) The NMR signal obtained from small molecules in solution relaxes slowly; it has a long transverse relaxation time (T_2). A large T_2 value translates into narrow line widths ($\Delta\nu$) in the NMR spectrum after Fourier transformation (FT) of the NMR signal. B) By contrast, for larger molecules, the decay of the NMR signal is faster (T_2 is smaller). This results in both in a weaker signal measured after the NMR pulse sequence and in broad lines in the spectra. C) Using TROSY, the transverse relaxation can be substantially reduced, which results in improved spectral resolution and improved sensitivity for large molecules; <i>taken from</i> ^[27]	3
3	Arrows indicate NOE interactions that can be observed in polypeptide chains.	5
4	Magnetization transfer paths of the following 3D-NMR methods: HNCA, HN(CO)CA and CBCA(CO)NH; <i>adapted from</i> ^[31]	6
5	CSP for the BRD4-BD1 protein upon binding to a ligand; <i>taken from</i> ^[38]	7
6	Fmoc-pTyr building blocks: A) Fmoc-Tyr(PO_3H_2)-OH, B) Fmoc-Tyr($\text{PO}(\text{OBzl})\text{OH}$)-OH, C) Fmoc-Tyr($\text{PO}(\text{NMe}_2)_2$)-OH	16
7	A) general reaction scheme of a Negishi cross-coupling reaction and B) proposed mechanism for the Negishi coupling; <i>adapted from</i> ^[90]	19
8	Fmoc-phosphoTyrosine building blocks 11 and 18 containing an isolated ^{12}C -H spin system	20
9	Aliphatic amino acid precursors for isoleucine 24 and valine and leucine 30	20
10	A new isotopologue of imidazole pyruvic acid 39 - a metabolic precursor for histidine	21
11	^1H NMR spectra of compounds 5 (top), 6 (middle) and 7 (bottom) in CDCl_3 revealing deuteration grade of 6 of 80%. <i>Top:</i> Signal E corresponds to the protons on positions 2 and 6 of compound 5 . <i>Middle:</i> Signal E corresponds to the protons on positions 2 and 6 of compound 6 , revealing a deuteration level of 90% on positions 2 and 6. <i>Bottom:</i> Signal D corresponds to the protons on positions 2 and 6 of compound 7 , which reveals no change in deuteration level from 6 to 7 .	23
12	The structure of 11 with atom labels is shown (<i>top</i> , referring to the ^1H spectra on the <i>bottom</i>).	25

13	^1H NMR spectra of 18 from the synthesized crude product (<i>top</i>) and the commercial sample (<i>bottom</i>)	27
14	Retrosynthetic design of Fmoc-protected spin-isolated amino acids; <i>adapted from</i> ^[4]	28
15	[4- ^{13}C ; 4- CD_3] Ketoisovaleric acid 30	30
16	[4- $^{13}\text{C}_2$, 3- D_2] 2-Ketobutanoic acid 24	30
17	^1H NMR spectrum of 34 picturing the incorporation of nitrogen-15 and carbon-13 by the depicted couplings.	33
18	^1H NMR spectrum showing the Enol-product of 34 with a carbon-13 content of 70%	34
19	Isotopically labeled pTyr moiety in protein	35
20	Protein with selectively labeled histidine	35

List of Schemes

1	Biosynthesis of the aliphatic amino acids in <i>E. coli</i> : A) isoleucine biosynthesis; B) valine and leucine biosynthesis; <i>adapted from</i> ^[57]	11
2	Histidine metabolism in <i>E. coli</i> : A) showing the histidine metabolism from the pentose phosphate pathway; B) and the major and minor degradation pathway in <i>E. coli</i> with the reversible conversion of histidine to imidazole-pyruvic acid. <i>from</i> ^[9,57]	13
3	One-pot synthesis of protected phosphorylated amino acids; <i>adapted from</i> ^[6]	17
4	Introduction of <i>N,N'</i> -dimethyldiaminophosphinoyl moiety to the tyrosine side-chain; <i>adapted from</i> ^[5]	17
5	Synthesis of <i>N</i> - α -Fmoc- <i>O</i> -(bis-dimethylaminophosphono)-[3,5- $^{13}\text{C}_2$ -2,6- D_2]- L-tyrosine 11 : <i>a)</i> PPh_3 , Imidazole, I_2 , DCM, 0°C - RT, 87%; <i>b)</i> Nitromalonaldehyde, H_2O , NaOH, 6 d, 4°C , 54%; <i>c)</i> H_2 balloon, Pd/C, MeOH, 2 h, RT, quant.; <i>d)</i> D_2O , HCl, Microwave 37 min, 180°C , quant.; <i>e)</i> 1) NaNO_2 , H_2SO_4 , DMSO, 1 h, $0 - 5^\circ\text{C}$, 2) NaI, overnight, RT, 44%; <i>f)</i> 2 , Zn, I_2 , $\text{Pd}_2(\text{dba})_3$, SPhos, DMF, overnight, 40°C , 81%; <i>g)</i> DBU, DMAP, bis-(dimethylamino)phosphorylchloride, DCM, 2 h, 0°C - RT, 90%, <i>h)</i> H_2 , Pd/C, MeOH, 2 h, 40°C , quant.; <i>i)</i> Fmoc-OSu, Acetone, NaHCO_3 , 4 h, 0°C - RT, 75%	22

- 6 Synthesis of *N*- α -Fmoc-*O*-benzyl-L-phosphotyrosine **18**: *a*) Fmoc-Cl, Na₂CO₃, Dioxane, 2 h, 0°C - RT, 95%; *b*) H₂SO₄, MeOH, 3 h, reflux, 89%; *c*) PPh₃, Imidazole, I₂, DCM, 0°C - RT, 87%; *d*) 1) NaNO₂, H₂SO₄, DMSO, 1 h, 0 - 5°C, 2) NaI, overnight, RT, 78%; *e*) **15**, Zn, I₂, Pd₂(dba)₃, SPhos, DMF, overnight, 40°C, 82%; *f*) NaOH/CaCl₂, Isopropanol/H₂O, 7 h, RT, 90%, *g*) 1) PCl₃, BnOH, 2,6-Lutidine or Pyridine, THF, 90 min 0 - 5°C, 2) NaBr/NaBrO₃, H₂O; not purified from crude 26
- 7 Synthesis of the aliphatic amino acid precursors [4-¹³C₂, 3-D₂] 2-ketobutanoic acid **24** and [4-¹³C; 4-CD₃] ketoisovaleric acid **30**: *a*) formaldehyde, K₂CO₃, H₂O, overnight, 40°C, 91%; *b*) PBr₃, diethyl ether, - 10°C - RT, 62%; *c*) Mg, ¹³CH₃I, diethyl ether, 80%; *d*) O₃, DCM, PPh₃, overnight, - 78°C - RT, 69%; *e*) 1) HCl(g), DCM/ diethyl ether, 1 h, 0°C - RT, 2) 0.5 M NaOD, D₂O, 2 h, RT, 92%; *f*) Mg, CD₃I, diethylether, 78%, *g*) = *d*), 58%; *h*) dimethylhydrazine, diethyl ether, overnight, RT, 78%; *i*) diisopropylamine, nBuli, ¹³CH₃I, THF, 2 h, - 78°C - RT, 85%; *j*) 1 M HCl, THF, 1 h, RT, 72%; *k*) = *e*), then 6 M HCl, 2 h, reflux, 0.5 M NaOD, D₂O, 65% 29
- 8 Synthesis of [1,3-¹⁵N, 2-¹³C]4-(2-carboxy-2-hydroxyvinyl)-1*H*-imidazolium chloride **39**: *a*) pTsOH, diethyl phosphonate, toluene, reflux, overnight, 65%; *b*) imidazole, TBDMSCl, DCM, 0°C - RT, overnight, 63%; *c*) ¹⁵NH₄Cl, CuCO₃.Cu(OH)₂, MW 100°C, 1 h, *then* thioacetamide, H₂O, 50°C, 2 h, 62%; *d*) MnO₂, MeOH, reflux, overnight, 80%; *e*) DMAP, (Boc)₂O, ACN, 0°C - RT, overnight, 23%; *f*) (Boc)₂O, Et₃N, DMF, THF, 0°C - RT, overnight, 74%; *g*) MnO₂, MeOH, RT, 48 h, 84%; *h*) LiHMDS, **33**, THF, -78°C - RT, overnight, 58%; 38%; *i*) 6 M HCl, RT, overnight, quant. 31

List of Tables

- 1 SAIL Strategy 14
- 2 Synthesis of (imidazole-1*H*-5-yl)-methanol: NaOH: +) NaOH was added up to a pH 12, -) no NaOH was added; work-up: *i*) precipitation of CuS using thioacetamide for 2 h at 50°C, *ii*) precipitation of CuS using thioacetamide for 2 h at 50°C with the addition of 1 M HCl 32

References

- [1] J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, N. J. Skelton, *Protein NMR spectroscopy: principles and practice*, Academic press, **1996**.
- [2] S.-y. Ohki, M. Kainosho, *Progress in Nuclear Magnetic Resonance Spectroscopy* **2008**, *53*, 208–226.
- [3] R. Lichtenecker, *Organic & Biomolecular Chemistry* **2014**, *12*, 7551–7560.
- [4] B. M. Young, P. Rossi, P. J. Slavish, Y. Cui, M. Sowaileh, J. Das, C. G. Kalodimos, Z. Rankovic, *Organic letters* **2021**, *23*, 6288–6292.
- [5] H.-G. Chao, B. Leiting, P. D. Reiss, A. L. Burkhardt, C. E. Klimas, J. B. Bolen, G. R. Matsueda, *The Journal of Organic Chemistry* **1995**, *60*, 7710–7711.
- [6] D. E. Petrillo, D. R. Mowrey, S. P. Allwein, R. P. Bakale, *Organic letters* **2012**, *14*, 1206–1209.
- [7] R. Lichtenecker, M. L. Ludwiczek, W. Schmid, R. Konrat, *Journal of the American Chemical Society* **2004**, *126*, 5348–5349.
- [8] S.-M. Liao, Q.-S. Du, J.-Z. Meng, Z.-W. Pang, R.-B. Huang, *Chemistry Central Journal* **2013**, *7*, 1–12.
- [9] J. Schörghuber, L. Geist, G. Platzer, R. Konrat, R. J. Lichtenecker, *ChemBioChem* **2017**, *18*, 1487–1491.
- [10] J. L. Wilson, *Biochemistry*; (Stryer, Lubert), **1988**.
- [11] rcsb protein data bank, <https://www.rcsb.org>, Accessed: 2022-07-04.
- [12] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, D. C. Phillips, *Nature* **1958**, *181*, 662–666.
- [13] D. Eisenberg, C. P. Hill, *Trends in biochemical sciences* **1989**, *14*, 260–264.
- [14] M.-A. Delsuc, M. Vitorino, B. Kieffer in *Structural Biology in Drug Discovery*, John Wiley Sons, Ltd, **2020**, Chapter 13, pp. 295–323.
- [15] D. P. Frueh, A. C. Goodrich, S. H. Mishra, S. R. Nichols, *Current Opinion in Structural Biology* **2013**, *23*, Protein-carbohydrate interactions / Biophysical methods, 734–739.
- [16] L. Banci, I. Bertini, C. Luchinat, M. Mori, *Progress in Nuclear Magnetic Resonance Spectroscopy* **2010**, *56*, 247–266.
- [17] In Annual Reports on NMR Spectroscopy, Academic Press, **2002**, pp. 31–69.
- [18] F. Bloch, *Physical review* **1946**, *70*, 460.
- [19] E. M. Purcell, H. C. Torrey, R. V. Pound, *Physical review* **1946**, *69*, 37.

- [20] M. Pellecchia, D. S. Sem, K. Wüthrich, *Nature Reviews Drug Discovery* **2002**, *1*, 211–219.
- [21] G. Wider, S. Macura, A. Kumar, R. Ernst, K. Wüthrich, *Journal of Magnetic Resonance (1969)* **1984**, *56*, 207–234.
- [22] N. K. Goto, L. E. Kay, *Current Opinion in Structural Biology* **2000**, *10*, 585–592.
- [23] N. L. Richard R. Ernst, Nuclear Magnetic Resonance Fourier Transform Spectroscopy, <https://www.nobelprize.org/prizes/chemistry/1991/ernst/lecture/>, Accessed: 2022-07-27.
- [24] G. C. K. Roberts in *Encyclopedia of Biophysics*, (Ed.: G. C. K. Roberts), Springer Berlin Heidelberg, Berlin, Heidelberg, **2013**, pp. 2027–2033.
- [25] B. T. Farmer, R. A. Venters in *Modern Techniques in Protein NMR*, (Eds.: N. R. Krishna, L. J. Berliner), Springer US, Boston, MA, **2002**, pp. 75–120.
- [26] J.-H. Ardenkjaer-Larsen, G. S. Boebinger, A. Comment, S. Duckett, A. S. Edison, F. Engelke, C. Griesinger, R. G. Griffin, C. Hilty, H. Maeda, et al., *Angewandte Chemie International Edition* **2015**, *54*, 9162–9185.
- [27] C. Fernández, G. Wider, *Current opinion in structural biology* **2003**, *13*, 570–580.
- [28] K. H. Gardner, M. K. Rosen, L. E. Kay, *Biochemistry* **1997**, *36*, 1389–1401.
- [29] K. Wüthrich, *Europhysics News* **1986**, *17*, 11–13.
- [30] N. E. Jacobsen, *NMR spectroscopy explained: simplified theory, applications and examples for organic chemistry and structural biology*, John Wiley & Sons, **2007**.
- [31] Protein NMR - A Practical Guide, <https://protein-nmr.org.uk/solution-nmr/spectrum-descriptions/hncoca/>, Accessed: 2022-12-04.
- [32] L. E. Kay, M. Ikura, R. Tschudin, A. Bax, *Journal of Magnetic Resonance (1969)* **1990**, *89*, 496–514.
- [33] T. Diercks, M. Coles, H. Kessler, *Current opinion in chemical biology* **2001**, *5*, 285–291.
- [34] C. Aguirre, O. Cala, I. Krimm, *Current protocols in protein science* **2015**, *81*, 17–18.
- [35] A. Kumar in *Proceedings of the Indian Academy of Sciences-Chemical Sciences*, *Vol. 95*, Springer, **1985**, pp. 1–8.
- [36] S. Maity, R. K. Gundampati, T. K. Suresh Kumar, *Natural Product Communications* **2019**, *14*, 1934578X19849296.

REFERENCES

- [37] M. P. Williamson, *Progress in nuclear magnetic resonance spectroscopy* **2013**, *73*, 1–16.
- [38] G. Platzer, M. Mayer, A. Beier, S. Brüschweiler, J. E. Fuchs, H. Engelhardt, L. Geist, G. Bader, J. Schörghuber, R. Lichtenecker, et al., *Angewandte Chemie* **2020**, *132*, 14971–14978.
- [39] M. Mayer, B. Meyer, *Angewandte Chemie International Edition* **1999**, *38*, 1784–1788.
- [40] C. D. L. Lima, PhD thesis, Universidade NOVA de Lisboa (Portugal), **2021**.
- [41] J. J. Ziarek, D. Baptista, G. Wagner, *Journal of Molecular Medicine* **2018**, *96*, 1–8.
- [42] K. H. Gardner, L. E. Kay, *Annual Review of Biophysics and Biomolecular Structure* **1998**, *27*, PMID: 9646872, 357–406.
- [43] R. Verardi, N. J. Traaseth, L. R. Masterson, V. V. Vostrikov, G. Veglia in *Isotope labeling in Biomolecular NMR*, (Ed.: H. S. Atreya), Springer Netherlands, Dordrecht, **2012**, pp. 35–62.
- [44] C. Klammt, D. Schwarz, F. Löhr, B. Schneider, V. Dötsch, F. Bernhard, *The FEBS Journal* **2006**, *273*, 4141–4153.
- [45] C. G. Hoogstraten, J. E. Johnson Jr., *Concepts in Magnetic Resonance Part A* **2008**, *32A*, 34–55.
- [46] K. H. Gardner, L. E. Kay, *Annual review of biophysics and biomolecular structure* **1998**, *27*, 357–406.
- [47] D. M. Kushlan, D. M. LeMaster, *Journal of Biomolecular NMR* **1993**, *3*, 701–708.
- [48] S. Grzesiek, J. Anglister, H. Ren, A. Bax, *Journal of the American Chemical Society* **1993**, *115*, 4369–4370.
- [49] D. M. LeMaster, *Progress in Nuclear Magnetic Resonance Spectroscopy* **1994**, *26*, 371–419.
- [50] S. Mondal, D. Shet, C. Prasanna, H. S. Atreya, **2013**.
- [51] W. Morgan, A. Kragt, J. Feeney, *Journal of Biomolecular NMR* **2000**, *17*, 337–347.
- [52] H. Takahashi, I. Shimada, *Journal of Biomolecular NMR* **2010**, *46*, 3–10.
- [53] M. Kainosho, P. Güntert, *Quarterly reviews of biophysics* **2009**, *42*, 247–300.
- [54] R. J. Lichtenecker, N. Coudevylle, R. Konrat, W. Schmid, *ChemBioChem* **2013**, *14*, 818–821.

- [55] R. J. Lichteneker, K. Weinhäupl, L. Reuther, J. Schörghuber, W. Schmid, R. Konrat, *Journal of Biomolecular NMR* **2013**, *57*, 205–209.
- [56] N. K. Goto, K. H. Gardner, G. A. Mueller, R. C. Willis, L. E. Kay, *Journal of Biomolecular NMR* **1999**, *13*, 369–374.
- [57] MetaCyc Database, <https://metacyc.org>, Accessed: 2022-08-26.
- [58] J. Schörghuber, L. Geist, G. Platzer, M. Feichtinger, M. Bisaccia, L. Scheibelberger, F. Weber, R. Konrat, R. J. Lichteneker, *Journal of Biomolecular NMR* **2018**, *71*, 129–140.
- [59] K. Uchida, *Amino acids* **2003**, *25*, 249–257.
- [60] A. Martínez, *Amino Acids* **1995**, *9*, 285–292.
- [61] A. Doğan, A. D. Özel, E. Kılıç, *Amino Acids* **2009**, *36*, 373–379.
- [62] U. D. Priyakumar, M. Punnagai, G. K. Mohan, G. N. Sastry, *Tetrahedron* **2004**, *60*, 3037–3043.
- [63] D. Vijay, G. N. Sastry, *Physical Chemistry Chemical Physics* **2008**, *10*, 582–590.
- [64] S. Grimme, *Angewandte Chemie International Edition* **2008**, *47*, 3430–3434.
- [65] M. Remko, D. Fitz, B. M. Rode, *Amino Acids* **2010**, *39*, 1309–1319.
- [66] M. Sargolzaei, M. Afshar, M. S. Sadeghi, H. Hamidian, *Journal of Structural Chemistry* **2014**, *55*, 1627–1634.
- [67] S. Talab, K. K. Taha, J. Lugtenburg, *Molecules* **2014**, *19*, 1023–1033.
- [68] M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. Mei Ono, P. Güntert, *Nature* **2006**, *440*, 52–57.
- [69] S. Burley, G. A. Petsko, *Science* **1985**, *229*, 23–28.
- [70] S. Ramazi, J. Zahiri, *Database* **2021**, *2021*.
- [71] P. Levene, C. Alsberg, *Journal of Biological Chemistry* **1906**, *2*, 127–133.
- [72] T. Pawson, J. D. Scott, *Trends in biochemical sciences* **2005**, *30*, 286–290.
- [73] N. Makukhin, A. Ciulli, *RSC Medicinal Chemistry* **2021**, *12*, 8–23.
- [74] T. Bilbrough, E. Piemontese, O. Seitz, *Chemical Society Reviews* **2022**.
- [75] T. Hunter, *Cold Spring Harbor perspectives in biology* **2014**, *6*, a020644.
- [76] L. Tautz, D. A. Critton, S. Grotegut, *Phosphatase Modulators* **2013**, 179–221.
- [77] R. Seger, E. G. Krebs, *The FASEB Journal* **1995**, *9*, 726–735.
- [78] B. A. Liu, K. Jablonowski, M. Raina, M. Arcé, T. Pawson, P. D. Nash, *Molecular cell* **2006**, *22*, 851–868.

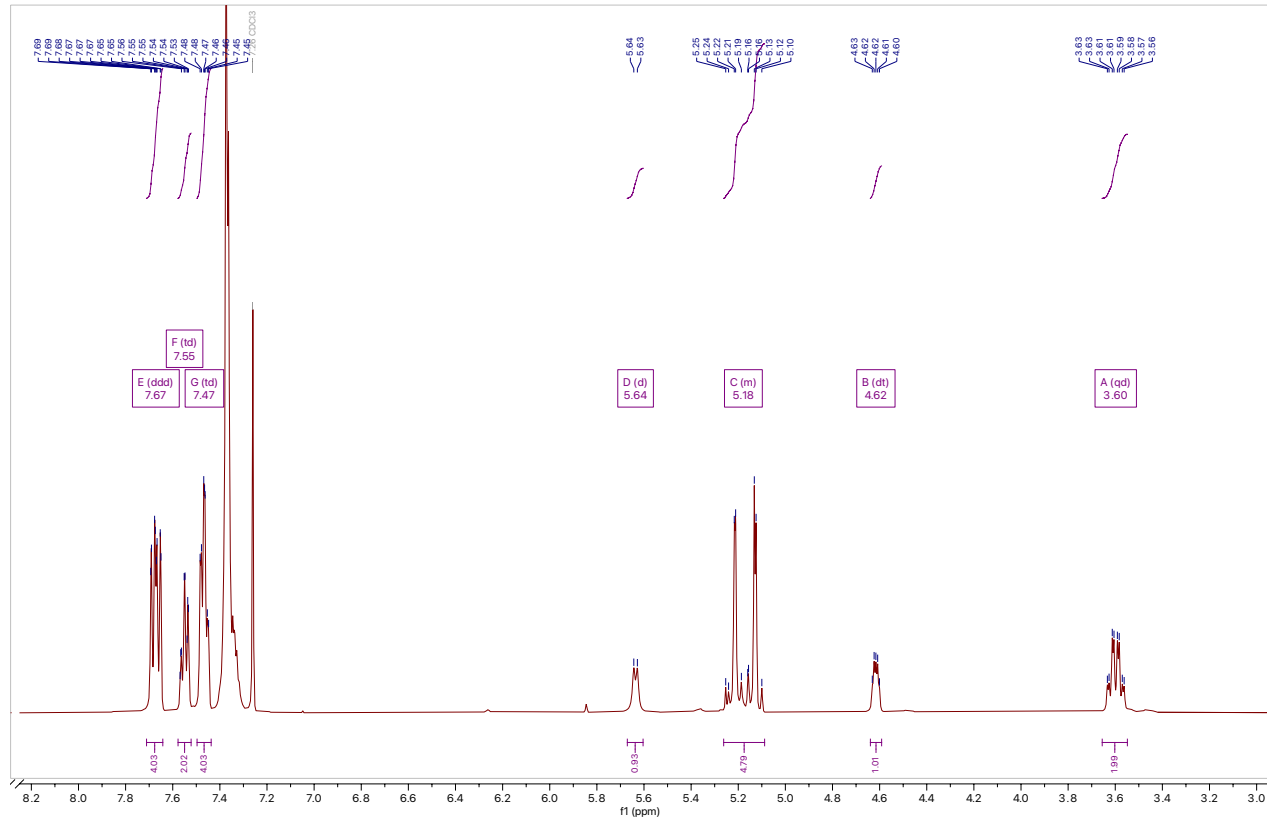
REFERENCES

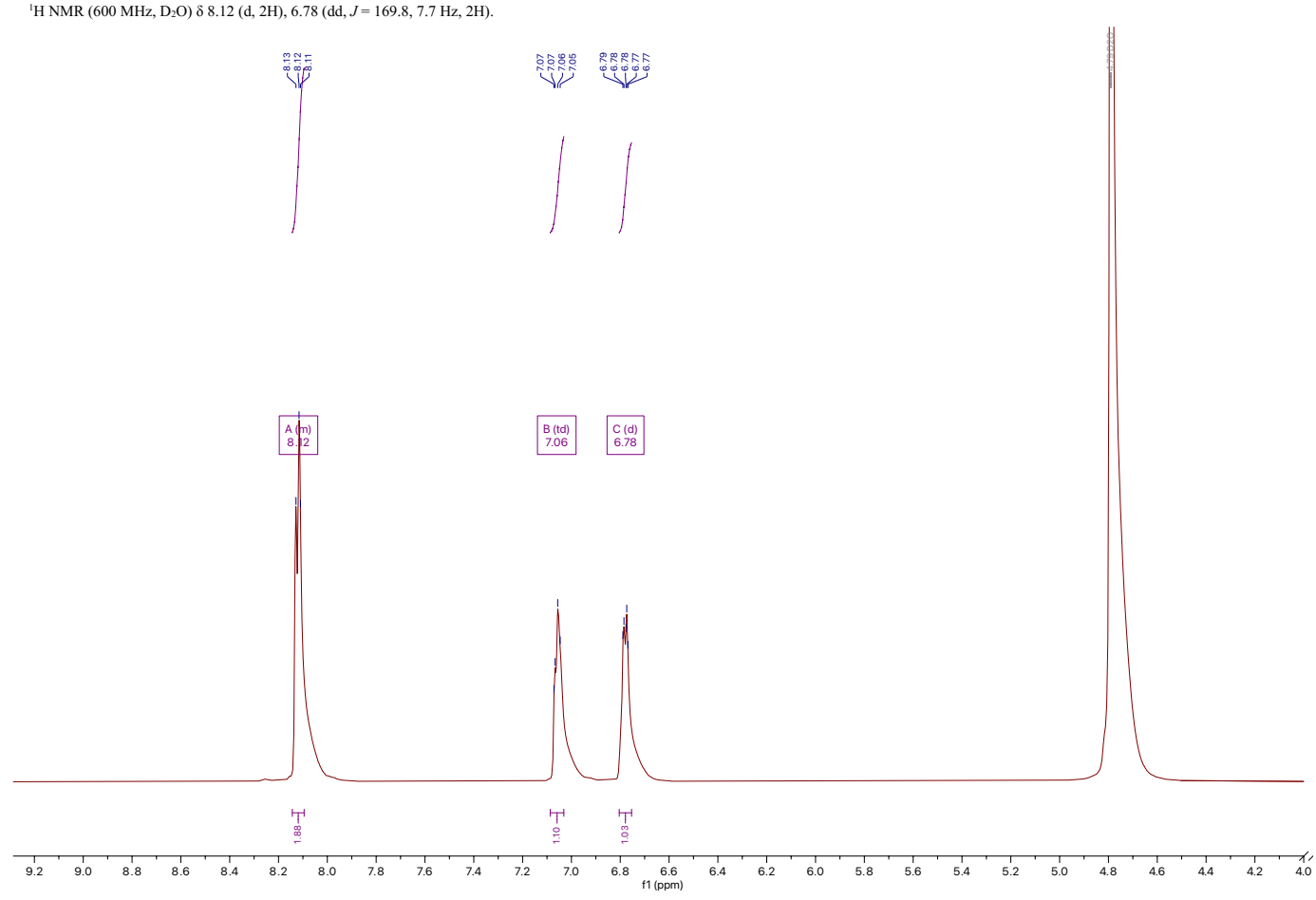
- [79] P. Filippakopoulos, S. Müller, S. Knapp, *Current opinion in structural biology* **2009**, *19*, 643–649.
- [80] J. Kuriyan, D. Cowburn, *Annual review of biophysics and biomolecular structure* **1997**, *26*, 259–288.
- [81] R. Behrendt, P. White, J. Offer, *Journal of Peptide Science* **2016**, *22*, 4–27.
- [82] B. Handa, C. Hobbs, *Journal of Peptide Science: An Official Publication of the European Peptide Society* **1998**, *4*, 138–141.
- [83] M. Ueki, M. Goto, J. Okumura in *Peptide Science—Present and Future*, Springer, **1999**, pp. 613–614.
- [84] J. Rahil, P. Haake, *Journal of the American Chemical Society* **1981**, *103*, 1723–1734.
- [85] R. J. Lichtenecker, K. Weinhäupl, W. Schmid, R. Konrat, *Journal of Biomolecular NMR* **2013**, *57*, 327–331.
- [86] M. Takeda, A. M. Ono, T. Terauchi, M. Kainosho, *Journal of Biomolecular NMR* **2010**, *46*, 45–49.
- [87] M. M. Heravi, E. Hashemi, N. Nazari, *Molecular diversity* **2014**, *18*, 441–472.
- [88] K. Nicolaou, P. G. Bulger, D. Sarlah, *Angewandte Chemie International Edition* **2005**, *44*, 4442–4489.
- [89] E. Negishi, A. O. King, N. Okukado, *The Journal of organic chemistry* **1977**, *42*, 1821–1823.
- [90] J. L. Serrano, T. R. Girase in *Palladacycles*, (Eds.: A. R. Kapdi, D. Maiti), Elsevier, **2019**, pp. 175–224.
- [91] J. Pelton, D. A. Torchia, N. Meadow, S. Roseman, *Protein Science* **1993**, *2*, 543–558.
- [92] M.-Y. Xu, W.-T. Jiang, Y. Li, Q.-H. Xu, Q.-L. Zhou, S. Yang, B. Xiao, *Journal of the American Chemical Society* **2019**, *141*, 7582–7588.
- [93] S. Mitchell, R. Waring, *Kirk-Othmer Encyclopedia of Chemical Technology* **2000**.
- [94] S. Shimodaira, M. Iwaoka, *Organic Chemistry* **2017**, 260–271.
- [95] R. Pascal, R. Sola, *Tetrahedron letters* **1998**, *39*, 5031–5034.
- [96] J. Trotter, W. Darby, *Organic Syntheses, Collect Vol III*, **1955**.
- [97] R. Kimura, T. Nagano, H. Kinoshita, *Bulletin of the Chemical Society of Japan* **2002**, *75*, 2517–2525.

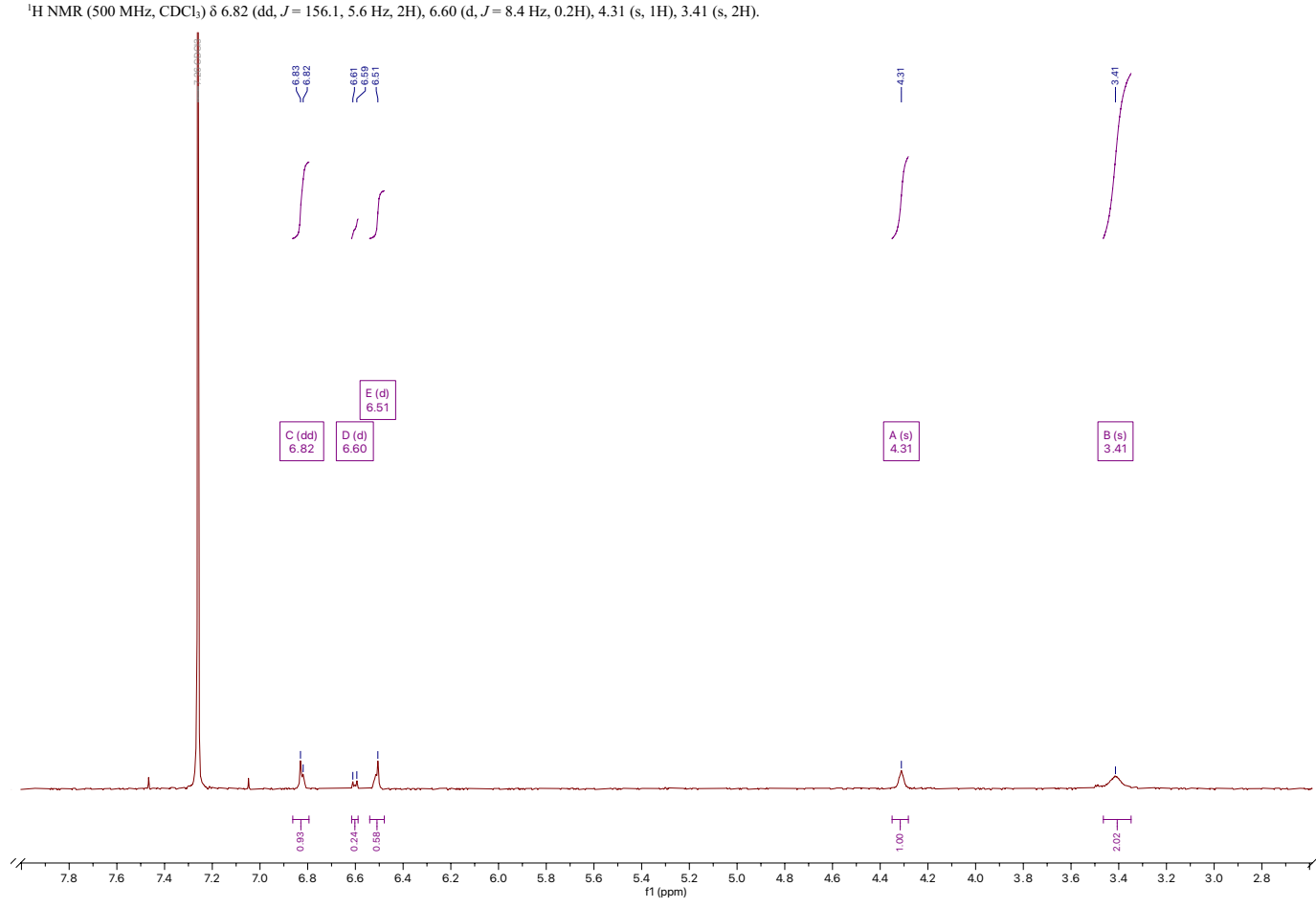
NMR Spectra

(R)-2-Benzyloxycarbonylamino-3-iodo-propionic acid benzyl ester 2

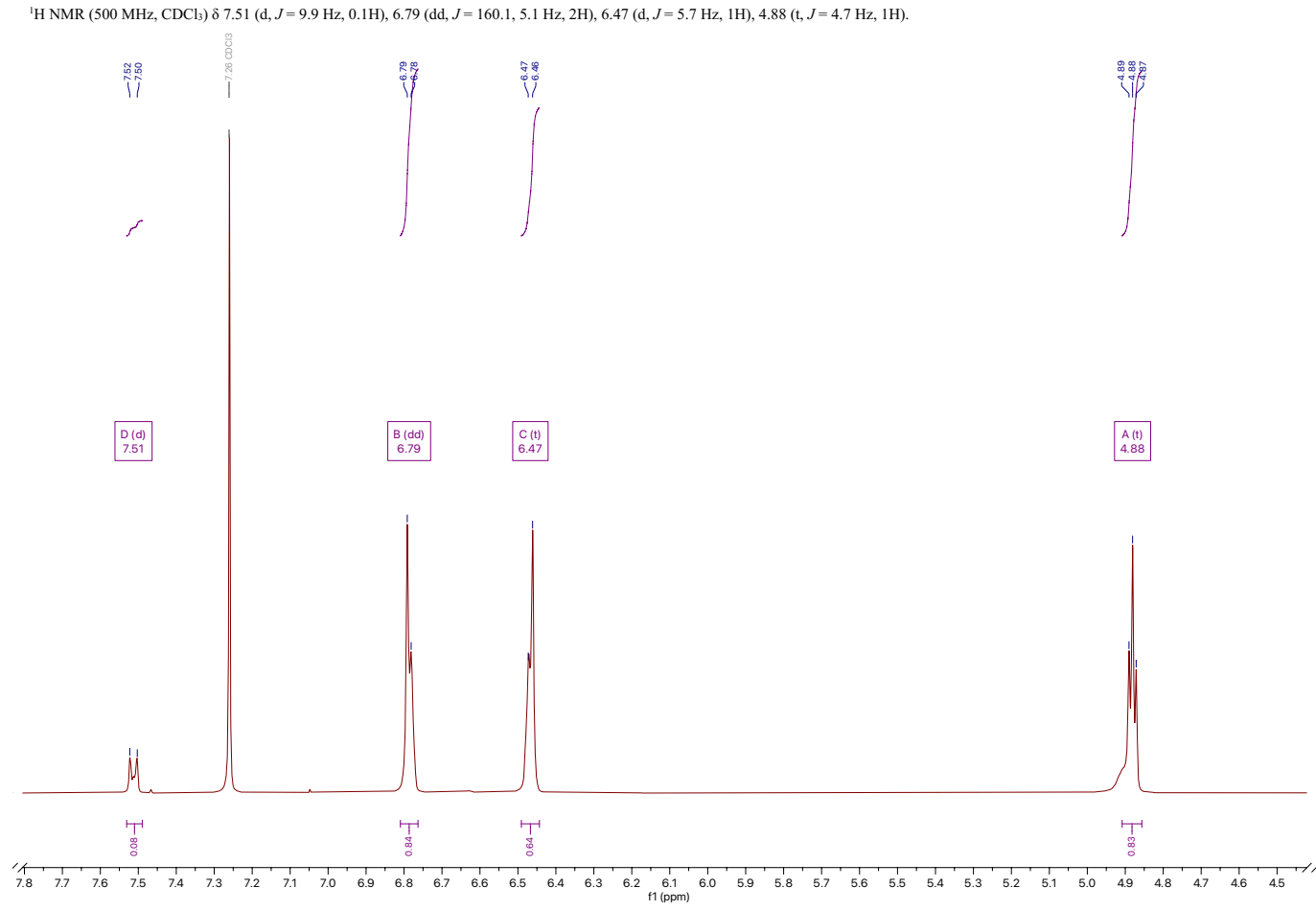
¹H NMR (500 MHz, CDCl₃) δ 7.67 (ddd, *J* = 12.0, 8.2, 1.5 Hz, 4H), 7.55 (td, *J* = 7.3, 1.5 Hz, 2H), 7.47 (td, *J* = 7.5, 2.8 Hz, 4H), 5.64 (d, *J* = 7.6 Hz, 1H), 5.26 – 5.09 (m, 4H), 4.62 (dt, *J* = 7.6, 3.8 Hz, 1H), 3.60 (qd, *J* = 10.4, 3.8 Hz, 2H).



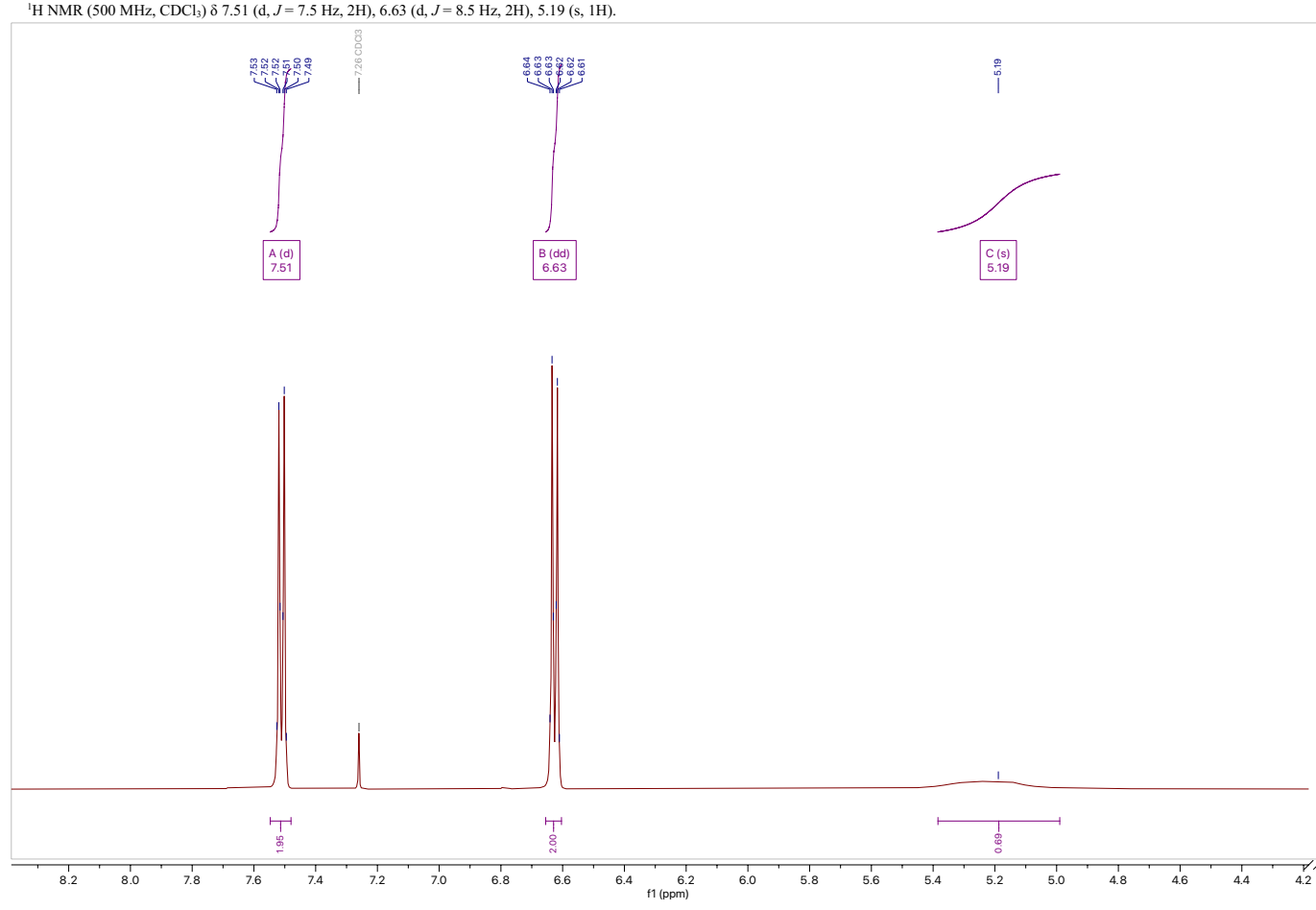
[2,6-¹³C₂]4-Nitrophenol 4

[2,6-¹³C₂]3,5-Dideuterio-4-aminophenol 6

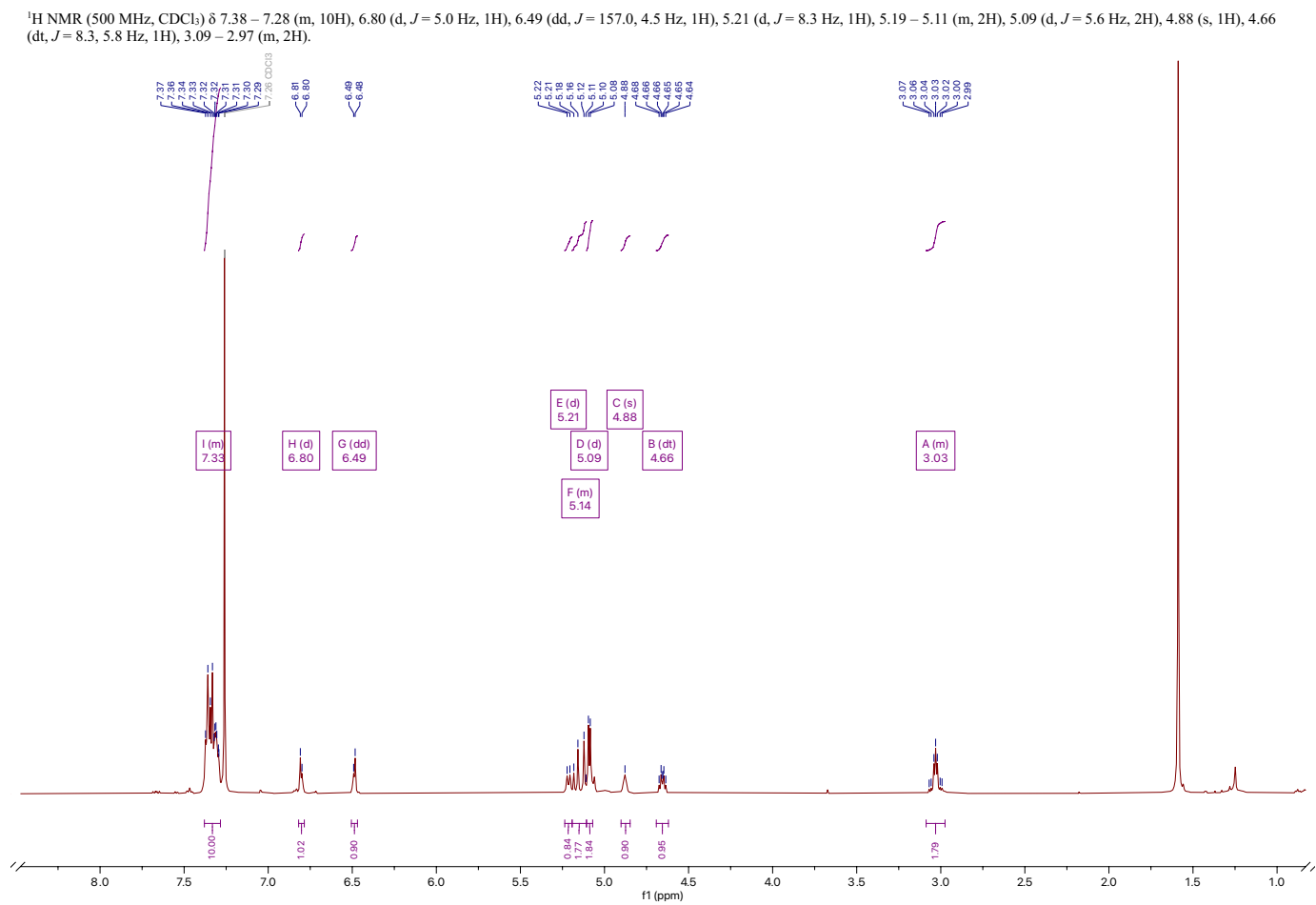
[2,6-¹³C₂]3,5-Dideuterio-4-iodophenol 7



4-Iodophenol 7b

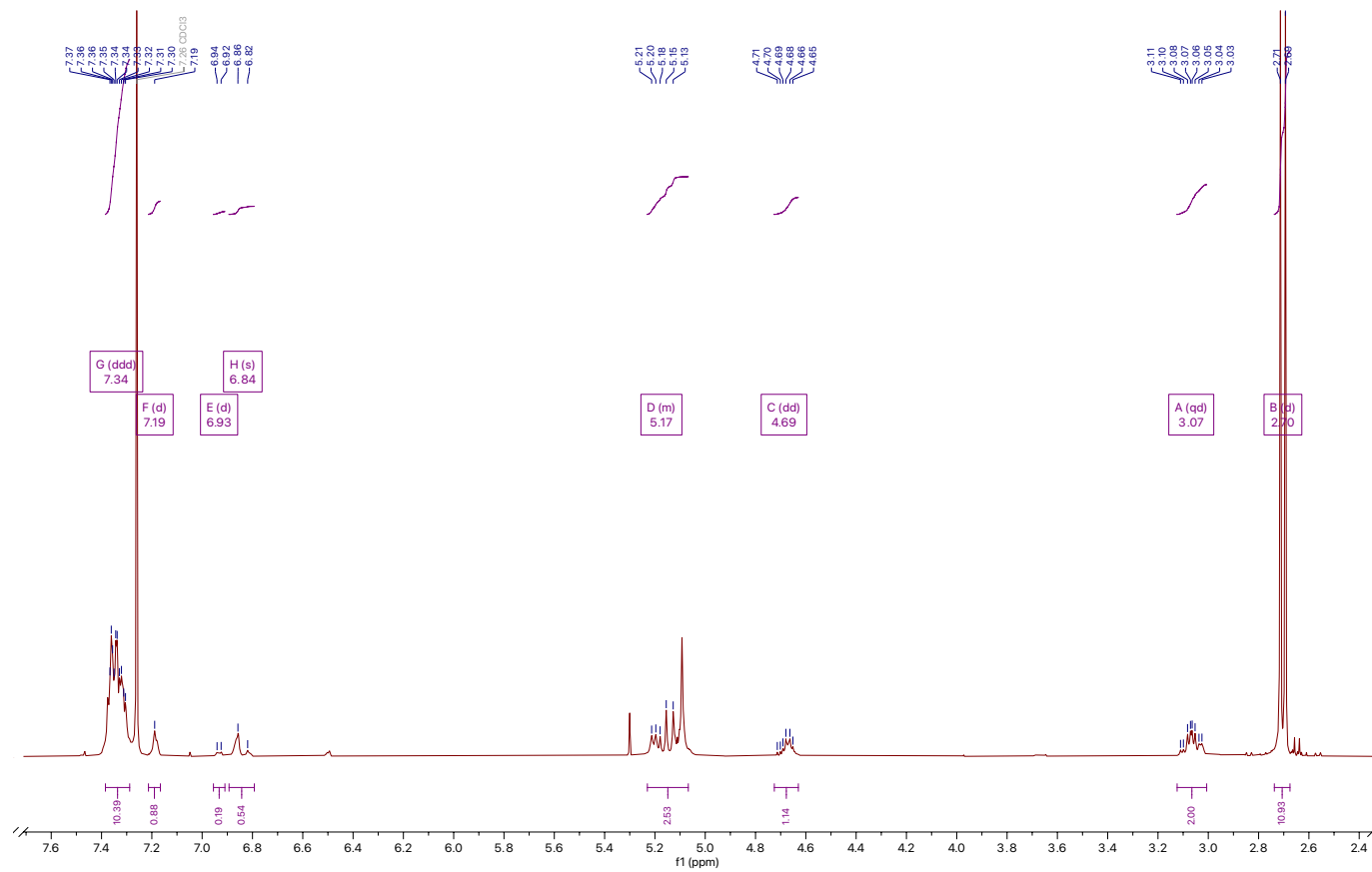


N-Carbobenzoxy[3,5-¹³C₂-2,6-D₂]-L-tyrosine benzyl ester 8



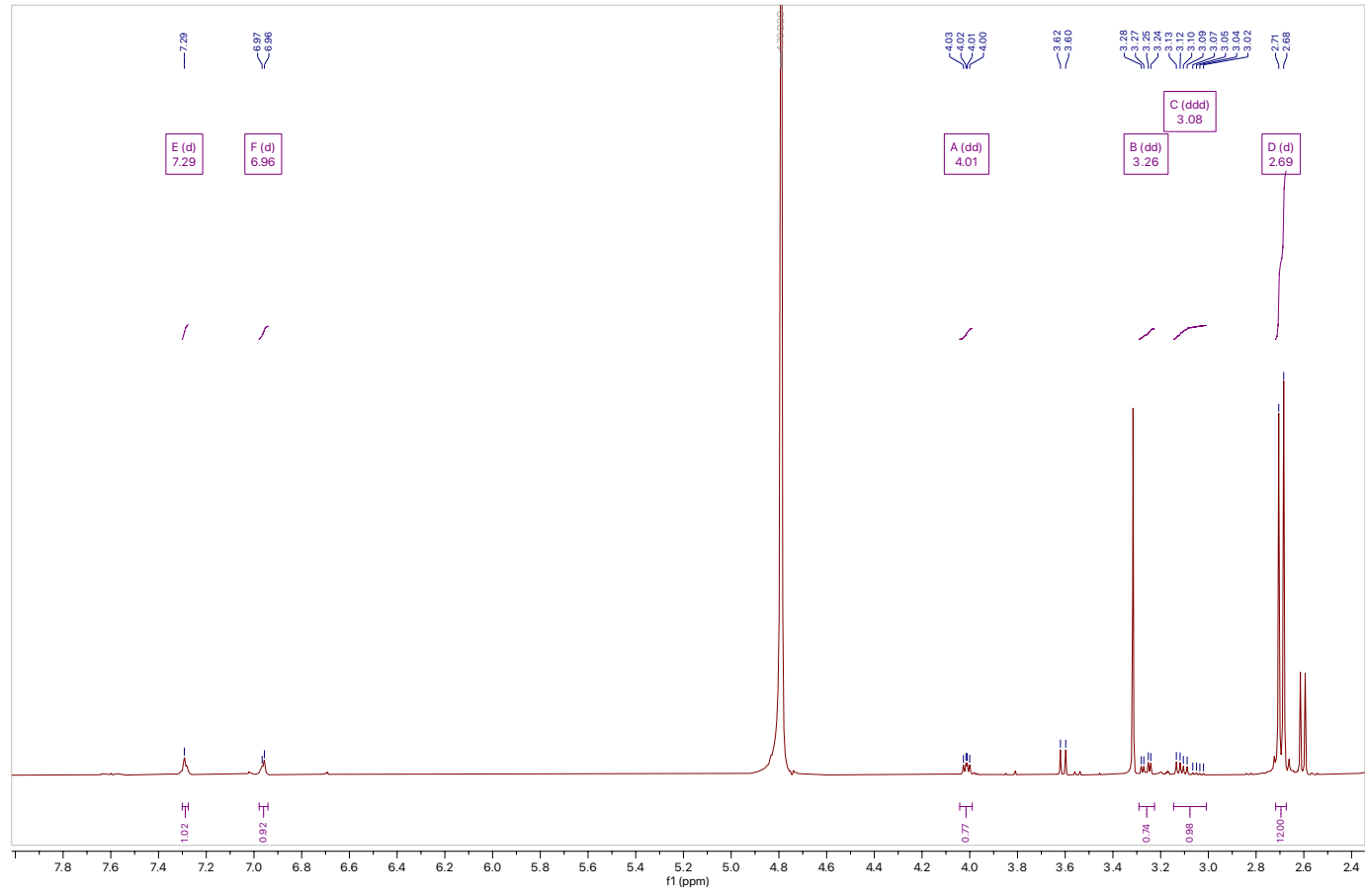
N-Carbobenzoxy[3,5-¹³C₂-2,6-D₂]-L-tyrosine[P(O)(NMe₂)₂] benzyl ester **9**

¹H NMR (500 MHz, CDCl₃) δ 7.34 (ddd, *J* = 13.1, 6.8, 3.6 Hz, 10H), 7.19 (dd, *J* = 159.3, 6.2 Hz, 2H), 6.93 (d, *J* = 8.1 Hz, 0.2H), 5.23 – 5.11 (m, 2H), 4.69 (dd, *J* = 15.5, 9.6 Hz, 1H), 3.07 (qd, *J* = 13.8, 5.8 Hz, 2H), 2.70 (d, *J* = 10.0 Hz, 12H).



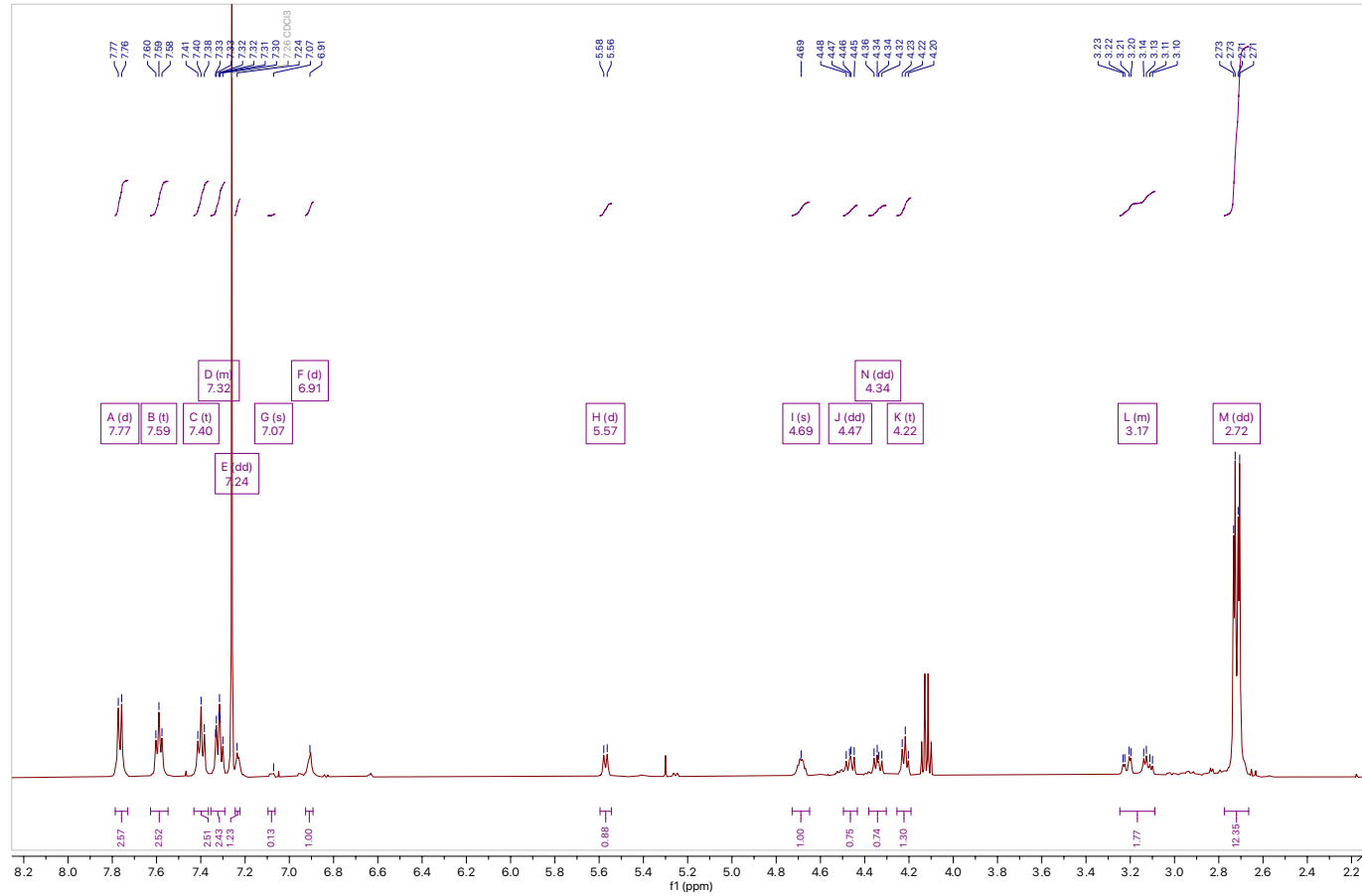
[3,5-¹³C₂-2,6-D₂]-L-Tyrosine[P(O)(NMe₂)₂] 10

¹H NMR (500 MHz, D₂O) δ 7.29 (dd, *J* = 160.3, 5.7 Hz, 2H), 4.01 (dd, *J* = 7.9, 5.3 Hz, 1H), 3.26 (dd, *J* = 14.7, 5.3 Hz, 1H), 3.08 (ddd, *J* = 34.3, 14.7, 7.8 Hz, 1H), 2.69 (d, *J* = 10.4 Hz, 12H).



N-Fmoc[3,5-¹³C₂-2,6-D₂]-L-tyrosine[P(O)(NMe₂)₂] 11

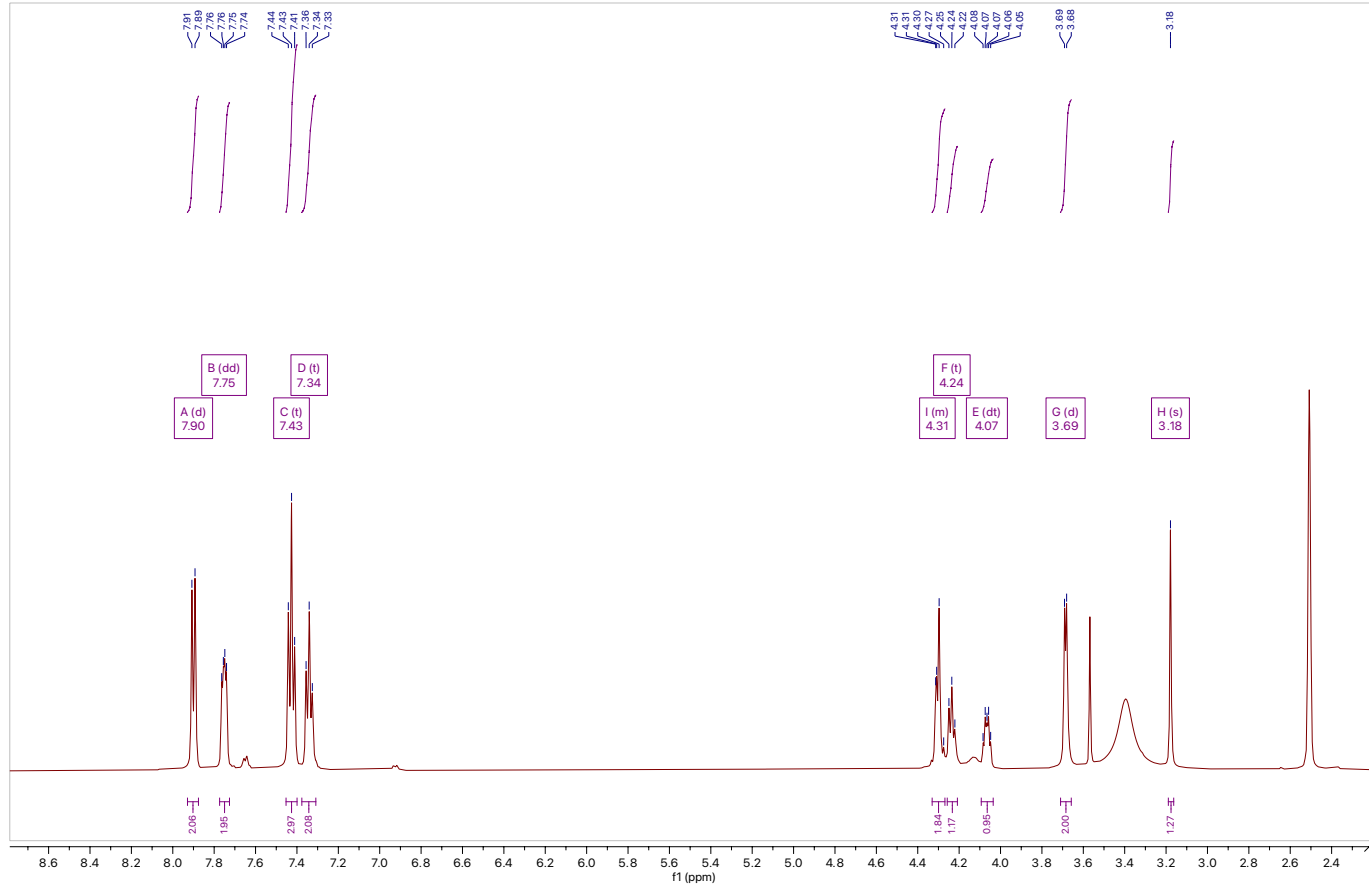
¹H NMR (500 MHz, CDCl₃) δ 7.77 (d, *J* = 7.6 Hz, 3H), 7.59 (t, *J* = 6.9 Hz, 3H), 7.40 (t, *J* = 7.4 Hz, 3H), 7.35 – 7.29 (m, 2H), 7.24 (dd, *J* = 165.3, 4.2 Hz, 1H), 7.07 (s, 0H), 5.57 (d, *J* = 7.7 Hz, 1H), 4.69 (s, 1H), 4.47 (dd, *J* = 10.6, 7.4 Hz, 1H), 4.34 (dd, *J* = 10.6, 7.0 Hz, 1H), 4.22 (t, *J* = 6.9 Hz, 1H), 3.25 – 3.09 (m, 2H), 2.72 (dd, *J* = 10.2, 3.5 Hz, 12H).



XX

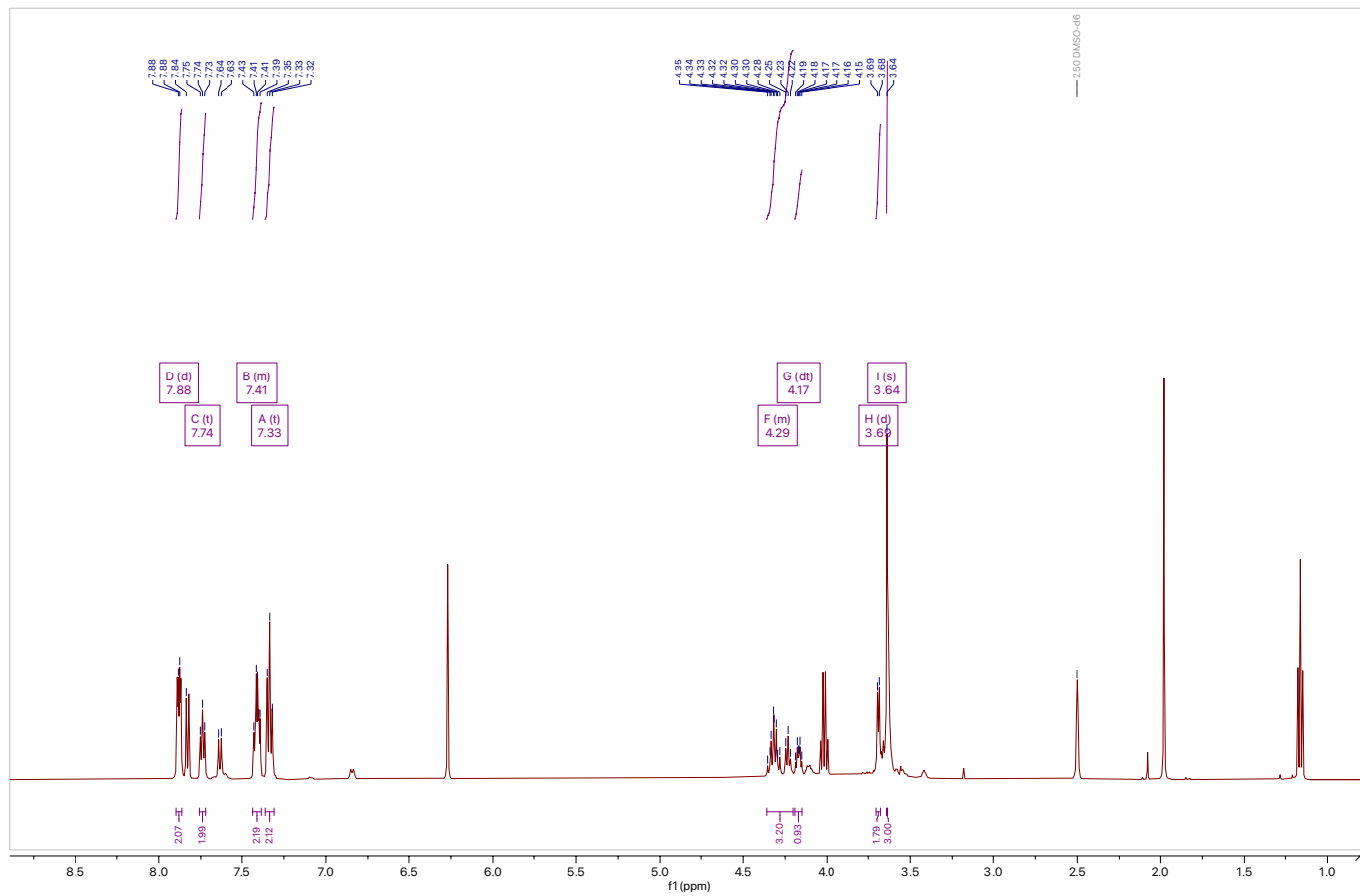
N-Fmoc-L-serine 13

^1H NMR (500 MHz, DMSO) δ 7.90 (d, $J = 7.5$ Hz, 2H), 7.75 (dd, $J = 7.6, 4.1$ Hz, 2H), 7.43 (t, $J = 7.7$ Hz, 2H), 7.34 (t, $J = 7.5$ Hz, 2H), 4.33 – 4.27 (m, 2H), 4.24 (t, $J = 7.2$ Hz, 1H), 4.07 (dt, $J = 8.8, 5.1$ Hz, 1H), 3.69 (d, $J = 5.1$ Hz, 2H), 3.18 (s, 1H).



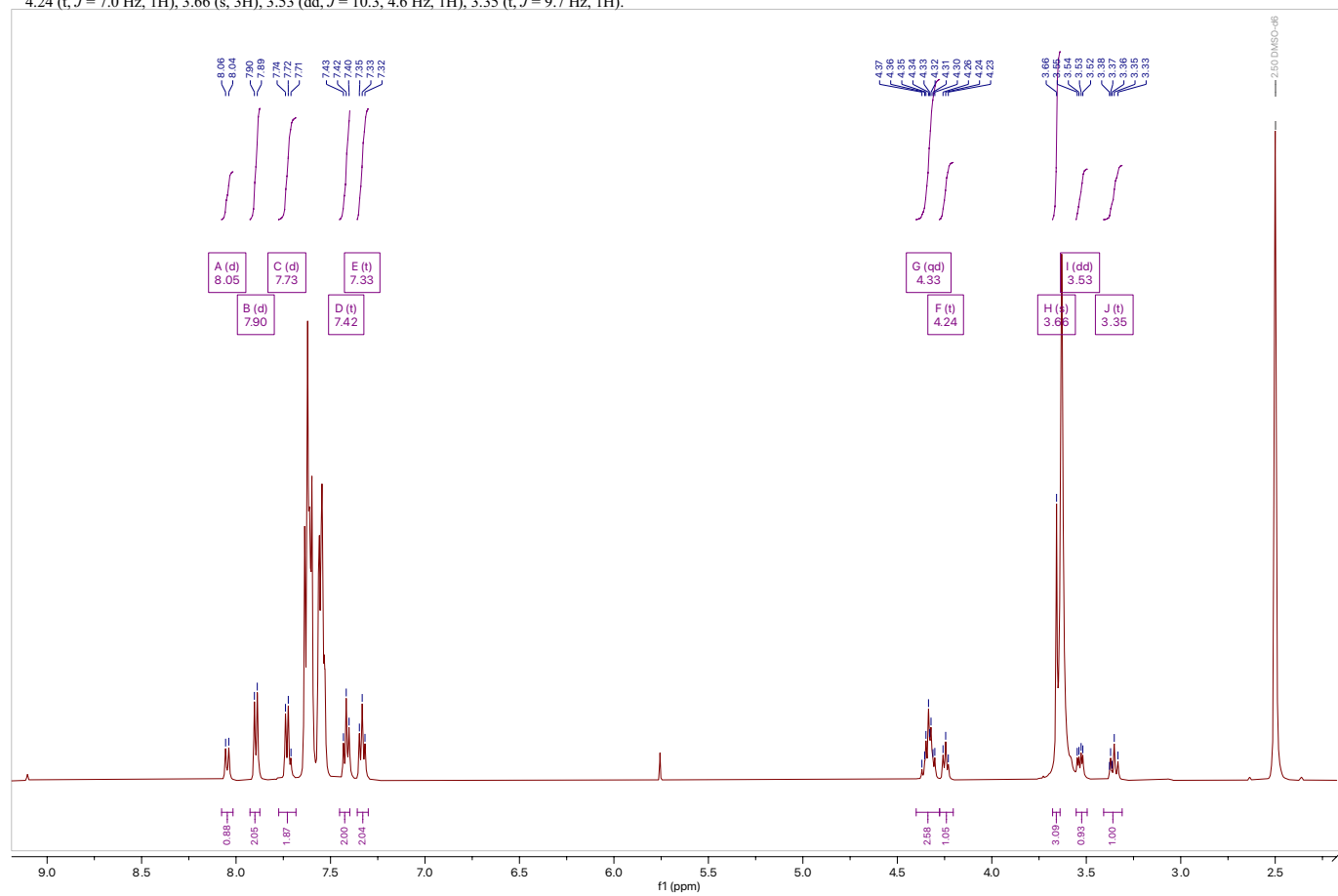
N-Fmoc-L-serine methyl ester 14

$^1\text{H NMR}$ (500 MHz, DMSO) δ 7.88 (d, $J = 4.1$ Hz, 2H), 7.74 (t, $J = 6.6$ Hz, 2H), 7.44 – 7.38 (m, 2H), 7.33 (t, $J = 7.6$ Hz, 2H), 4.36 – 4.20 (m, 3H), 4.17 (dt, $J = 8.0, 5.2$ Hz, 1H), 3.69 (d, $J = 5.3$ Hz, 2H), 3.64 (s, 3H).

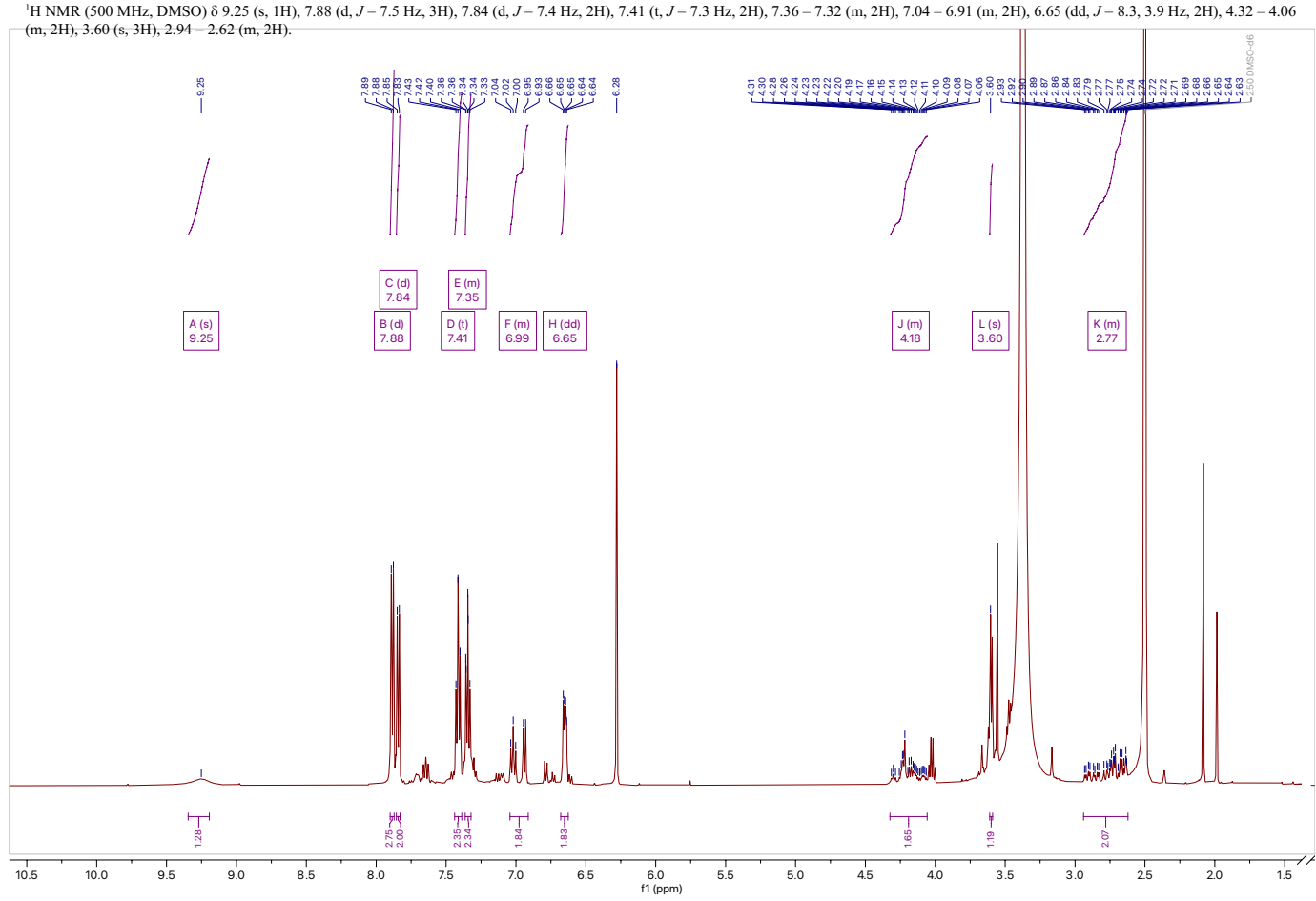


N-Fmoc-3-iodo-L-alanine methyl ester 15

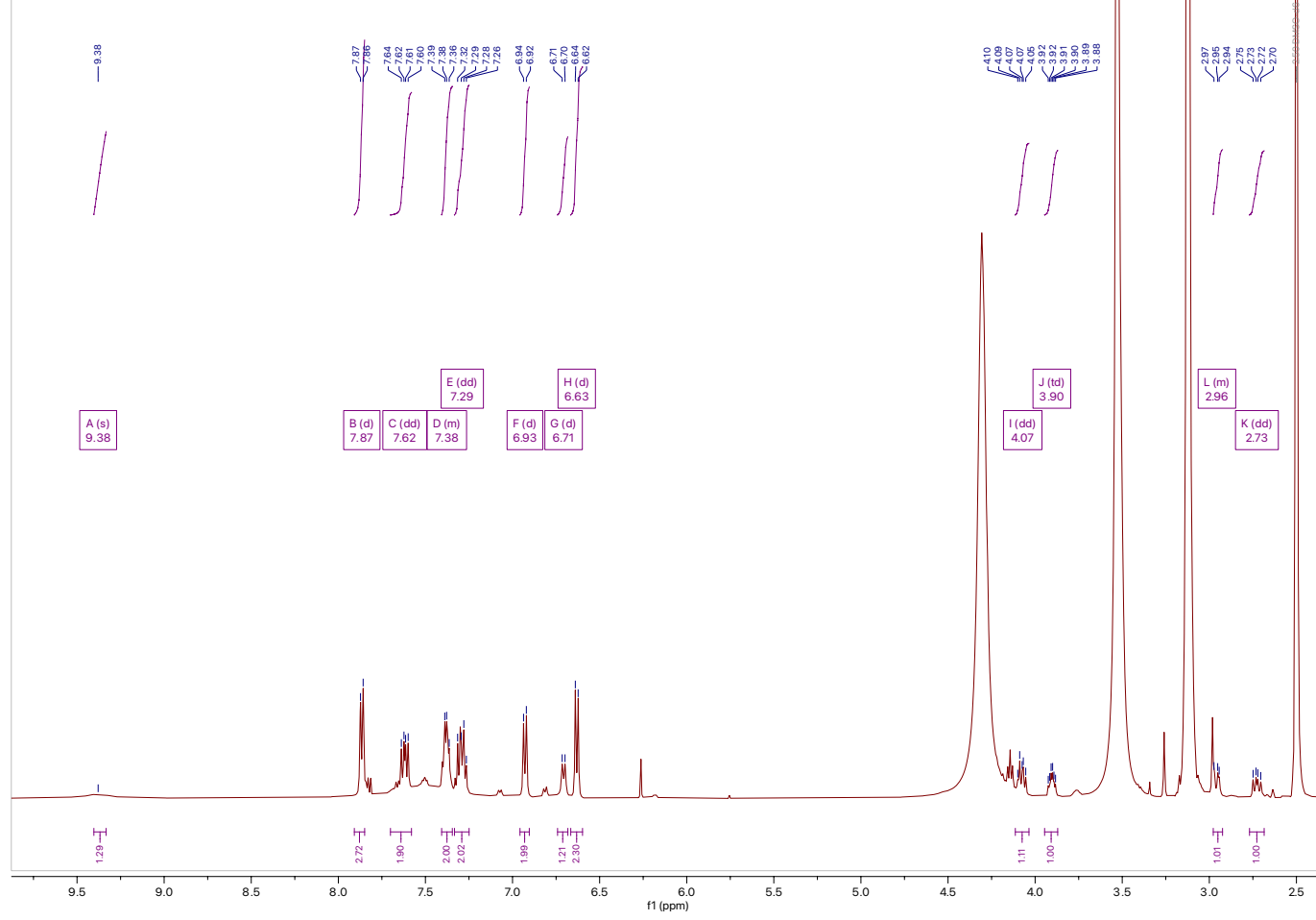
$^1\text{H NMR}$ (500 MHz, DMSO) δ 8.05 (d, $J = 8.3$ Hz, 1H), 7.90 (d, $J = 7.5$ Hz, 2H), 7.73 (d, $J = 7.5$ Hz, 2H), 7.42 (t, $J = 7.4$ Hz, 2H), 7.33 (t, $J = 7.4$ Hz, 2H), 4.33 (qd, $J = 10.4, 6.2$ Hz, 3H), 4.24 (t, $J = 7.0$ Hz, 1H), 3.66 (s, 3H), 3.53 (dd, $J = 10.3, 4.6$ Hz, 1H), 3.35 (t, $J = 9.7$ Hz, 1H).

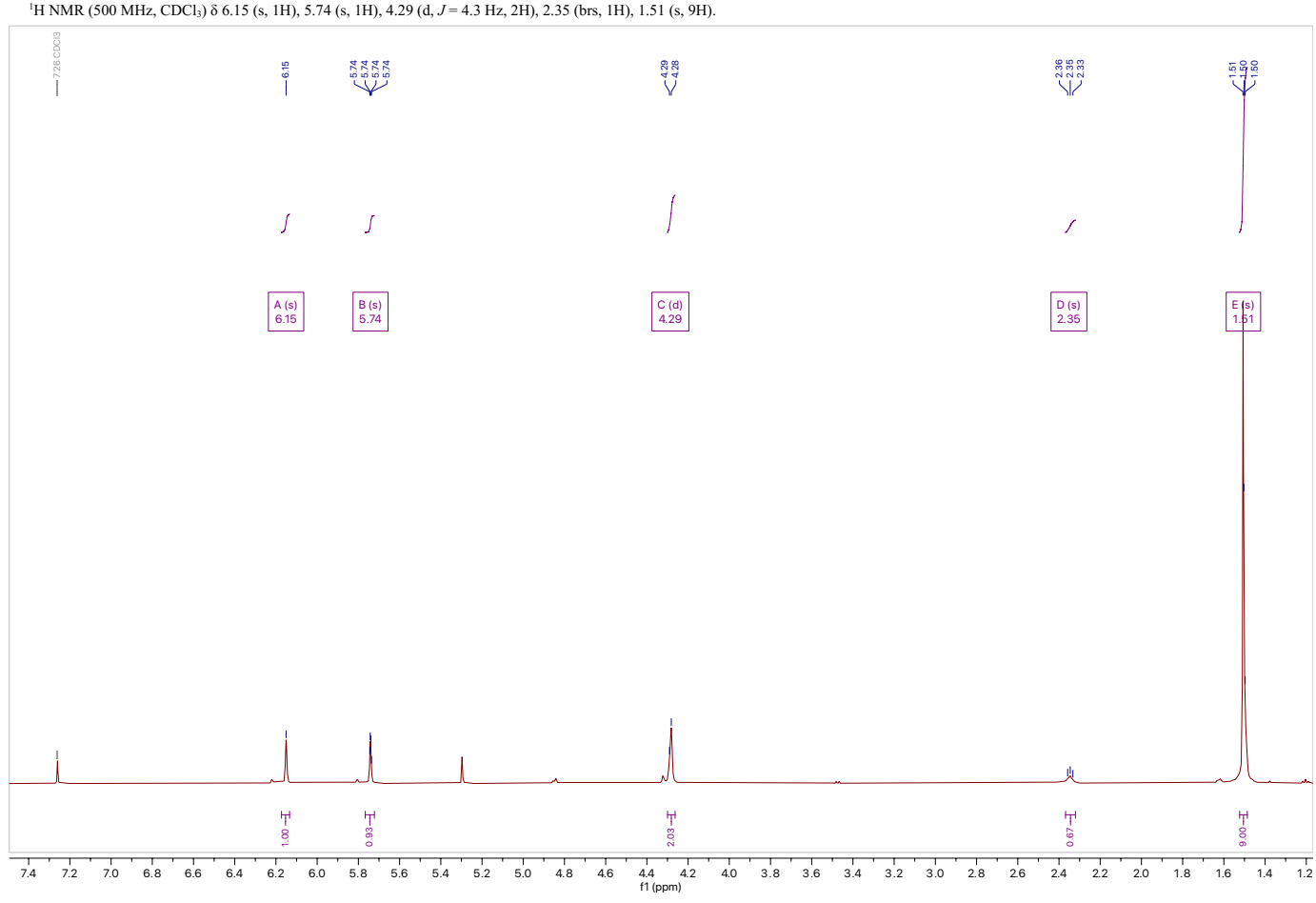


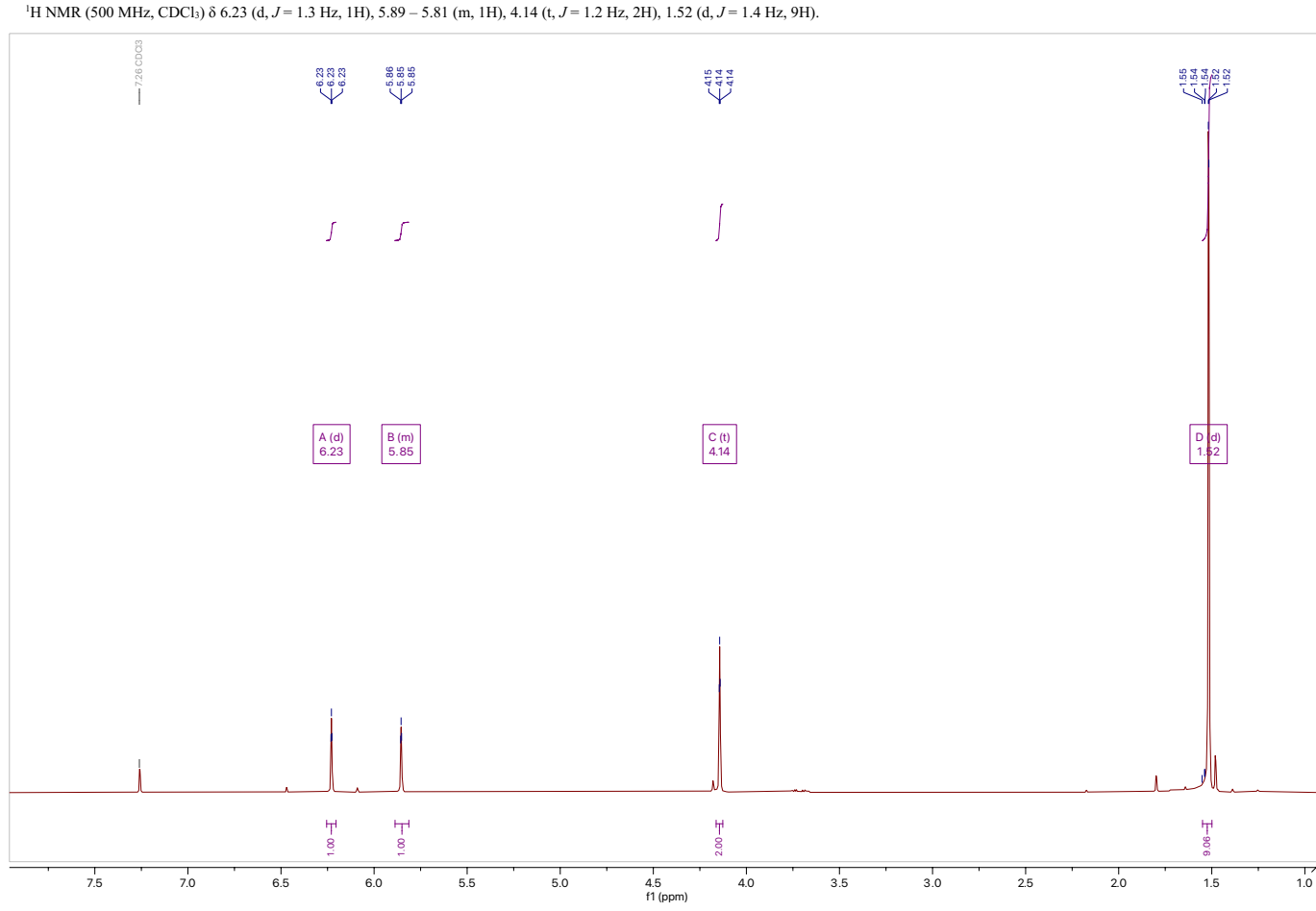
N-Fmoc-L-tyrosine methyl ester 16



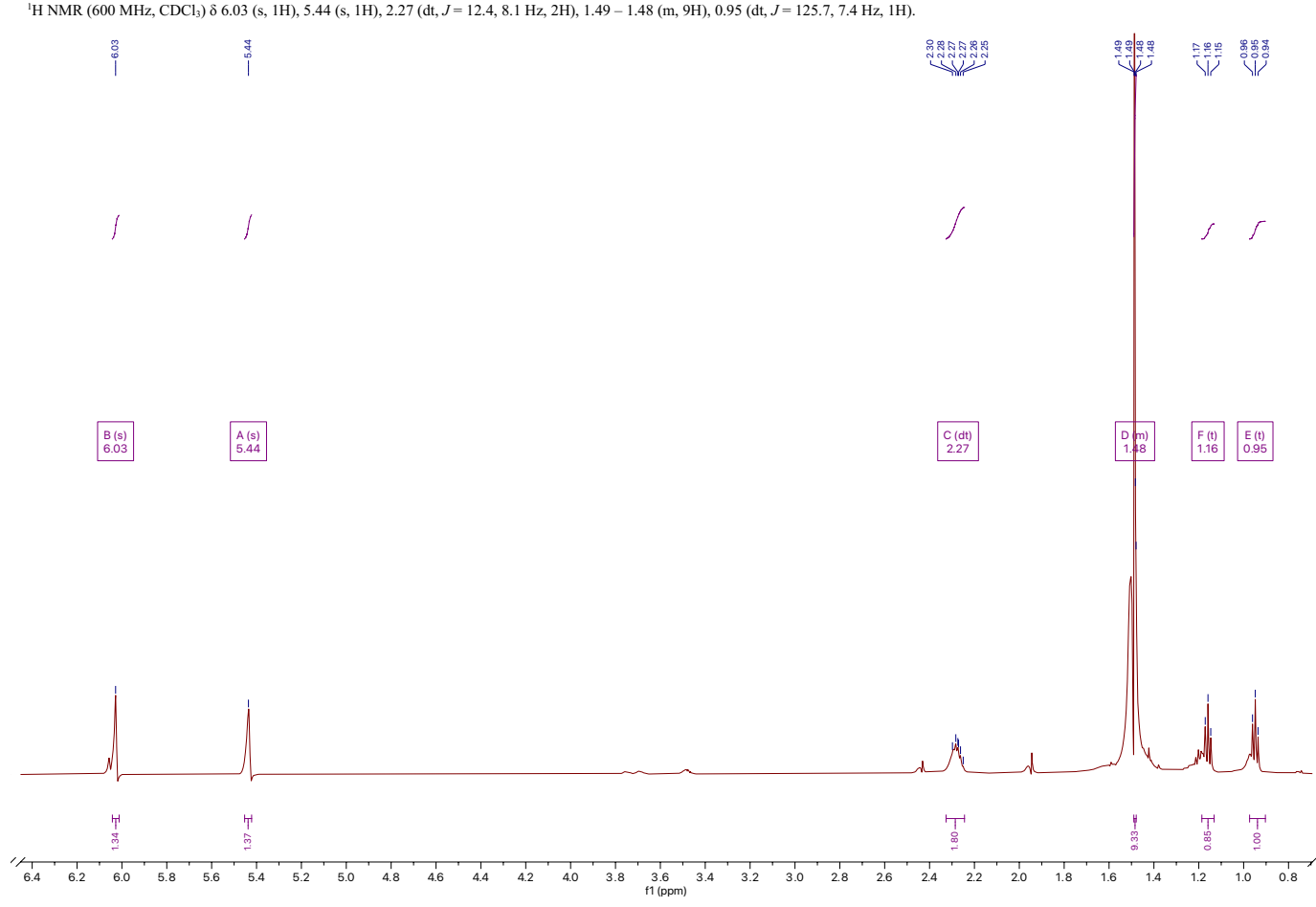
N-Fmoc-L-tyrosine 17



tert-Butyl 2-(hydroxymethyl)acrylate **20**

tert-Butyl 2-(bromomethyl)acrylate 21

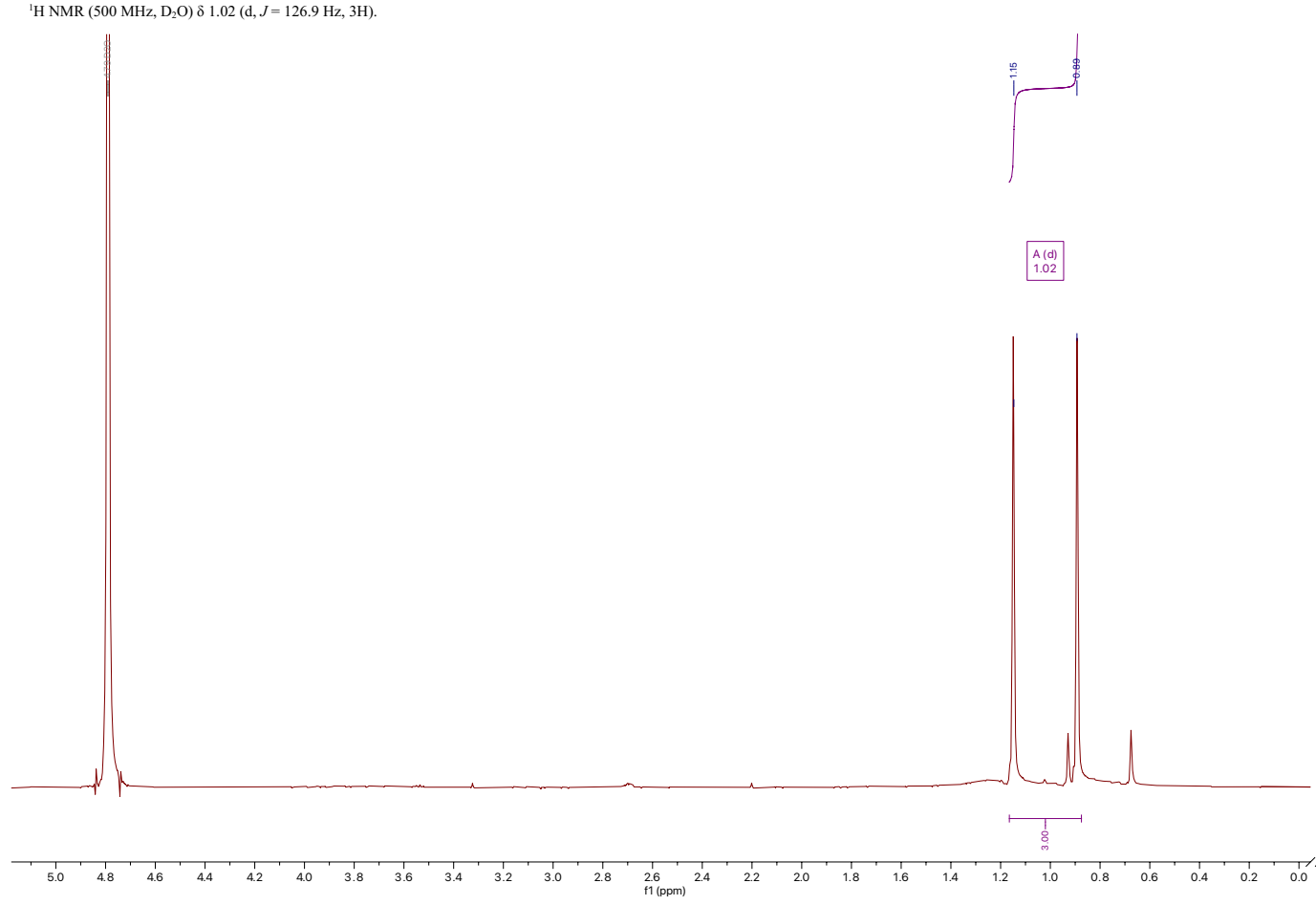
ii:XXX

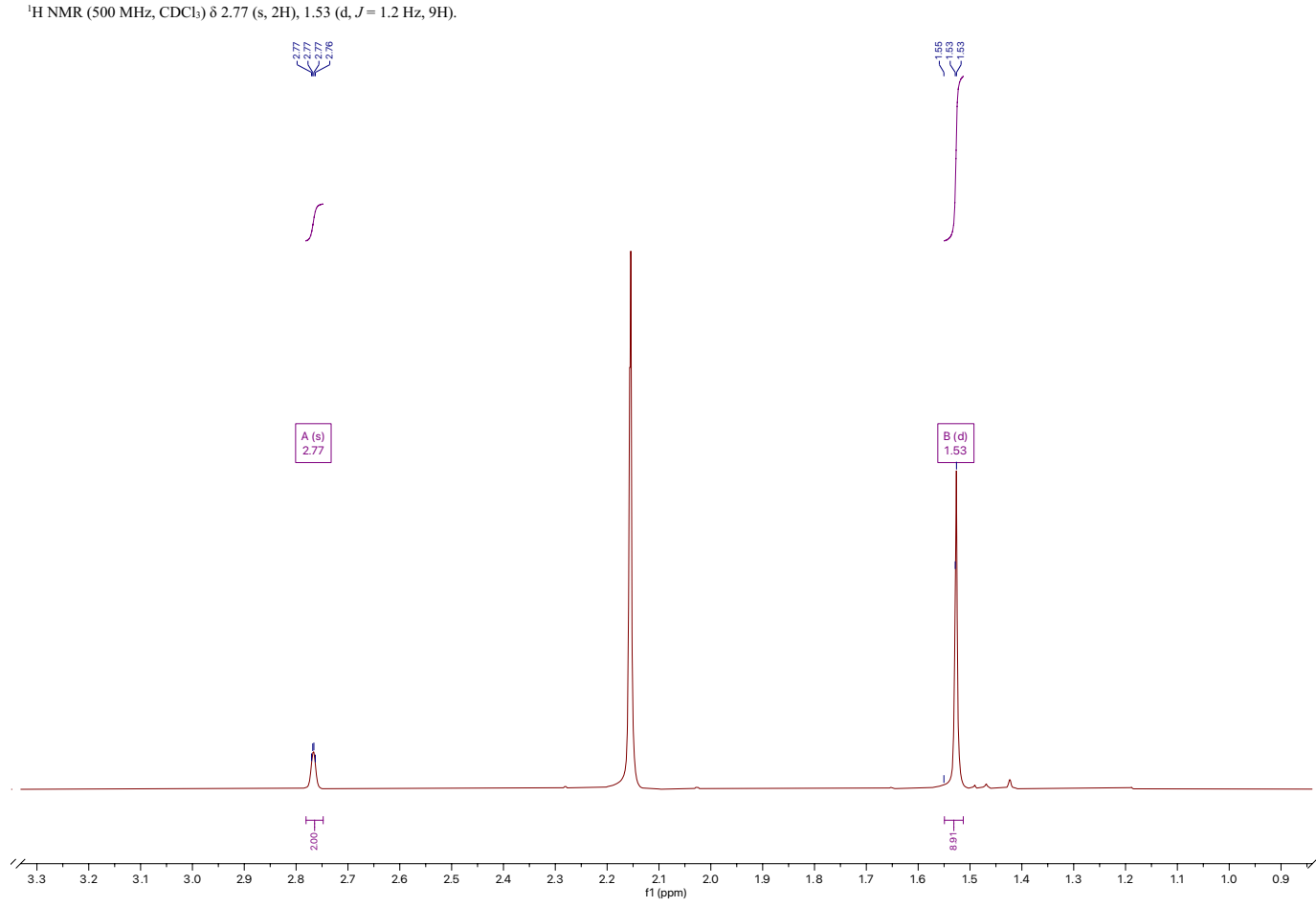
[4-¹³C₂] *tert*-Butyl 2-methylenebutanoate 22

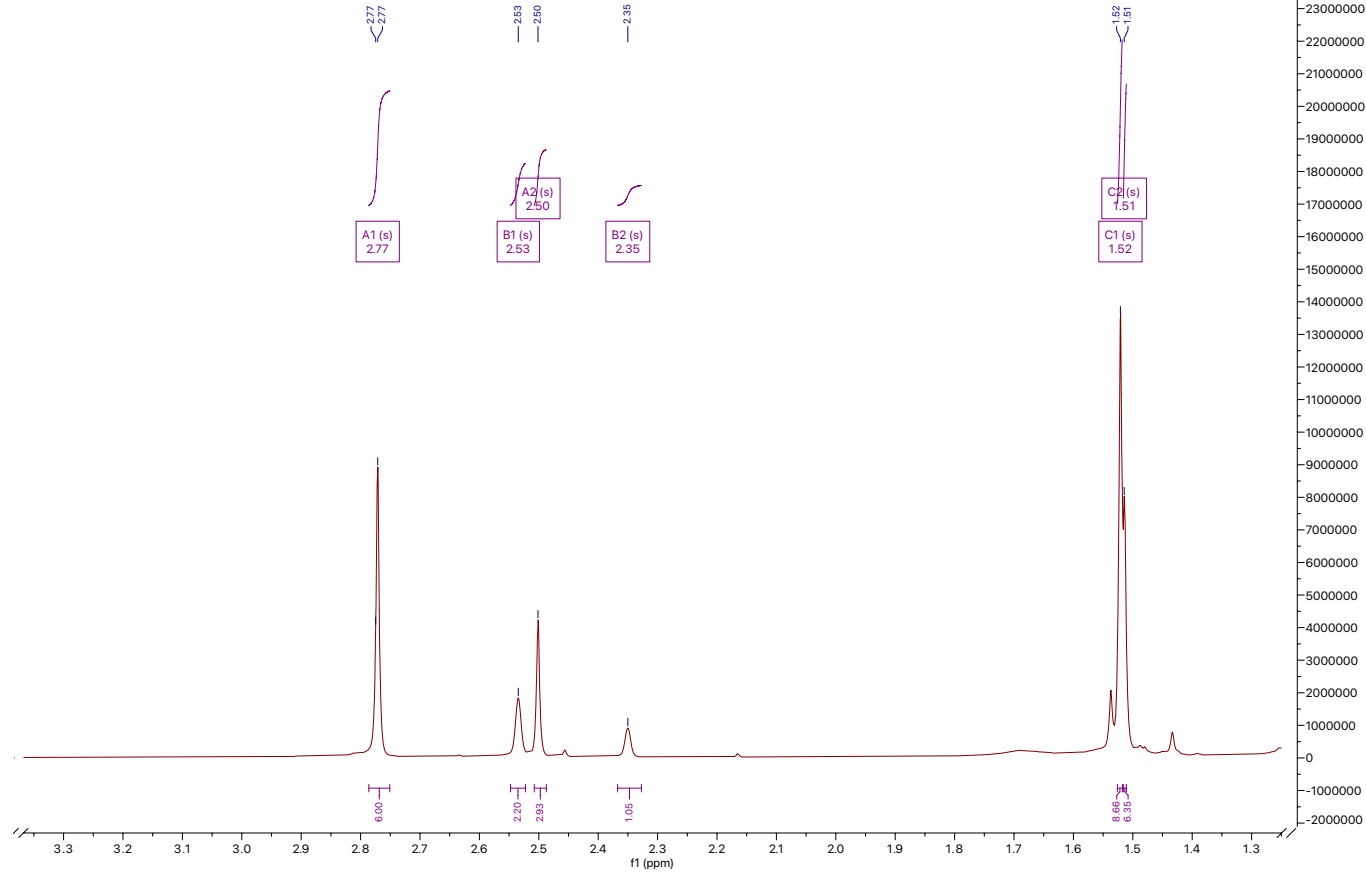
[4-¹³C₂] *tert*-Butyl 2-ketobutanoate **23**

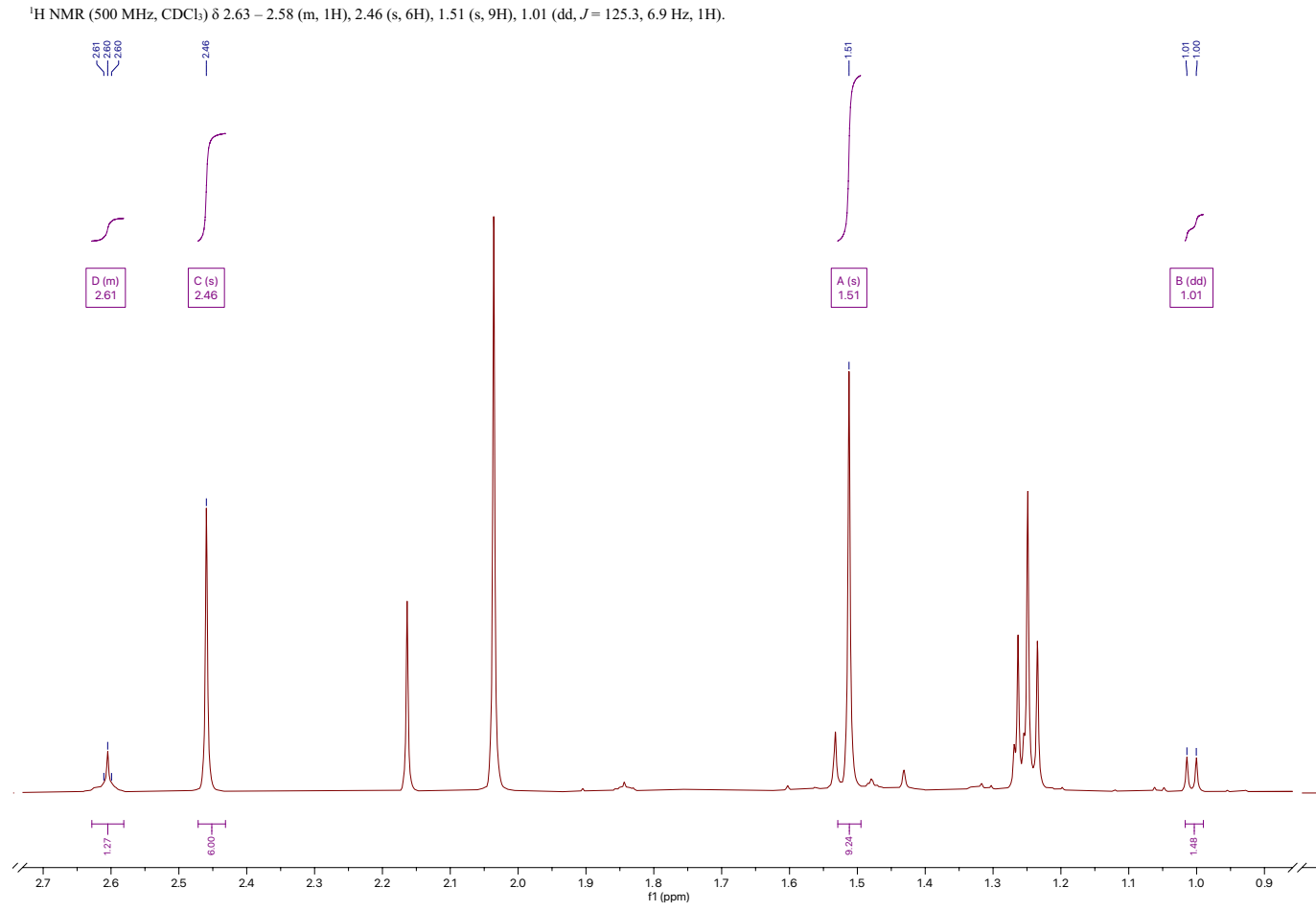


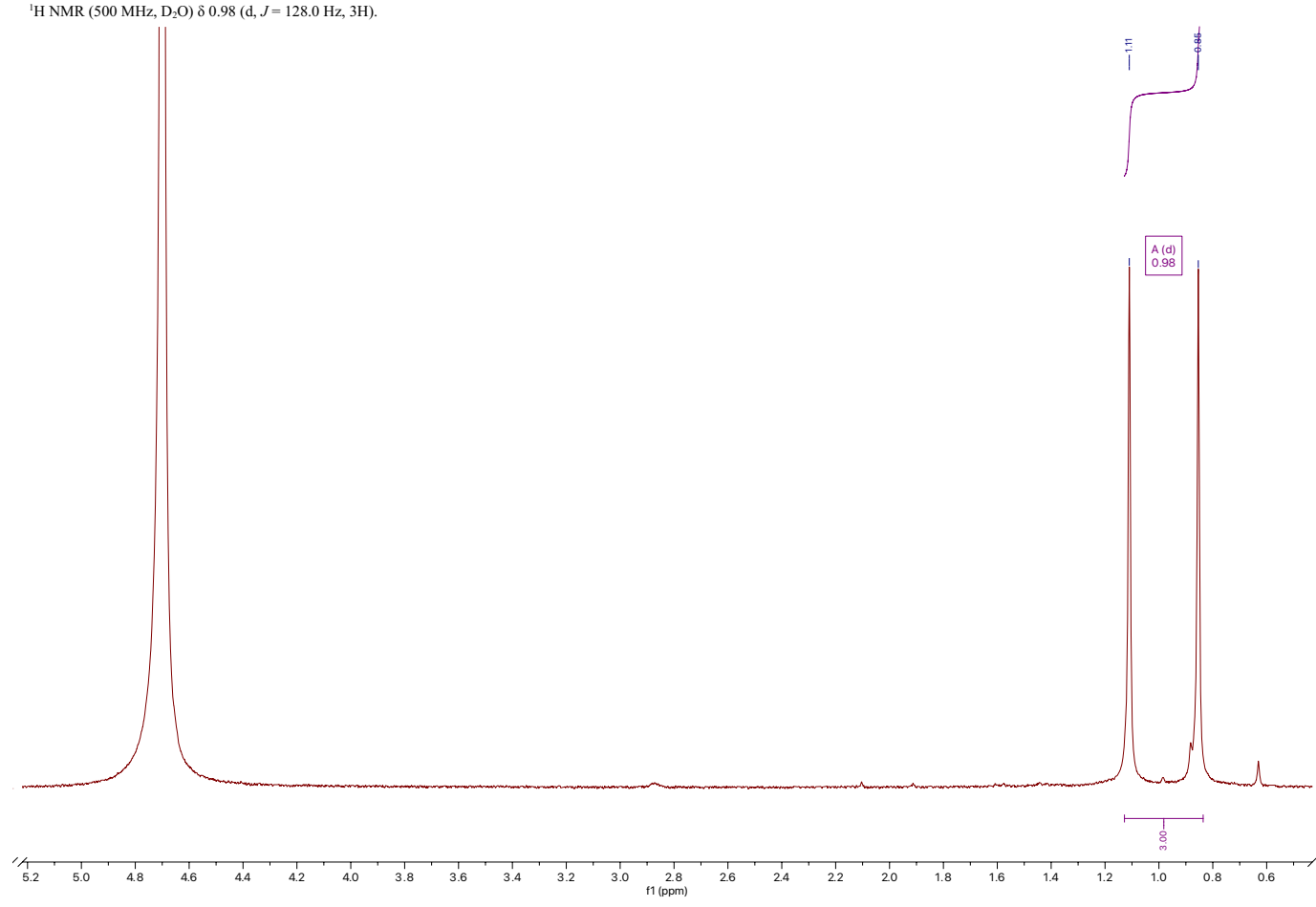
[4-¹³C₂, 3-D₂] 2-Ketobutanoic acid 24

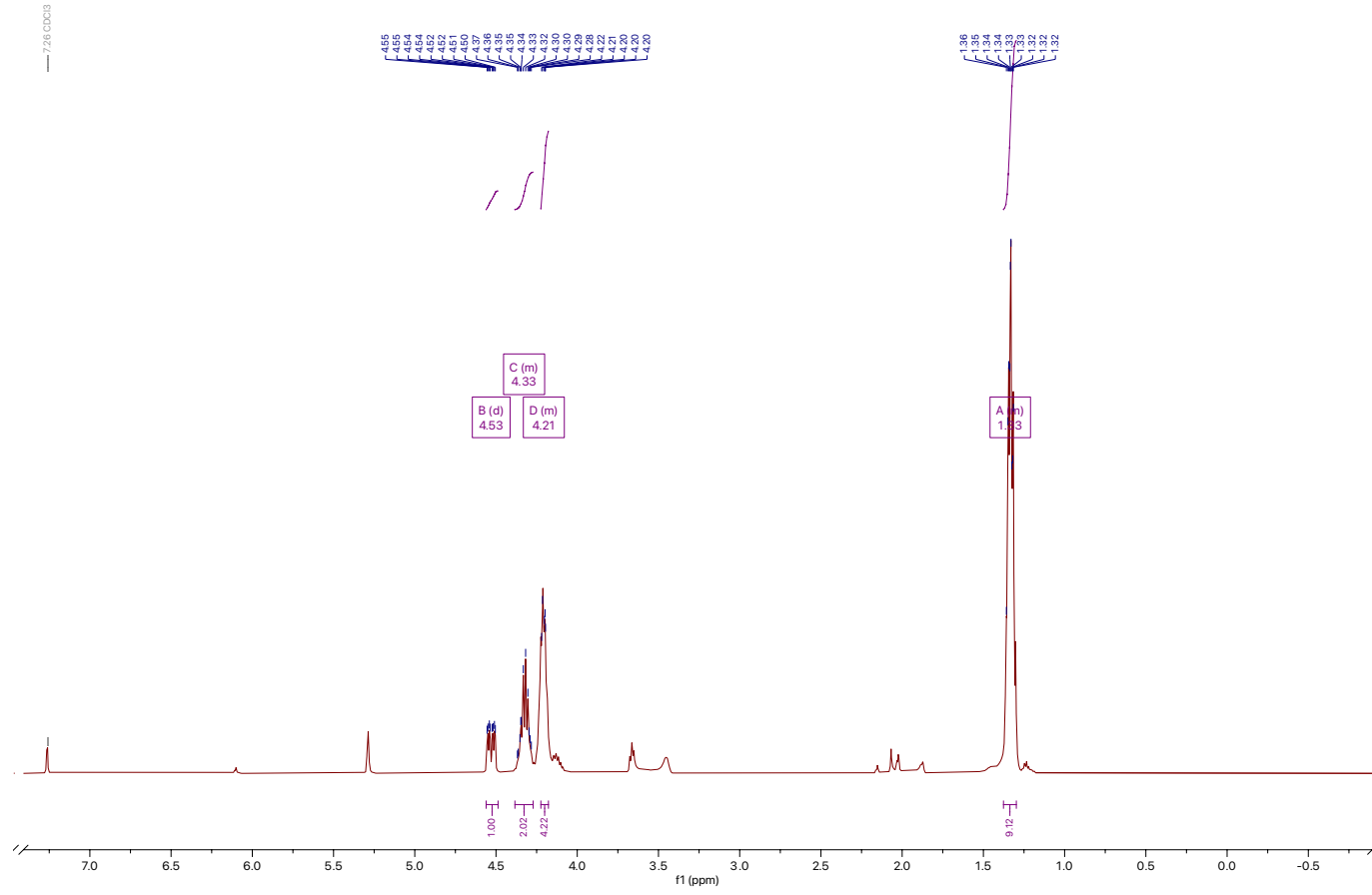


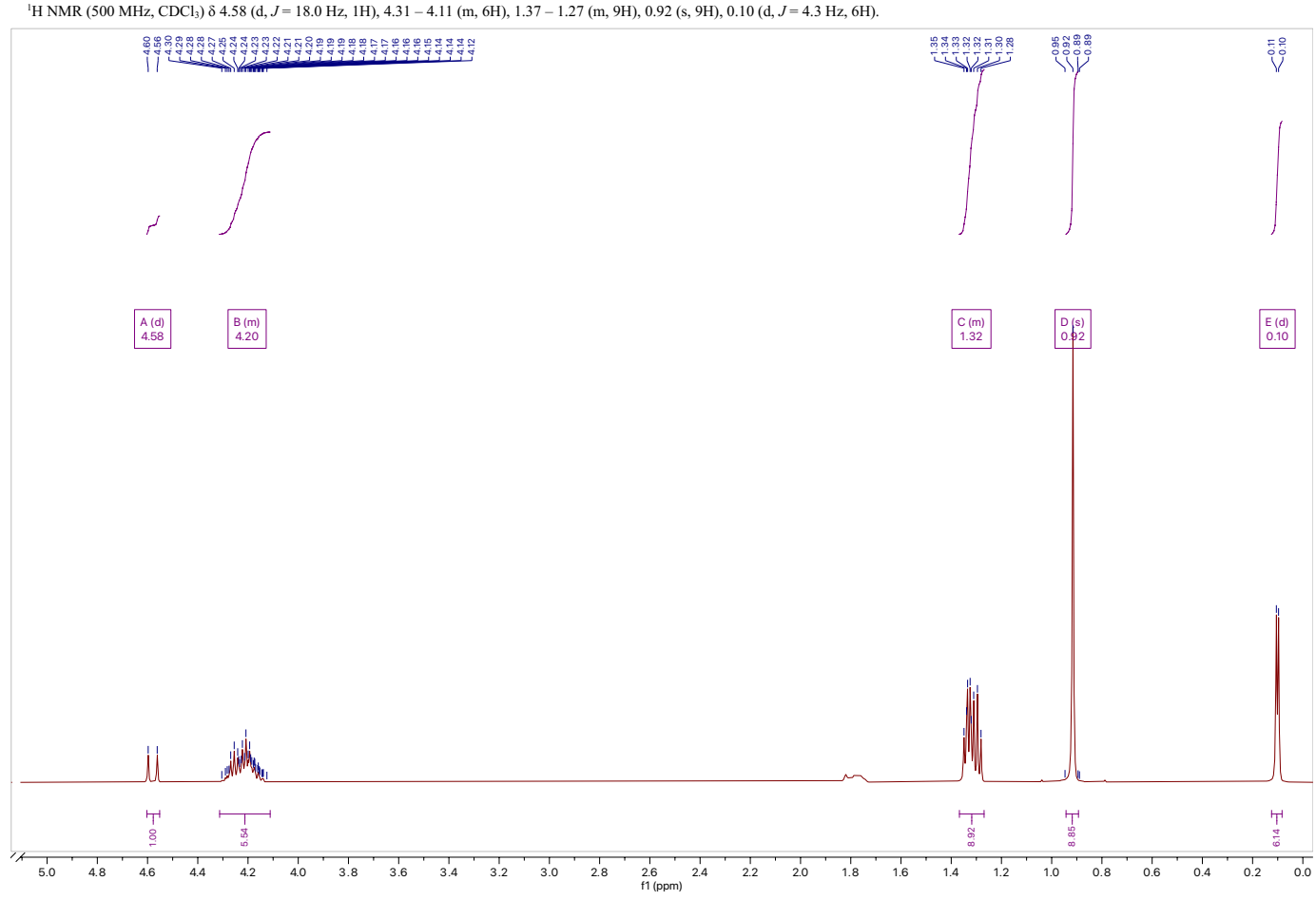
[4-CD₃] *tert*-Butyl 2-ketobutanoate 26

[4-CD₃] *tert*-Butyl 2-(2,2-dimethylhydrazono)butanoate 27¹H NMR (500 MHz, CDCl₃) δ 2.77 (s, 6H), 2.53 (s, 2H), 1.52 (s, 9H).¹H NMR (500 MHz, CDCl₃) δ 2.50 (s, 6H), 2.35 (s, 2H), 1.51 (s, 9H).

[4-¹³C; 4-CD₃] *tert*-Butyl 2-(2,2-dimethylhydrazono)-3-methylbutanoate 28

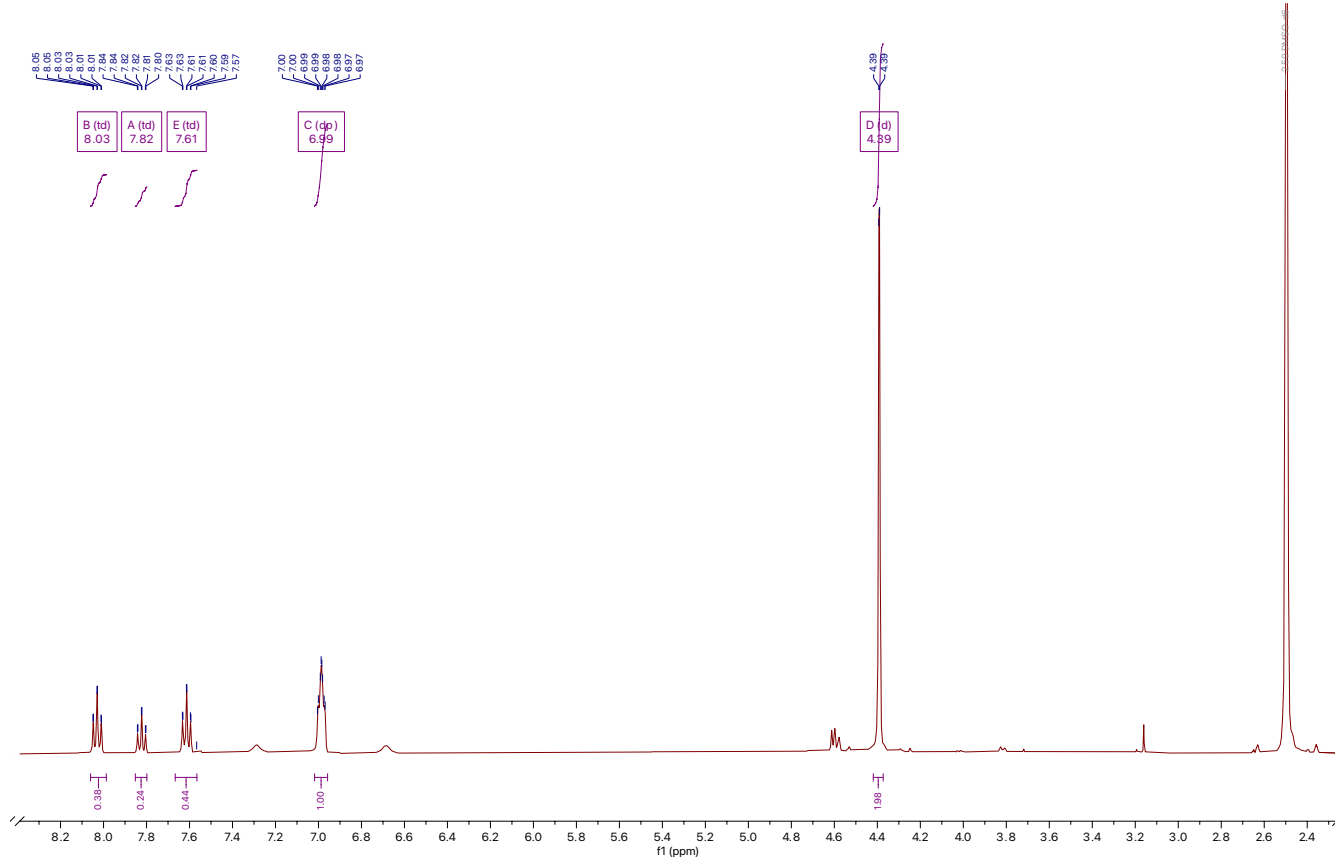
[4-¹³C; 4-CD₃] Ketoisovaleric acid 30

Ethyl 2-(diethoxyphosphoryl)-2-hydroxyacetate **32** $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 4.53 (d, $J = 15.3$ Hz, 1H), 4.38 – 4.27 (m, 2H), 4.22 – 4.18 (m, 4H), 1.38 – 1.30 (m, 9H).

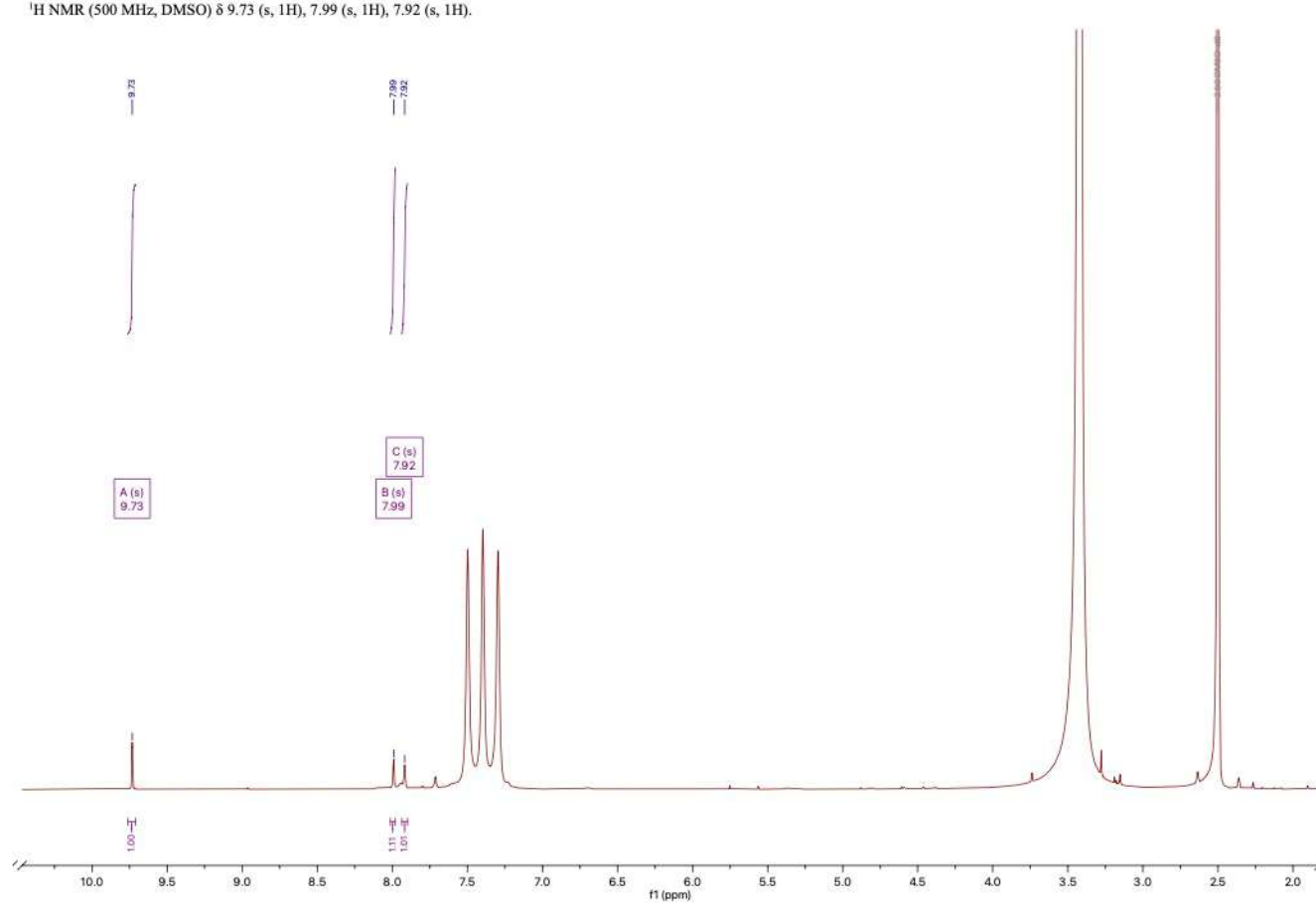
Ethyl 2-((*tert*-butyldimethylsilyl)oxy)-2-(diethoxyphosphoryl)acetate **33**

[1,3-¹⁵N, 2-¹³C](Imidazol-1*H*-5-yl)methanol 34

¹H NMR (500 MHz, DMSO) δ 8.03 (dtd, *J* = 209.5, 9.4, 1.2 Hz, 0.75H), 7.82 (td, *J* = 9.3, 1.2 Hz, 0.25H), 6.99 (dp, *J* = 7.6, 2.3 Hz, 1H), 4.39 (d, *J* = 2.0 Hz, 2H).

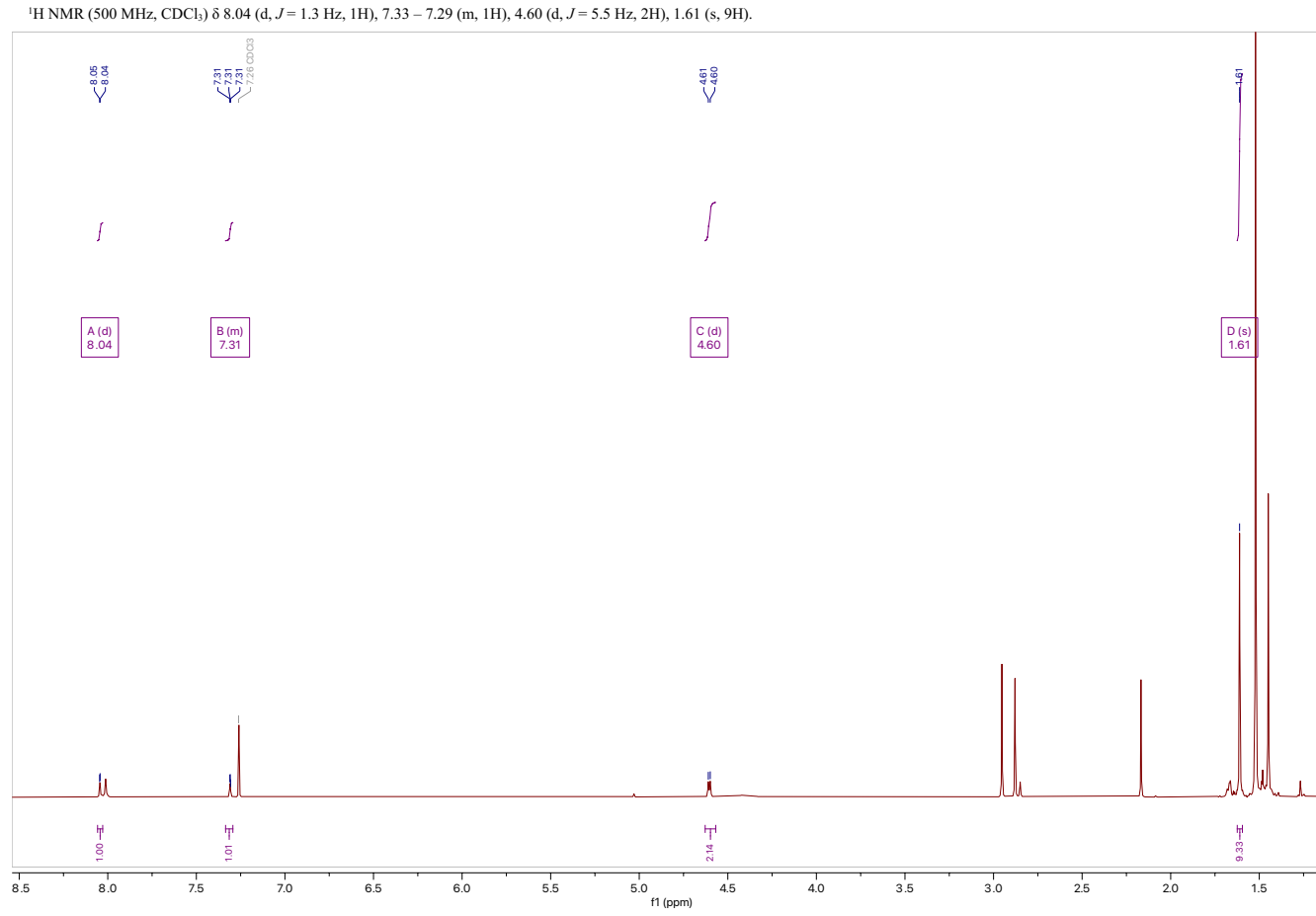


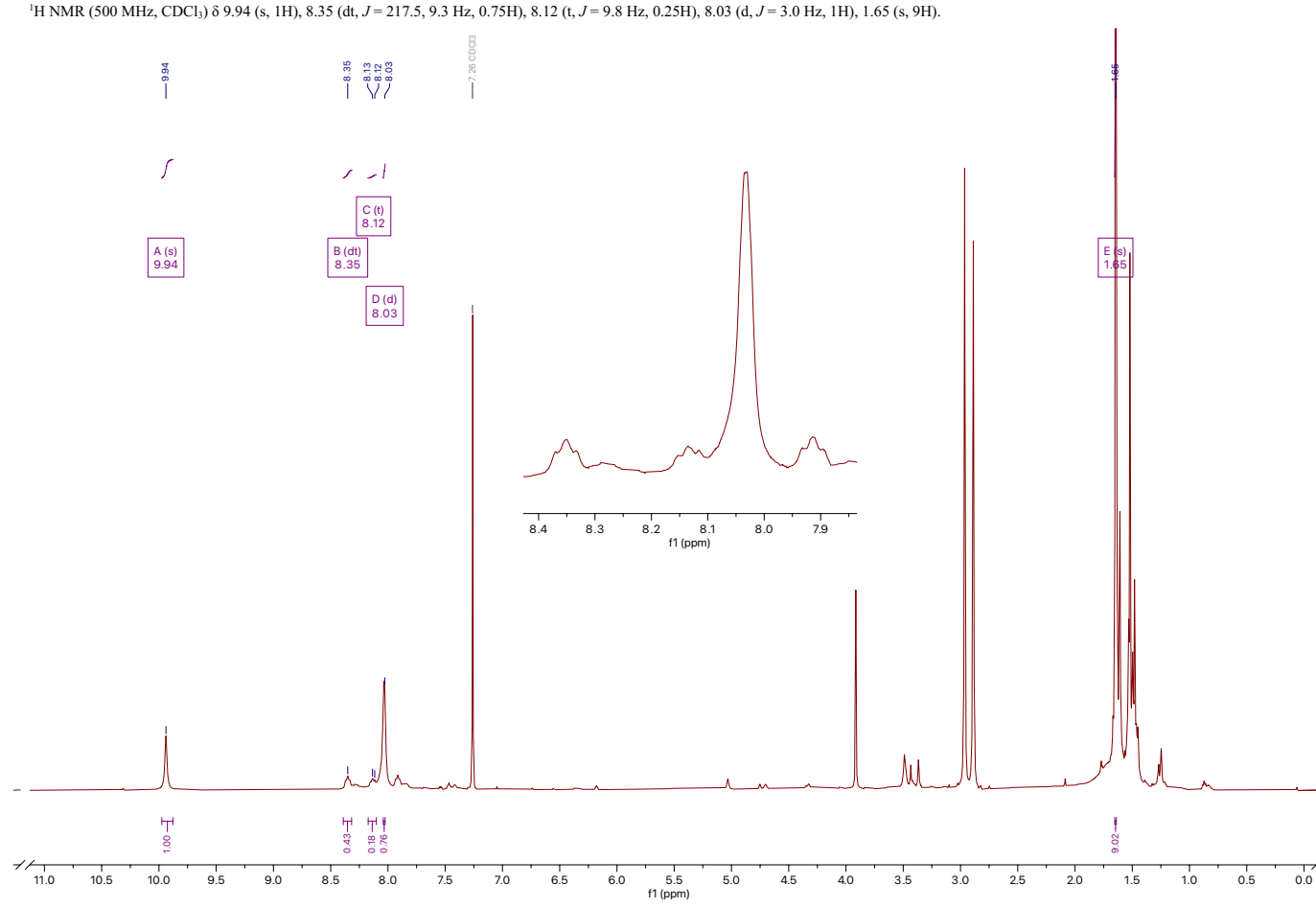
1H-imidazole-5-carbaldehyde 35

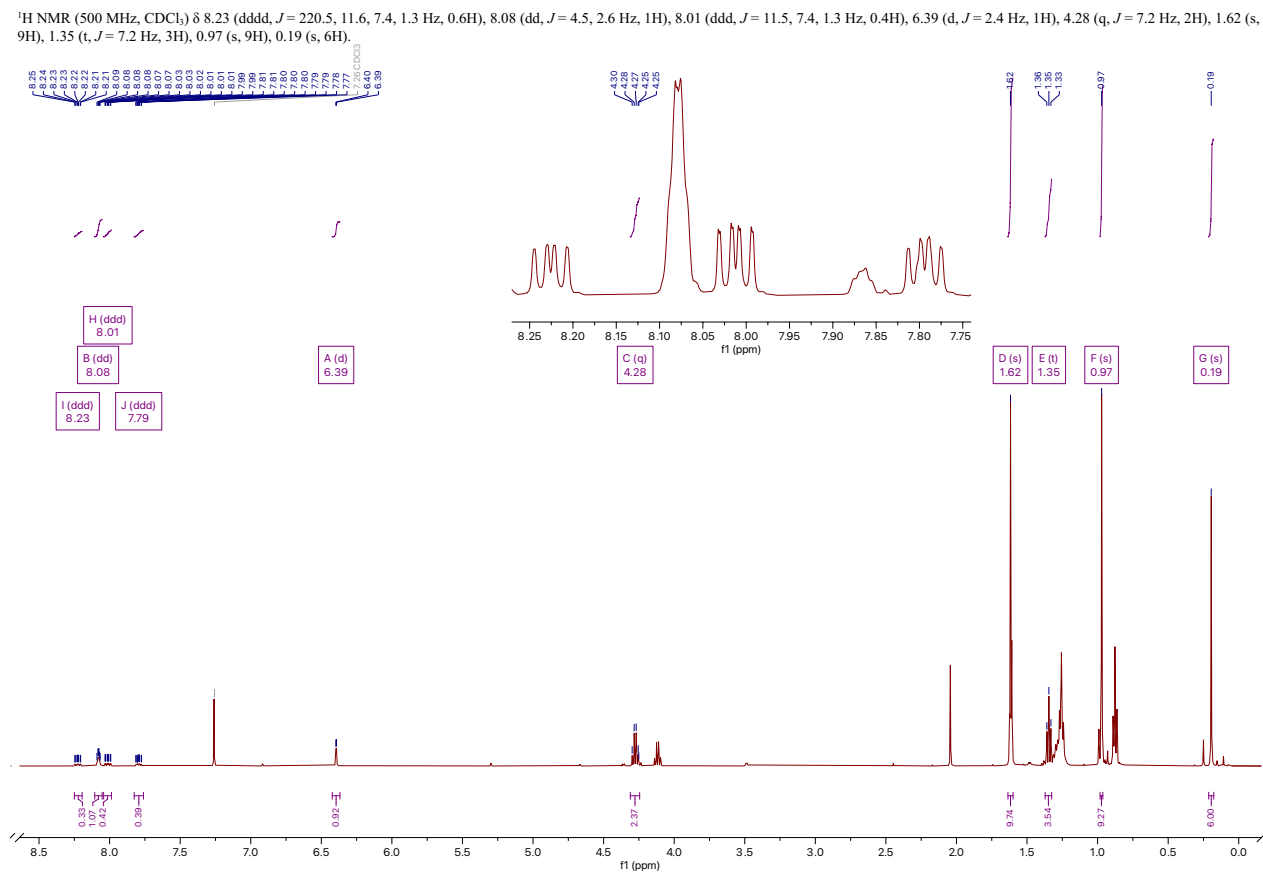


tert-Butyl-imidazole-4-yl-methanol **36**

The unlabeled analog of **36** is shown.



tert-Butyl-[1,3-¹⁵N, 2-¹³C]imidazole-1-carboxylate 37

tert-Butyl-4-(2-((*tert*-butyldimethylsilyl)oxy)-3-ethoxy-3-oxoprop-1-en-1-yl)-[1,3-¹⁵N, 2-¹³C]-imidazole-1-carboxylate 38

[1,3-¹⁵N, 2-¹³C]4-(2-Carboxy-2-hydroxyvinyl)-1*H*-imidazolium chloride 39

