universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„NMR Spectroscopy in Drug Discovery“

verfasst von / submitted by

Theresa Höfurthner, BSc MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktorin der Naturwissenschaften (Dr.rer.nat)

Wien, 2024 / Vienna 2024

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 796 605 419
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Chemie
Betreut von / Supervisor:	Univ.-Prof. Dr. Robert Konrat

SUPERVISOR:
Univ.-Prof. Dr. Robert Konrat

LOCATION:
Vienna, Austria

TIME FRAME:
October 2020 - February 2024

Dedicated To My Family

Walking with a friend in the dark
is better than walking alone in the light.

— Helen Keller

ACKNOWLEDGMENTS

Thank you for walking with me:

- to my supervisor Robert Konrat - for your encouragement, and your scientific and emotional support throughout the years. For the freedom and trust you gave me. For just being awesome!
- to my co-superivsor at Boehringer: Moriz Meier and Leonard Geist. Thank you for your expertise, coolness, and excitement. I will miss our meetings.
- to the people at UHN, UfT, and SGC: Cheryl Arrowsmith, Shili Duan, Scott Houliston, and all the people working there that made my stay unforgettable.
- to my co-authors: Roman Lichtenecker, Georg Kontaxis, Giorgia Toscano. Doing science with you was so much fun, thank you for your patience, guidance, and expertise!
- to my PhD-sister Irene Ceccolini: no words can describe how grateful I am to have you (and Salvatore and Ada) in my life.
- to my colleagues that turned into a wonderful friendship: Andreas Beier, Karin Ledolter, Julian Holzinger, Mario Migotti, Sonja Knödelstorfer, Thomas Schwarz, Virginia Busetto, Gerald Platzer, Clemens Kauffmann, Sven Brüschweiler, Ola Ptaszek. For the amazing human beings that you are!
- to my partner-in-crimes: Katharina Siess and Sonya Widen. You made the toughest days bearable and fun again. Thank you for spending so many hours with me in our second living room. Also, shout out to the best bartender ever - Pat!
- to my mother: for being my rock!
- to my siblings: Matthias and Lisa, Johanna, Laura and Alex, for always having my back!
- to my father: your laugh will forever resonate with me, I miss you!

PUBLICATIONS

1. G. Toscano, T. Höfurthner, B. Nagl, A. Beier, M. Mayer, L. Geist, DB. McConnell, H. Weinstabl, R.J. Lichtenecker, and R. Konrat, "$^{13}C\beta$-Valine and $^{13}C\gamma$-Leucine Methine Labeling To Probe Protein-Ligand Interaction", *ChemBioChem*, 2024, online ahead of print.

2. T. Höfurthner, G. Toscano, G. Kontaxis, A. Beier, M. Mayer, L. Geist, DB. McConnell, H. Weinstabl, R.J. Lichtenecker, and R. Konrat, "Synthesis of a ^{13}C-methylene-labeled isoleucine precursor as a useful tool for studying protein side-chain interactions and dynamics", *J Biomol NMR*, 2023, online ahead of print.

3. T.C. Schwarz, A. Beier, K. Ledolter, T. Gossenreiter, T. Höfurthner, M. Hartl, T.S. Baker, R.J. Taylor, and R. Konrat, "High-resolution structural information of membrane-bound α-synuclein provides insight into the MoA of the anti-Parkinson drug UCB0599", *PNAS*, vol. 120, no. 15, 2023.

4. T. Höfurthner, B. Mateos, and R. Konrat, "On-Cell NMR Contributions to Membrane Receptor Binding Characterization", *ChemPlusChem*, vol. 86, no. 6, pp. 938-945, 2021.

ABSTRACT

The aim of medicinal chemists is the design of the best suitable drug for a target protein, with efficient allocation of all available resources. Using structures of protein-ligand complexes generated with X-ray crystallography or cryo-electron microscopy (cryo-EM) as a starting point, the ligands are modified in order to achieve the highest binding affinity and efficacy. As these structures only depict a single snapshot of the protein-ligand binding event, important information about the thermodynamic properties of the complex formation gets lost. An additional source of information is therefore needed in order to properly describe the binding complex. Nuclear Magnetic Resonance (NMR) Spectroscopy is able to detect protein-ligand interactions at atomic resolution under near-physiological conditions, which delivers important data for drug discovery programs by extracting binding site and kinetic information. Moreover, the usage of selectively labeled precursors decreases the complexity of protein spectra, as only certain carbon and hydrogen atoms are isotopically labeled, thus allowing the investigation of side chain dynamics. Additionally, certain noncovalent interactions can also be detected by employing selective labeling schemes, such as CH-π interactions. Here, two novel precursors for Valine, Leucine and Isoleucine, and their incorporation into the bromodomain 1 of the bromodomain-containing protein 4 (Brd4-BD1) are described and their applicability to detect CH-π interactions, in which the CH groups of the protein serve as the donor and the ligand aromatic moieties as π acceptors. The direct probing of these noncovalent interactions by NMR Spectroscopy together with the usage of computational methods will accelerate drug discovery programs.

ZUSAMMENFASSUNG

Ziel der medizinischen Chemiker ist es, das am besten geeignete Arzneimittel für ein Zielprotein zu entwerfen und dabei alle verfügbaren Ressourcen effizient zu nutzen. Ausgehend von Strukturen von Protein-Ligand-Komplexen, die mit Röntgenkristallographie oder cryo-Elektronenmikroskopie (cryo-EM) erstellt wurden, werden die Liganden modifiziert, um die höchste Bindungsaffinität und Wirksamkeit zu erreichen. Da diese Strukturen nur eine einzige Momentaufnahme des Protein-Liganden-Bindungsvorgangs darstellen, gehen wichtige Informationen über die thermodynamischen Eigenschaften der Komplexbildung verloren. Daher wird eine zusätzliche Informationsquelle benötigt, um den Bindungskomplex richtig zu beschreiben. Die Kernspinresonanzspektroskopie (NMR) ist in der Lage, Protein-Ligand-Wechselwirkungen mit atomarer Auflösung unter nahezu physiologischen Bedingungen nachzuweisen, was wichtige Daten für Programme zur Entdeckung von Arzneimitteln liefert, indem Informationen über die Bindungsstelle und die Kinetik extrahiert werden. Darüber hinaus verringert die Verwendung selektiv markierter Ausgangsstoffe die Komplexität von Proteinspektren, da nur bestimmte Kohlenstoff- und Wasserstoffatome isotopisch markiert sind, was die Untersuchung der Seitenkettendynamik ermöglicht. Darüber hinaus können bestimmte nicht-kovalente Wechselwirkungen auch durch selektive Markierungsverfahren nachgewiesen werden, wie z. B. CH-π-Wechselwirkungen. Hier werden zwei neuartige Vorläufer für Valin, Leucin und Isoleucin, und ihr Einbau in die Bromodomäne 1 des Bromodomänen-enthaltenden Proteins 4 (Brd4-BD1) beschrieben und ihre Anwendbarkeit zum Nachweis von CH-π-Wechselwirkungen, bei denen die CH-Gruppen des Proteins als Donor und die aromatischen Ligandenanteile als π-Akzeptoren dienen. Die direkte Untersuchung dieser nicht-kovalenten Wechselwirkungen durch NMR-Spektroskopie zusammen mit dem Einsatz von computerunterstützten Berechnungsmethoden wird die Entdeckung von Medikamenten beschleunigen.

CONTENTS

LIST OF FIGURES

Part I

INTRODUCTION

1 INTRODUCTION

What is the most suitable drug for a target protein? Answering this question seems like a heavy-handed task and is dependent on multiple different events that take place in a protein-drug interaction. These events can range from protein folding/unfolding upon binding, the binding event itself, domain motions, protein side chain rotations, loop motions, ring flips, and bond rotation. In addition to the different types of events, they also occur on different timescales, from ps to s, which are dependent on the thermodynamic properties of the involved proteins [1]–[5]. Medicinal chemists aim to design the very best molecule that interacts with the target protein by putting all the available information together, starting from the structure of the protein to optimizing the ligand with the goal of achieving the highest binding affinity and efficacy. To date, three-dimensional (3D) protein structures are predominately determined by X-ray crystallography and cryo-electron microscopy (cryo-EM). Together with the new field of artificial intelligence-based structure prediction tools, like AlphaFold, structural information will be available for every globular, properly folded protein target, which is a major advantage for all structure-based drug design programs [6]. Nevertheless, this approach suffers under the condition that they are static images and can therefore not include the dynamic properties of the ligand or protein. Additionally, structures derived from the above-mentioned methods do not detect hydrogen atoms, and they are therefore added during processing and refinement according to their expected geometries. This is sufficient for the amide hydrogen (H^N) and hydrogens of the side chains such as tryptophan, asparagine, and glutamine, but lacks important data or gives only indirect information about other hydrogen atoms present in the protein, especially in terms of hydrogen bond formation [7].

Nuclear Magnetic Resonance (NMR) Spectroscopy has become one of the most essential techniques in drug discovery programs, as it is very feasible to probe structural and dynamical properties of protein-ligand interactions [8]–[13]. As binding events take place at different time scales, depending on what kind of motions are involved, they can be probed by different NMR experiments (Figure 1).

This doctoral thesis and the presented papers try to highlight the importance of hydrogen atoms in the context of ligand binding and their binding affinity by using selective labeling strategies and robust NMR experiments. Moreover, it showcases the incorporation of two novel precursors for leucine and valine methine residues ([3-^{13}C,4,4'-

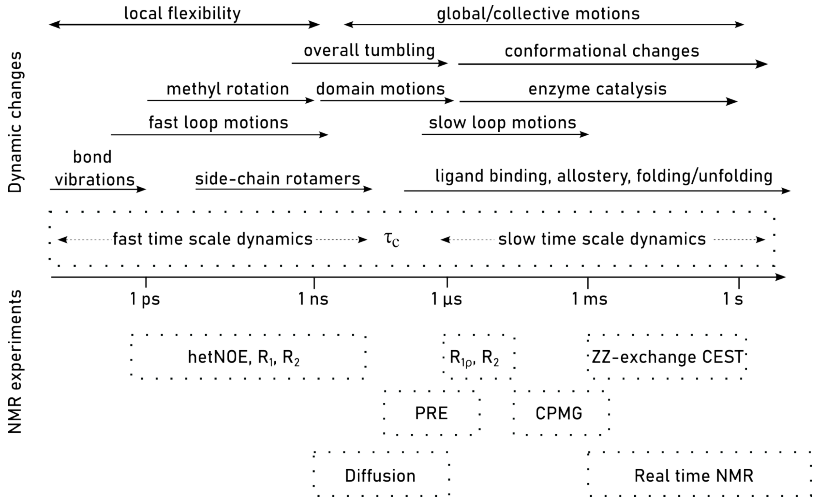


Figure 1: Overview of different time scales and NMR experiments adapted from [2]

$^{2}H_6$] ketoisovaleric acid) and isoleucine methylene ([3-^{13}C,4,4,4-$^{2}H_3$] 2-ketobutyrate) residues into the bromodomain 1 of the bromodomain-containing protein 4 (Brd4-BD1) and their applicability to probe weak hydrogen bonds, in particular CH-π interactions. The following section aims to give the reader a short but comprehensive introduction to the field of protein-ligand interactions by NMR Spectroscopy.

1.1 LIGAND BINDING

Molecular recognition is the mechanism by which two or more molecules form a specific complex through non-covalent interactions (NCIs) and therefore an essential part of life [14]. In this context, two important key players have to be mentioned: specificity and affinity of the interaction partners. First, binding affinity and its underlying thermodynamic properties will be discussed briefly.

1.1.1 *Binding Affinity*

For a one-site binding event, its kinetics are usually described as in Eq. (1), where the free protein and ligand are in equilibrium with the formed complex. The formation of the complex is usually dependent on their assembly and disassembly rate, k_{on} and k_{off}, respectively. If the system is in equilibrium, the forward binding reaction is balanced by the reverse unbinding reaction (Eq. (2)).

$$[P] + [L] \rightleftharpoons [PL] \tag{1}$$

$$k_{on}[P][L] = k_{off}[PL] \tag{2}$$

The binding rate, or association rate (k_{on}), and the dissociation rate (k_{off}) form the binding affinity (Eq. 3).

$$K_D = \frac{k_{off}}{k_{on}} \tag{3}$$

The above-depicted equations give the impression that protein-ligand interactions are taking place in "splendid isolation", which is likely never true in a biological context. It must be considered that the protein and the ligand are in aqueous matter. Thermodynamically speaking, the process of ligand binding is therefore dependent on various different factors, such as buffer components, pH, and temperature, which yield to one simplified read-out, the Gibbs free energy (ΔG). The binding kinetics of an interaction is linked to its Gibbs free energy via its rate constants, k_{on} and k_{off} or dissociation constant K_D. If thermodynamically controlled, an interaction will only be formed if ΔG is negative, and we can therefore assume that these parameters determine the stability of the formed complex [15]. As depicted in Eq. (4), ΔG can also be dissected into its enthalpic (ΔH) and entropic term (-TΔS). These two thermodynamic properties are highly dependent on the system-specific properties.

$$\Delta G = \Delta H - T\Delta S \tag{4}$$

In order to have a better understanding of what actually determines the binding affinity, we shall dissect these two properties even more. The enthalpy contribution is generally treated as the formation of non-covalent interactions resulting from the protein-ligand interaction. As mentioned above, interactions are taking place in a solute, meaning that pre-existing non-covalent interactions must be disrupted or reorganized in order for the protein binding to a ligand, e.g. gain/loss of hydrogen bonds [16]–[18]. This is also true for the entropic contribution to the overall binding free energy. ΔS can be re-written into these three terms:

$$\Delta S = \Delta S_{solv} + \Delta S_{conf} + \Delta S_{r/t} \tag{5}$$

All three terms determine the net entropy change that contribute to favorably or unfavorably binding free energy, in which ΔS_{solv} represents changes in the solvent. ΔS_{conf} describes the contribution that arises from changes in the conformational freedom of protein and ligand binding and $\Delta S_{r/t}$ represents the loss of translational and rotational degrees of freedom upon protein-ligand complex formation [19]–[21]. All of these different kinds of contributions have to be taken into account when finding the best suitable ligand for a target. Any given modification on the ligand might be favorable in terms of enthalpy, but it might be very unfavorable in terms of its entropy. Protein folding and protein-ligand complex formation are therefore a

mixture of attractive and repulsive forces. The *Enthalpy-Entropy Compensation*, which is a fundamental property in all biological reactions, is therefore unavoidable when designing a potential drug [22].
Enthalpy and entropy changes determine the sign and magnitude of the binding free energy and are therefore very closely related. Tight binding that results from multiple favorable non-covalent interactions leads to a large negative ΔH. However, this is usually accompanied by a negative entropic change, due to decreased mobility of the protein, ligand, and solute, which then leads to a medium-magnitude change in the binding free energy or vice versa [23]. As the read-out, ΔG, does not change that much, the discrimination between both contributions is important in the field of drug design. The ideal strategy is to identify what contributes to which thermodynamic factor and maximize the favorable enthalpic or entropic contribution, while minimizing its penalty [24], [25].

1.1.2 *Specificity or Non-covalent Interactions*

Multiple different non-covalent interactions (NCIs) affect the enthalpic contribution of the binding free energy, and therefore also affect the binding affinity of a ligand to its target protein [26]. The most prominent NCIs are Van der Waals, hydrogen bonds, and electrostatic interactions. The focus in this section is dedicated to a specific kind of hydrogen bond, the so-called weak hydrogen bonds. According to IUPAC [27], a hydrogen bond is formed between a hydrogen atom covalently linked to a molecule (X-H), in which X is more electronegative than H, and an atom or a group of atoms (Y) in the same or different molecule, in which there is evidence of bond formation. The hydrogen bond acceptor Y is an electron-rich region due to its lone or π electron pair orbital. This definition is also valid for weak hydrogen bonds. A study in 2017 by Freitas et. al [28] conducted a search of all protein-ligand entries deposited in the Protein Data Bank (PDB) and their underlying interactions. The fourth most frequent interactions found were weak hydrogen bonds. Figure 2 shows all hydrogen bonds found in protein-ligand interactions in the PDB. It is no coincidence that in this thesis, we focus on the most abundant interaction: the aliphatic CH to aromatic-π hydrogen bond, in which the CH group of aliphatic residues acts as the hydrogen-bond donor, whereas the π-system of the ligand acts as the hydrogen-bond acceptor [29]–[31].

1.1.3 *Fragment-based Drug Discovery (FBDD)*

The goal of medicinal chemists is to find the best suitable drug available for a specific protein target. Usually these drugs were found through High-Throughput-Screening (HTS) methods, in which thou-

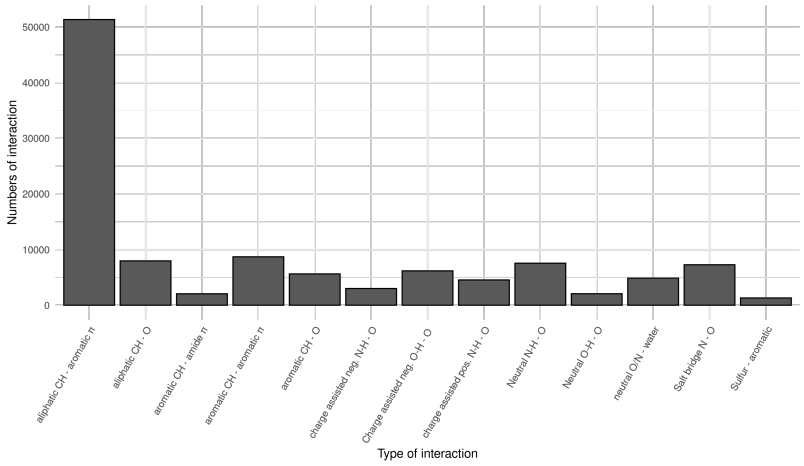


Figure 2: Types of Hydrogen bonds found in Protein-Ligand Interactions [28]

sands of different ligands, in so called ligand libraries, are screened against a target by using different types of biochemical and -physical methods. While HTS is still very commonly used in drug discovery programs, in the recent years, it seems that there is a paradigm shift towards FBDD and therefore more difficult protein targets are now accessible [32]–[34]. The underlying principle of this drug design method is the usage of a small library of small molecules, called fragments, which are screened against a target and modified throughout the process. These fragments profit from the fact that they have a low degree of complexity, are usually smaller than 300 g/mol in size, and can form a few directed favorable interactions with the target protein (less than five hydrogen bond donors and less than 10 acceptors). Usually, identified fragments represent therefore an ideal binding motif that is the perfect starting point for further optimization. Biophysical methods to characterize these low affinity binders have to be very accurate, in contrast to HTS methods, which should be accurate in identifying false positives. Once a fragment is identified, it will undergo specific chemical modifications to increase its binding affinity [12], [35]–[37].

It is of high importance to understand that we will need a method that can tease apart all the different thermodynamic contributions of the binding free energy. One single method alone is unlikely to be able to accomplish this, but a strategic combination of methods may have a very good chance of elevating drug discovery programs to the next level: Nuclear Magnetic Resonance (NMR) Spectroscopy together with computer-aided chemical shift calculations, which can be derived from Molecular Dynamics (MD) and Quantum Mechanics (QM/MM) simulations.

1.2 CHEMICAL SHIFT PERTURBATIONS

NMR Spectroscopy has always been a very useful method in drug discovery programs, as it can detect binding from the perspective of the ligand or protein using different experiments [38]. Chemical shift perturbations (CSPs) is an easy NMR technique to probe binding events by NMR spectroscopy from the target's perspective. Usually, a ^{15}N or ^{13}C uniformly labeled protein sample is measured and the ligand of interest is titrated into the protein. After each titration step, the effects upon binding are monitored by acquiring two-dimensional (2D) spectra, in which one observed atom is the hydrogen and the other one ^{15}N or ^{13}C. CSPs can be derived by a set of different 2D experiments, the most prominent being the heteronuclear single quantum coherence (HSQC) or the heteronuclear multiple-quantum coherence (HMQC). The resulting change in the chemical shift is a highly sensitive reporter of changes in the environment of the investigated nucleus, which is its sum of magnetic anisotropy around bonds, aromatic ring current effects and local electric fields from charged groups [39]. Using this accurate read-out, it is possible to derive not only the binding affinity of an interaction but also which part of the protein is affected upon binding, assuming there is a full chemical shift assignment available. However, this means that the target protein should not be bigger than 30 kDa (kilo-Dalton) in size to achieve high sensitivity and resolution. In the last decades, major improvements have been made in terms of both NMR experiments and hardware development [40]–[49]. Additionally, selective protein labeling strategies constantly expand protein analysis and increase the size of attainable biomolecules, by introducing ^{13}C, ^{15}N, and ^{2}H isotope positions in amino acid side chains . Selective isotope labeling has been developed for aromatic and aliphatic residues, while stereoselective methyl labeling for Leucine and Valine residues has also led to specifically identifying these residues [50]–[64]. The majority of available aliphatic precursors focus on introducing the ^{13}C isotope at the methyl position of each residue, as the methyl group has favorable relaxation properties due to the -CH_3 threefold symmetry, thereby increasing signal intensities. Moreover, aliphatic residues are highly abundant in hydrophobic protein cores and binding pockets, making them beneficial reporters of binding events. As the binding affinity is also highly dependent on solvent-exposed residues and their hydrogen bond formation with the solute [65], [66]. By extending the labeling strategy from the methyl to the methylene and the methine group, solvent exposure of these specific residues is less likely. The methylene group of Isoleucine ($C\gamma_1$) and the methine group of Valine and Leucine ($C\beta$, $C\gamma$) were selectively ^{13}C-^{1}H labeled, whereas other non-exchangeable protons in the side chain were deuterated to suppress through-bond coupling effects. Chemical shifts (CS) are strongly de-

pendent on their chemical environment, especially that of hydrogen. Depending on the orientation of the observed hydrogen towards aromatic ring structures, the observed CS is moving down- or upfield, due to the additional local magnetic field that is induced by aromatic ring current effects [31], [67]–[70]. Moreover, the magnitude of the resulting CSP also correlates well with the binding energy of the interaction [71]. With that, simple 2D experiments provide indirect information about the binding contribution of individual interactions and also structural information, resulting in a better understanding of the *Structure-Affinity-Relationship* [72].
In order to validate the CSPs upon ligand interaction and therefore the existence of CH-π interactions, the measured hydrogen chemical shift is compared to the calculated one. The chemical shift calculation of the hydrogen is performed by Eq. (6),

$$\Delta\sigma = 10^6 \frac{ne^2a^2}{4\pi mc^2} \frac{3\cos^2\theta - 1}{r^3} \tag{6}$$

in which $\Delta\sigma$ is the change of the isotropic nuclear shielding constant in ppm. The first term of the equation contains constant values; n is the number of circulating electrons, e is the elementary charge (Franklin) and a is the diameter of the aromatic ring (cm). The second term contains the distance restraint between the hydrogen and the aromatic ring center (r) and the angular dependency between the ring normal through the aromatic ring center and the proton to the ring center vector (θ). This approximation describes the effect a single proton signal is experiencing by a secondary magnetic field, which is induced by the electric ring current via a magnetic dipole positioned at the ring center [73].

1.3 SUMMARY OF THE THESIS

This thesis explores the potential of two novel precursors and their applicability in detecting CH-π interactions. The precursor [3-^{13}C,4,4,4-$^{2}H_3$] 2-ketobutyrate to selectively label the methylene group of Isoleucine is described. Isoleucine Cγ_1-H_2 side chain dynamics were studied, as well as strong ligand binders, illustrating their feasibility in detecting sensitive information. In paper number two, synthesis route of [3-^{13}C,4,4'-$^{2}H_6$] ketoisovaleric acid was presented and successful incorporation into the methine groups of Valine and Leucine residues was demonstrated. In this paper, a total set of 10 ligands were used to study CH-π interactions.

Part II

SHOWCASE OF PUBLISHED WORK

The central part of this cumulative thesis consists of two scientific articles about site-specific labeling of different types of aliphatic residues in the protein Brd4-BD1 and their applicability to measure CH-π interactions. The articles have been published in peer-reviewed scientific journals. Aside from the continuous page numbering, the original layout design should be considered intellectual property of the respective journals. Supporting figures and tables are also included in the original format. Their respective digital object identifiers (DOIs) are provided together with a short statement of personal contribution at the beginning of each chapter.

2

^{13}C-METHYLENE-LABELED ISOLEUCINE PRECURSOR

T. Höfurthner, G. Toscano, G. Kontaxis, A. Beier, M. Mayer, L. Geist, D. B. McConnell, H. Weinstabl, R. Lichtenecker, and R. Konrat, "Synthesis of a 13C-methylene-labeled isoleucine precursor as a useful tool for studying protein side-chain interactions and dynamics," *Journal of Biomolecular NMR*, 2023

DOI: 10.1007/s10858-023-00427-2

PERSONAL CONTRIBUTION

Protein production, data analysis, validation, writing
Shared First Author with G. Toscano

Journal of Biomolecular NMR
https://doi.org/10.1007/s10858-023-00427-2

ARTICLE

Synthesis of a ^{13}C-methylene-labeled isoleucine precursor as a useful tool for studying protein side-chain interactions and dynamics

Theresa Höfurthner[1,3] · Giorgia Toscano[2,3] · Georg Kontaxis[4] · Andreas Beier[1] · Moriz Mayer[5] · Leonhard Geist[5] · Darryl B. McConnell[5] · Harald Weinstabl[5] · Roman Lichtenecker[2] · Robert Konrat[1,4]

Received: 5 July 2023 / Accepted: 28 September 2023

Abstract
In this study, we present the synthesis and incorporation of a metabolic isoleucine precursor compound for selective methylene labeling. The utility of this novel α-ketoacid isotopologue is shown by incorporation into the protein Brd4-BD1, which regulates gene expression by binding to acetylated histones. High quality single quantum ^{13}C−^{1} H-HSQC were obtained, as well as triple quantum HTQC spectra, which are superior in terms of significantly increased ^{13}C-T_2 times. Additionally, large chemical shift perturbations upon ligand binding were observed. Our study thus proves the great sensitivity of this precursor as a reporter for side-chain dynamic studies and for investigations of CH-π interactions in protein-ligand complexes.

Keywords Precursor labeling · Transversal relaxation studies · CH-π interaction · Ligand binding

Introduction

Nuclear Magnetic Resonance (NMR) Spectroscopy has matured into a powerful tool to characterize interactions between biological molecules and their ligands at atomic resolution. Although originally described as a versatile technique to screen for potential protein binders (Gossert and Jahnke 2016; Harner et al. 2017; Luchinat et al. 2020; Gronenborn 2022), more sophisticated detection schemes have now emerged. Ligand-based detection schemes, such as waterLOGSY, Saturation Transfer Difference (STD) spectroscopy and ^{15}N–^{13}C-filtered ^{1}H-1D spectra, can reveal valuable information about site specific interaction patterns and their mechanistic details (Dalvit et al. 2000; Mayer and Meyer 2001). Together with protein NMR-based mapping of ligand binding, this unveils unique information about the characteristics of protein-ligand complexes. The potential of protein NMR for drug discovery and design campaigns was realized by the discovery of Structure-Affinity-Relationship (SAR) by NMR by the Fesik group, in which ^{1}H–^{15}N-HSQCs on uniformly labeled proteins are recorded in order to investigate ligand binding interactions (Shuker et al. 1996). A particularly powerful application of SAR is Fragment-based drug discovery (FBDD) (Rees et al. 2004). FBDD starts with small fragments (< 300 g/mol) of weak binders (mM-100 µM), which are then structurally optimized throughout the process in order to maximize non-covalent interactions, most importantly the Van der Waals, electrostatic, and hydrogen bond interactions. Despite many past applications, backbone (^{1}H–^{15}N) detection schemes do not fully harness the potential of NMR spectroscopy in drug design campaigns due to limited experimental sensitivity, spectral overlap, and the ambiguous relationship between

✉ Roman Lichtenecker
roman.lichtenecker@univie.ac.at

✉ Robert Konrat
robert.konrat@univie.ac.at

1 Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Department of Structural and Computational Biology, Max Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, 1030 Vienna, Austria

2 Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Institute of Organic Chemistry, University of Vienna, Währingerstraße 38, 1090 Vienna, Austria

3 Vienna Doctoral School in Chemistry (DoSChem), University of Vienna, Währingerstraße 42, 1090 Vienna, Austria

4 Department of Structural and Computational Biology, Max Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, 1030 Vienna, Austria

5 Boehringer Ingelheim RCV GmbH & Co. KG, Dr. Boehringer Gasse 5-11, 1121 Vienna, Austria

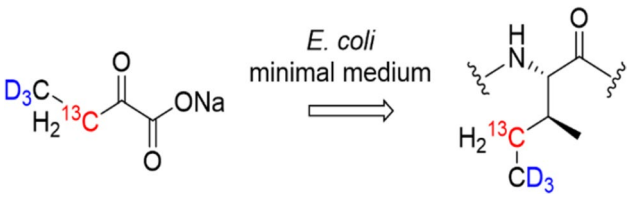
Published online: 11 October 2023

ligand binding modes and $^{1}H^{N}$–^{15}N chemical shift perturbations (Williamson 2013). These problems can be largely overcome by resorting to ^{1}H–^{13}C side-chain detection exploiting the exquisite sensitivity and spectral dispersion obtained with Methyl-TROSY approaches and multiple quantum coherence experiments, which allow applications to slowly tumbling, high molecular weight proteins (Tugarinov et al. 2003, 2004). Selective labeling techniques for aliphatic and/or aromatic residues have been described and are by now well established (Lichtenecker et al. 2013a, b). Labeling of the aliphatic amino acids Isoleucine, Valine and Leucine was one of the first methods of selective ^{13}C incorporation reported in the literature (Cardillo et al. 1977). These amino acids are highly abundant in the hydrophobic cores of proteins and thus represent valuable targets to introduce ^{13}C- and ^{2}H-nuclei with the intention to reduce signal overlap, increase signal-to-noise ratio and optimize magnetization transfer pathways. Moreover, metabolic α-ketoacid precursors can be added to the minimal growth media of *E.coli*-based overexpression systems and are effectively metabolized into the corresponding target aliphatic amino acids *in-vivo* within defined metabolic pathways (Gardner and Kay 1997; Goto et al. 1999). So far, ^{13}C incorporation in aliphatic residues has mainly focused on methyl groups to benefit from their advantageous relaxation properties due to fast C–C bond rotation and the reduction of signal overlap in the otherwise crowded spectral region of methyl signals. Methyl CH_3 groups and aromatic CHs are sensitive reporters of the ligand binding mode and changes of side-chain structural dynamics upon ligand binding. A special and particularly interesting case of non-covalent interactions are CH-π interactions, in which the CH group of the aliphatic or aromatic residue acts as the hydrogen-bond donor, whereas the π-system is the hydrogen-bond acceptor (Hunter 2004). We have recently shown that CH-π interactions are determinants for the affinity of a drug molecule to its target protein and can efficiently be probed by selective amino acid labeling (Platzer et al. 2020).

Encouraged by our previously obtained results, we suggest extending the selective labeling methodology to the methylene group (CH_2) of Isoleucine.

Herein, we show that γ1-^{13}C Isoleucine patterns are effectively induced by the corresponding precursor [3-^{13}C, 4,4,4-$^{2}H_3$] 2-ketobutyrate (Scheme 1) and can be applied as sensitive tools to probe interactions between aryl-ligand CH-π acceptors and protein Isoleucine residue CH-π donors.

We additionally describe the possibility to use [3-^{13}C, 4,4,4-$^{2}H_3$] 2-ketobutyrate in D_2O *E.coli* overexpression medium containing *U*-^{2}H glucose to increase the deuteration grade in the β and γ2 positions of the Ile-side chains. This additional deuteration attenuates line-broadening due to ^{1}H–^{1}H dipolar interaction and reduces the number of signals resulting from natural abundance methyl correlations.



Scheme 1 Conversion of [3-^{13}C; 4,4,4-$^{2}H_3$] 2-ketobutyrate in *E. coli* minimal media to Isoleucine

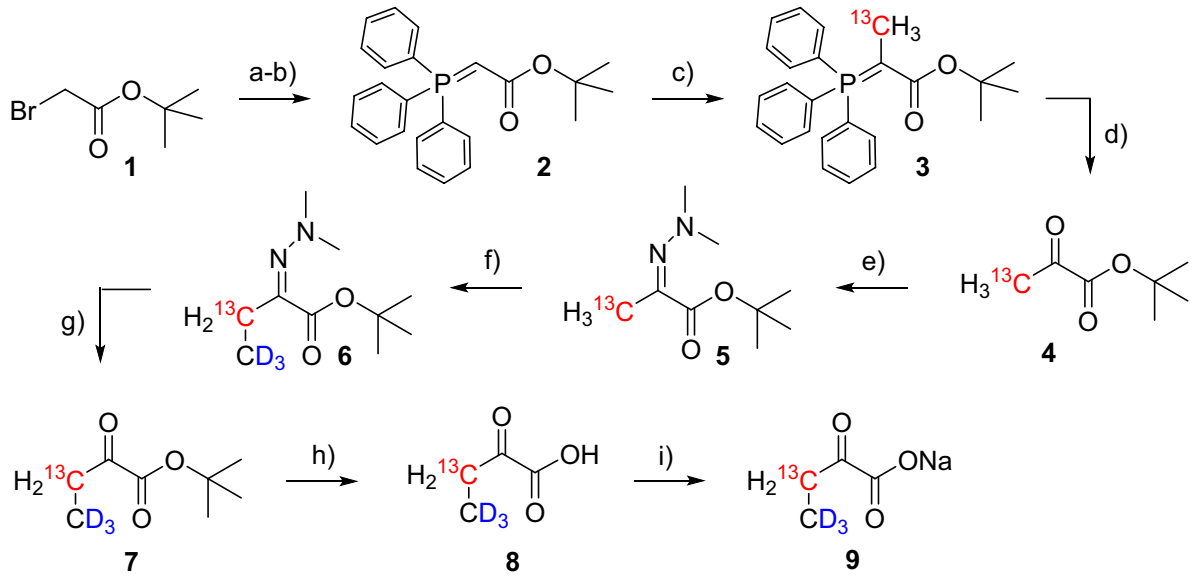
Materials and methods

Precursor synthesis

Isotopologues of α-ketobutyric acid are common additives in minimal media of bacterial protein overexpression to achieve highly selective Isoleucine labeling. While most of these methods aim for ^{13}C methyl groups, we established a novel synthetic route to access an Ile γ_1-$^{13}CH_2$/δ-CD_3 pattern (Scheme 2). Our approach combines a modified literature-known protocol to synthesize 3-^{13}C pyruvate (Werkhoven et al. 1999) with an optimized procedure of dimethylhydrazone monomethylation. Starting from *tert*-butyl bromoacetate **1**, the phosphorous ylide **2** was formed and subsequently methylated using ^{13}C-iodomethane. Ozonolysis with careful bulb-to-bulb distillation of the product from the viscous reaction residue yielded [3-^{13}C] *tert*-butyl pyruvate **4**. This compound was further converted to the dimethylhydrazone **5**, which was then deprotonated at the alpha-position using sodium hexamethyldisilazane (NaHMDS) and subsequently transformed to the hydrazone **6** upon [$^{2}H_3$] iodomethane addition. Short reaction times in this methylation step effectively minimized the formation of side-products (see SI for experimental details). Alternative protocols using lithium diisopropylamide (LDA) as a base reported in literature could not be reproduced in comparable yields (Hajduk et al. 2000). Final hydrazone hydrolysis and ester cleavage under acidic conditions resulted in the formation of the target compound [3-^{13}C; 4,4,4-$^{2}H_3$] α-ketobutyric acid **8**. Compound **8** can be directly applied to protein overexpression as a free acid or lyophilized to a white powder of the corresponding sodium salt **9**, respectively.

Ligand synthesis

Ligands **A** (CAS: 1610045-04-9) and **B** (CAS: 1772600-27-7) were synthesized as reported in earlier work (Platzer et al. 2020). The corresponding structural details of these two compounds are shown in the supporting information.



Scheme 2 Synthesis of [3-^{13}C; 4,4,4-$^{2}H_3$] sodium α-ketobutyrate **9**. Reagents and conditions: **a** PPh_3, toluene, rt, overnight; quant.; **b** $NaOH_{aq}$ (1 M), rt, 15 min, quant.; **c** $^{13}CH_3I$, DCM, rt, overnight; **d** O_3, −78 °C, DCM, 5 min, then PPh_3, 3 h, rt, 70% over two steps; **e** H_2NNMe_2, Et_2O, 18 h, rt, 76%; **f** NaHMDS, CD_3I, THF, −78 °C, 5 min, 80%; **g** HCl_{aq} (1 M), Et_2O, 95%; **h** HCl_g, DCM/Et_2O 1:1, 18 h, 0 °C-rt, 90%; **i** $NaOH_{aq}$ (1 M), lyophilisation, quant. Experimental details, as well as NMR characterization of products and intermediates are given in the Supplementary Information (SI)

Expression and purification of Brd4-BD1

Protein overexpression was performed as described previously (Schörghuber et al. 2017). Briefly, recombinant human Brd4-BD1 (bromodomain 1 of Bromodomain containing protein 4) was expressed in *E.coli* BL21 (DE3), which contains an N-terminal TEV-cleavable His6-tag (plasmid was kindly provided by Boehringer Ingelheim GmbH & Co. KG). Uniformly ^{15}N and ^{13}C-labeled Brd4-BD1 was expressed following the expression protocol for efficient isotopic labeling of recombinant proteins using a fourfold cell concentration in M9 minimal medium, supplemented with 1 g/L $^{15}NH_4Cl$, 3 g/L $^{13}C_6$-D-glucose. Uniformly ^{15}N and selective γ_1-^{13}C Isoleucine labeled Brd4-BD1 was expressed following the same protocol, M9 minimal medium was supplemented with 1 g/L $^{15}NH_4Cl$, 3 g/L $^{12}C_6$-D-glucose and 130 mg/L [3-^{13}C, 4,4,4-$^{2}H_3$] 2-ketobutyrate (Marley et al. 2001). Perdeuterated, uniformly ^{15}N, D-Glucose-1,2,3,4,5,6,6-d_7 and selective γ_1-^{13}C Isoleucine Brd4-BD1 expression (concentrations in [g/l] used as above) was initialized by taking several colonies and inoculating them in a 10 mL M9-H_2O minimal medium for 8 h at 37 °C shaking. From that culture, 250 µL were taken and inoculated in 10 mL fresh M9-D_2O minimal medium, which was shaken overnight at 37 °C. At a cell density of ~2.5, the culture was taken and transferred to the final expression culture in a total 300 mL M9-D_2O medium. The cells were grown until an OD_{600nm} of 0.7 and induced with 0.4 mM IPTG. Cells were harvested by centrifugation, lysed by sonication and the lysates were subsequently centrifuged. Proteins were purified from the supernatant by Ni^{2+} affinity chromatography. The purified protein was treated with TEV protease and again loaded onto a Ni^{2+} column to bind the cleaved His6-tag and the His6-tagged TEV protease. The flow-through containing Brd4-BD1 was concentrated and stored in 10 mM sodium phosphate buffer pH 7.5, 100 mM sodium chloride and 1 mM dithiothreitol (DTT). In the case of perdeuterated Brd4-BD1, after concentrating, the buffer was exchanged to D_2O buffer (10 mM sodium phosphate buffer pH 7.5, 100 mM sodium chloride and 1 mM DTT). Purity was analyzed by SDS-Page. NMR samples of Brd4-BD1 were prepared in 10 mM sodium phosphate buffer containing 0.1–0.5 mM protein, 100 mM sodium chloride, 10% D_2O, and 1 mM DTT. NMR samples of perdeuterated Brd4-BD1 were prepared in 10 mM sodium phosphate D_2O buffer containing 0.2 mM protein, 100 mM sodium chloride, and 1 mM deuterated Tris(2-Carboxyethyl)phosphine:DCl-D16 (TCEP).

NMR measurements

Carbon relaxation studies. All protein NMR measurements were conducted at 298 K on a Bruker Neo 600 MHz spectrometer equipped with a TXI RT probe head with perdeuterated and selective γ_1-^{13}C Isoleucine Brd4-BD1 sample concentration of 200 µM. For ^{13}C relaxation studies, pulse sequence "hsqcctetgpsp" for the constant time ^{1}H–^{13}C HSQC with the constant time delays of 0.0133, 0.0266, 0.04 and 0.0532 s were used (Vuister and Bax 1992). For the

relaxation measurement of ^{13}C–$^{1}H_2$ heteronuclear triple/single quantum coherence, the pulse sequence "hmqcctetgp.2" (slightly modified) for the constant time ^{1}H–^{13}C HTQC was used (Marino et al. 1997). The constant time delays were set to 0.028, 0.042, 0.056, 0.07, 0.084, 0.098, 0.112, 0.128, 0.14, 0.154 s. To allow direct comparison of intensities, all spectra in a series were acquired with the same spectral parameters and the same settings for the receiver gain. Specifically, spectra were recorded using 106 (f1) × 1024 (f2) real points (CT-HSQC) and 144 (f1) × 1024 (f2) real points (CT-HTQC) with acquisition times of 0.06 × 0.01 s. 128 scans per fid were recorded with a recycle delay of 1 s. The pulse scheme of the CH_2-TROSY experiment was applied as described by Miclet et al. (2004) and the corresponding spectrum recorded with a time delay of 0.028 s.

Protein-Ligand interaction studies. Protein-Ligand interaction measurements were conducted at 298 K on a Bruker Avance HD3 + 800 MHz spectrometer equipped with a TXI RT probe head with Brd4-BD1 sample concentration of 200 µM and ligand concentration of 1 mM. 2D ^{1}H–^{13}C HSQC spectra were acquired using the pulse sequence "hsqcetfpgpsi2" of the Bruker library (Palmer et al. 1992; Kay et al. 1992; Grzesiek and Bax 1993; Schleucher et al. 1994). Spectra were recorded using 128 (f1) × 1024 (f2) real points and acquisition times of 0.05 × 0.01 s. 32 scans were recorded per t1 increment with a recycle delay of 1 s.

Analysis

NMR spectra were processed and analyzed with NMRPipe (Delaglio et al. 1995) and SPARKY (Goddard and Kneller 2006) and CCPNmr (Vranken et al. 2005). Unambiguous sequential assignment of the Isoleucine Cγ1 signals was performed with a series of three-dimensional NMR experiments (HNCACB, HN(CO)CACB, hCC-TOCSY-coNH and Hcc-TOCSY-coNH). ^{13}C relaxation studies were performed by fitting the logarithmic intensities of the peaks to the linear regression model in RStudio (RStudio Team 2020) by taking the linearized logarithmic function as $\log[I(t)] = \log(A) - t/T_2$.

Results

To demonstrate the potential of our selectively labeled isoleucine precursor in protein dynamics studies, we incorporated [3-^{13}C; 4,4,4-$^{2}H_3$] α-ketobutyric acid **8** into the bromodomain 1 of Bromodomain-containing protein 4 (Brd4-BD1).

Brd4-BD1 belongs to the family of bromodomain and extra-terminal domain (BET) proteins, which act as chromatin readers by binding to acetylated histones and therefore regulating gene expression (Hu et al. 2022). Brd4-BD1 contains seven Isoleucine residues, theoretically yielding 14 peaks. The incorporation of the ^{13}C label into Brd4-BD1 was confirmed by a ^{1}H–^{13}C-HSQC spectrum (Fig. 1, in red). As expected, two cross peaks are obtained for the diastereotopic methylene protons of each Ile $C\gamma_1$ differing in the hydrogen dimension, but with the same carbon chemical shift. Note that the peaks of two of the Isoleucine $C\gamma_1$ groups are overlapping (presumably I100 and I101). However, spurious cross peaks are found in the methyl group region of the ^{1}H–^{13}C HSQC spectra (Fig. 1). In order to validate that these peaks are indeed natural abundance correlations of methyl CH_3 obtained from the expression with D-Glucose (^{12}C) in H_2O medium (SI Fig. 1, in black), we incorporated the [3-^{13}C; 4,4,4-$^{2}H_3$] α-ketobutyric acid also in D_2O minimal media (SI Fig. 1, in red). As shown in SI Fig. 1 in red, the natural abundance peaks are suppressed by deuteration, and additionally, no metabolic scrambling of the precursor can be observed.

The availability of $^{13}CH_2$ labeled Ile-residues offers the possibility to create heteronuclear multiple-quantum (triple/single-quantum, T(S)QC) coherences in a straightforward manner. It is well-known that triple/single quantum $^{13}CH_2$ coherences relax more slowly than the corresponding single quantum coherences of the individual ^{13}C and ^{1}H spins (Grzesiek and Bax 1995; Grzesiek et al. 1995; Marino et al. 1997; Ruschak et al. 2010; Tugarinov and Kay 2013). Effective relaxation rates were measured by recording a series of ^{1}H–^{13}C-CT-HSQCs and ^{1}H–^{13}C-CT-HTQCs spectra (Fig. 2) with different relaxation delays (see experimental section). Figure 3 shows a comparison of the rates extracted from the ^{1}H–^{13}C-CT-HSQC and ^{1}H–^{13}C-CT-HTQC, respectively. It can clearly be seen that the ^{13}C relaxation times for heteronuclear T/SQC extracted from the ^{1}H–^{13}C-CT-HTQC were

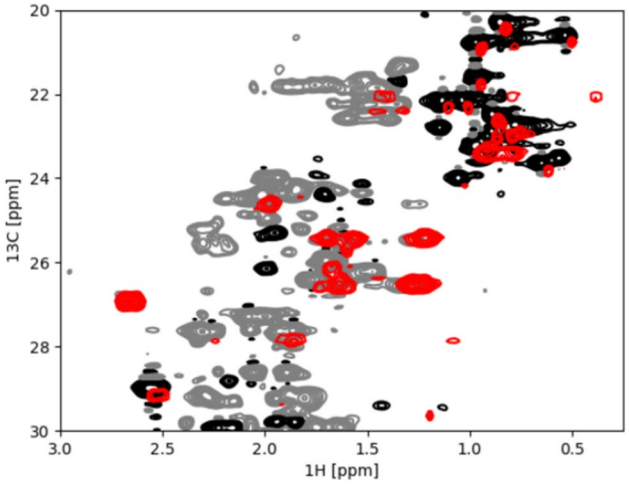


Fig. 1 Overlay of ^{1}H–^{13}C-HSQC spectra of selectively Ile $^{13}CH_2$ labeled Brd4-BD1 (red) onto the CT-^{1}H–^{13}C-HSQC of ^{13}C-uniformly labeled Brd4-BD1 (^{13}CH and $^{13}CH_3$ signals in black and $^{13}CH_2$ signals in grey), zoomed into the CH_2 region

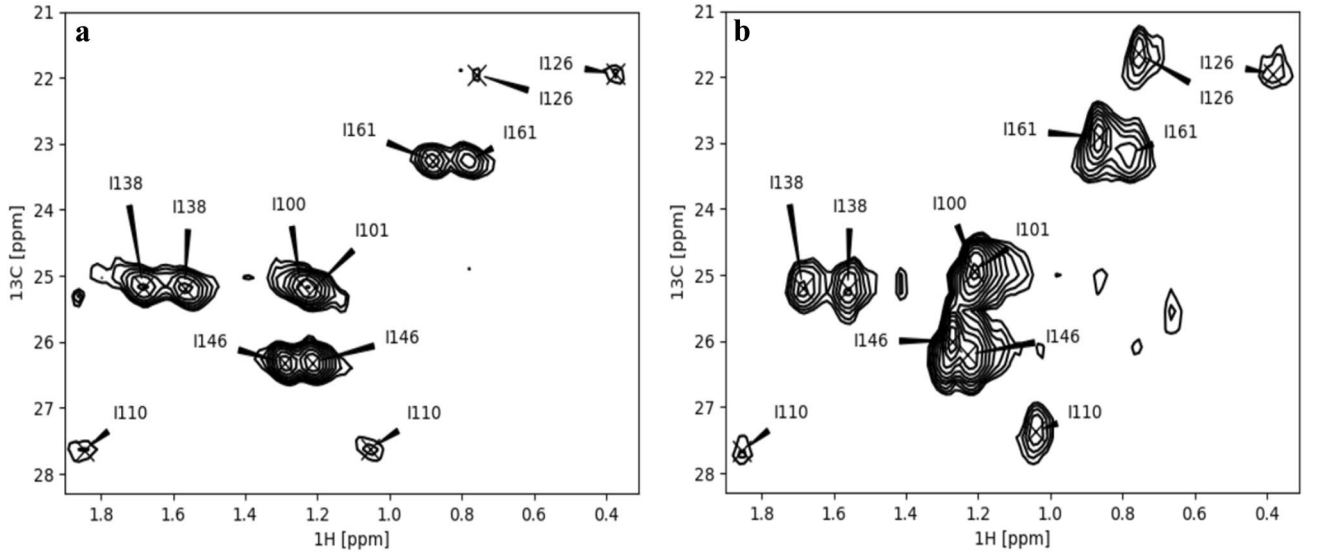


Fig. 2 Comparison between CT-HTQC (ct delay 0.028 s) (**a**) and CT-HSQC (ct delay 0.013 s) (**b**) spectra of perdeuterated ^{15}N-Brd4-BD1-Ile-$^{13}C\gamma 1$

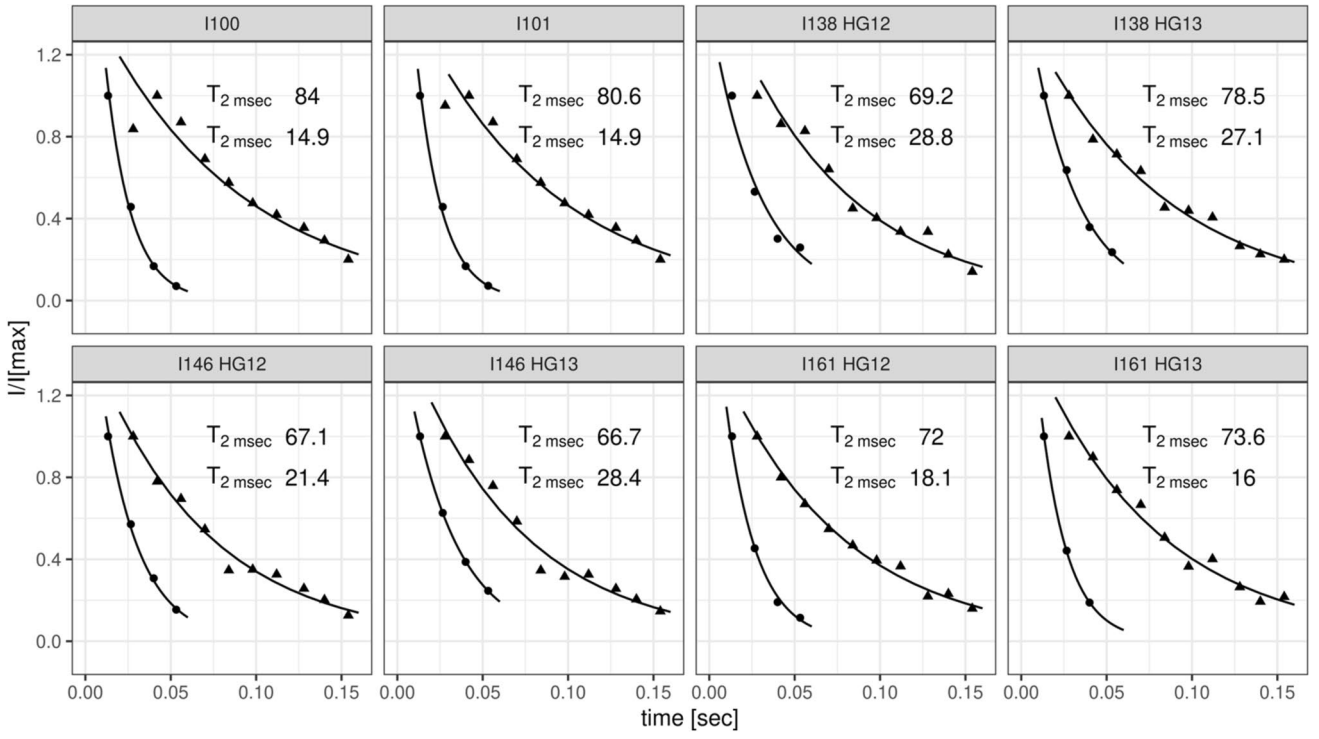


Fig. 3 Extracted ^{13}C relaxation times for specific Ile residues in perdeuterated Brd4-BD1. Dots represent relaxation times extracted from CT-HSQC, whereas pyramids represent rates extracted from CT-HTQCs

increased on average by a factor of three compared to the ^{13}C-T_2 from the ^{1}H–^{13}C-CT-HSQC, as expected from theory (Grzesiek and Bax 1995; Grzesiek et al. 1995; Marino et al. 1997; Ruschak et al. 2010; Tugarinov and Kay 2013). It should be noted that data for residues I110 and I126 could not be analyzed as their signals are barely above the noise level. Interestingly, these two residues are also buried inside the hydrophobic core of Brd4-BD1, whereas the other five

Isoleucine residues are found rather on the surface or at the flexible parts of the protein, thus suggesting the existence of considerable exchange broadening due to conformational averaging.

Selectively Ile-labeled and otherwise highly deuterated samples are moreover particularly interesting for applying specific transverse-relaxation-optimized NMR spectroscopy (TROSY) methods, which have been introduced to improve the sensitivity and resolution of methylene NMR studies (Miclet et al. 2004). We recorded a corresponding CH_2-TROSY spectrum of the Ile $^{13}CH_2$ Brd-BD1 sample expressed in D_2O/deuterated glucose, which shows additional gain in spectral resolution – even relative to the HTQC spectrum (especially in the ^{1}H dimension by virtually eliminating the large geminal $^2J_{HH}$) as a result of the exclusive TROSY selection of the slowest relaxing multiplet component. Upon comparing the signal-to-noise ratios of the different experiments, it becomes evident that HTQC performs better in this regard compared to both CH_2-TROSY and HSQC (SI Fig. 2).

Finally, in order to test whether the methylene group of Isoleucine can be used as a reporter of CH-π interactions, we decided to measure ^{1}H–^{13}C-HSQCs with ligands **A** and **B**, which are nanomolar Brd4-BD1 binders (Platzer et al. 2020). Figure 4a shows the chemical shift changes of ^{15}N-Brd4-BD1-Ile-^{13}Cγ1 with 1 mM ligand **A** (red) and **B** (blue). The ligand induced proton chemical shift perturbation (CSP) of the Ile146 methylene resonances result in 1.71 ppm for ligand **A** and 1.78 ppm for ligand **B** for one hydrogen atom and 0.59 ppm and 0.66 ppm for the other hydrogen, respectively. These significant CSPs point towards an ideally oriented CH-π interaction between the Ile146-$C_{\gamma1}H_2$ and the ligand aromatic system, as favorable CH-π stacking orientations are correlated with high upfield shifts. Taking a closer look at the crystal structure of Brd4-BD1 in complex with both ligands (Fig. 4b), the hydrogen of the methylene group (pro-S) is directly positioned over the centroid of the triazole ring, whereas the pro-R hydrogen atom is further away from the ligand. This information enables us therefore to also stereo-specifically assign the hydrogens of the isoleucine methylene groups. In order to correlate the observed CSPs of the Ile146 methylene resonance with the relative orientation of the ligand in the binding pocket of Brd4-BD1, we used the Pople equation (Pople 1956; Platzer et al. 2020). The geometric parameters were taken from the published X-ray structures (see SI Table 2), resulting in a calculated CSP for the pro-S hydrogen of Ile146 of 1.1 ppm and 1.5 ppm for ligand **A** and **B**. For the hydrogen atom in pro-R configuration, the calculated CSPs are 0.5 ppm and 0.6 ppm, respectively. This finding shows that the experimentally derived CSPs match well with the calculated, again supporting the presence of a beneficial CH-π stacking orientation.

Discussion and conclusion

The novel α-ketobutyric isotopologue reported here provides an economic tool to implement methylene labeling in isoleucine side chains. This precursor is directly applicable to *E.coli*-based expression systems, as shown by incorporation into the human protein Brd4-BD1, allowing observation of seven distinct Isoleucine Cγ1–H_2 pairs in an otherwise

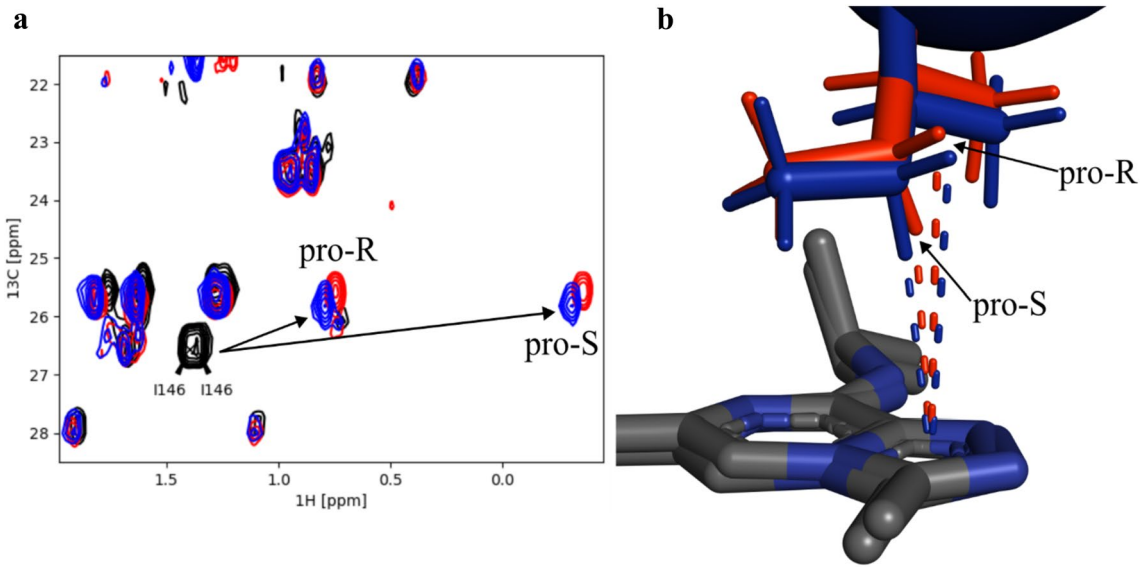


Fig. 4 **a** ^{1}H-^{13}C-HSQC overlay of Brd4-BD1-Ile-^{13}Cγ1 (black) with Ligand **A** (red) and Ligand **B** (blue). **b** Zoom into the Brd4-BD1 binding pocket with an overlay of Ligand **A** (red, 6XV3) and **B** (blue, 6XUZ). Dashed lines indicate the distances between the ligand ring center and the two diastereotopic protons of Ile146. The prochiral designation of the methylene group is indicated in both figures

crowded ^{1}H–^{13}C-HSQC spectra (Fig. 1). The incorporation of the precursor into Brd4-BD1 by expression in D_2O minimal media results in simplified protein spectra against a deuterated background with favourable ^{13}C relaxation properties and enhanced signal-to-noise ratio. Our experiments further exemplify the benefit of the deuterated Isoleucine Cγ1 labeled Brd4-BD1 by studying its carbon transverse relaxation properties by measuring ^{13}C single and ^{13}C/^{1}H heteronuclear multiple quantum coherences. Comparison of the extracted effective ^{13}C relaxation times shows an increase by a factor of ~3.5 going from ^{13}C SQC to heteronuclear triple quantum coherences (TQC). The larger ^{13}C relaxation times in the TQC experiment might allow for interesting applications such as paramagnetic relaxation enhancement (PRE), residual dipolar couplings (RDC) and conformational exchange via CPMG-type measurements. In practice, this has to be balanced against other experimental factors such as potentially higher overall intrinsic signal intensity in the HSQC spectra (due to signal improvement by Rance/Kay type gradient schemes) but holds the promise of application to large molecular weight systems. Most importantly, however, the large upfield chemical shift observed for the pro-S methylene proton in the BRD4-ligand complex convincingly demonstrates the usefulness of the Isoleucine methylene group as an excellent reporter for CH-π interactions.

It can thus be anticipated that particularly protein-ligand binding studies will considerably benefit from the availability of the new Isoleucine precursor. The facile and efficient introduction of this novel precursor to realize new Isoleucine isotopologues in a target protein represents a valuable and general tool to fine-tune NMR studies and decipher protein dynamics, allosteric mechanisms, and binding interactions.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10858-023-00427-2.

Acknowledgements T. Höfurthner, G. Toscano and A. Beier were funded by the Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Austria. The authors would like to thank G. Platzer, S. Brüschweiler, I. Ceccolini and T. Schwarz for input and discussion.

Author contributions All authors contributed to the study conception, design and analysis. Material preparation and data collection were performed by TH and GT. All authors read and approved the final manuscript.

Funding Open access funding provided by University of Vienna.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

References

Cardillo R, Fuganti C, Ghiringhelli D, Grasselli P, Gatti G (1977) Pattern of incorporation of leucine samples asymmetrically labelled with 13 C in the isopropyl unit into the C5-isoprenoid units of echinuline and flavoglaucine. J Chem Soc Chem Commun. https://doi.org/10.1039/c39770000474

Dalvit C, Pevarello P, Tatò M, Veronesi M, Vulpetti A, Sundström M (2000) Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. J Biomol NMR 18:65–68. https://doi.org/10.1023/a:1008354229396

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293. https://doi.org/10.1007/BF00197809

Gardner KH, Kay LE (1997) Production and incorporation of ^{15} N, ^{13} C, ^{2}H (^{1}H-δ1 methyl) isoleucine into proteins for multidimensional NMR studies. J Am Chem Soc 119(32):7599–7600. https://doi.org/10.1021/ja9706514

Goddard TD, Kneller DG (2006) Sparky—NMR assignment and integration software. University of California, California

Gossert AD, Jahnke W (2016) NMR in drug discovery: a practical guide to identification and validation of ligands interacting with biological macromolecules. Prog Nucl Magn Reson Spectrosc 97:82–125. https://doi.org/10.1016/j.pnmrs.2016.09.001

Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE (1999) A robust and cost-effective method for the production of val, Leu, Ile (delta 1) methyl-protonated 15 N-, 13 C-, 2H-labeled proteins. J Biomol NMR 13(4):369–374. https://doi.org/10.1023/a:1008393201236

Gronenborn AM (2022) Small, but powerful and attractive: 19F in biomolecular NMR. Structure 30:6–14. https://doi.org/10.1016/j.str.2021.09.009

Grzesiek S, Bax A (1993) The importance of not saturating water in protein NMR: application to sensitivity enhancement and NOE measurements. J Am Chem Soc 115:12593–12594. https://doi.org/10.1021/ja00079a052

Grzesiek S, Bax A (1995) Spin-locked multiple quantum coherence for signal enhancement in heteronuclear multidimensional NMR experiments. J Biomol NMR 6:335–339. https://doi.org/10.1007/BF00197815

Grzesiek S, Kuboniwa H, Bax A, Hinck AP (1995) Multiple-quantum line narrowing for measurement of Hα—HβJ couplings in isotopically enriched proteins. J Am Chem Soc 117:5312–5315. https://doi.org/10.1021/ja00124a014

Hajduk PJ, Augeri DJ, Mack J, Mendoza R, Yang J, Betz SF, Fesik SW (2000) NMR-based screening of proteins containing 13 C-Labeled methyl groups. J Am Chem Soc 122:7898–7904. https://doi.org/10.1021/ja000350l

Harner MJ, Mueller L, Robbins KJ, Reily MD (2017) NMR in drug design. Arch Biochem Biophys 628:132–147. https://doi.org/10.1016/j.abb.2017.06.005

Hu J, Pan D, Li G, Chen K, Hu X (2022) Regulation of programmed cell death by Brd4. Cell Death Dis 13:1059. https://doi.org/10.1038/s41419-022-05505-1

Hunter CA (2004) Quantifying intermolecular interactions: guidelines for the molecular recognition toolbox. Angew Chem Int Ed 43:5310–5324. https://doi.org/10.1002/anie.200301739

Kay LE, Keifer P, Saarinen T (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J Am Chem Soc 114:10663–10665. https://doi.org/10.1021/ja00052a088

Lichtenecker RJ, Coudevylle N, Konrat R, Schmid W (2013) Selective isotope labelling of leucine residues by using α-Ketoacid precursor compounds. ChemBioChem 14:818–821. https://doi.org/10.1002/cbic.201200737

Lichtenecker RJ, Weinhäupl K, Reuther L, Schörghuber J, Schmid W, Konrat R (2013) Independent valine and leucine isotope labeling in *Escherichia coli* protein overexpression systems. J Biomol NMR 57:205–209. https://doi.org/10.1007/s10858-013-9786-y

Luchinat E, Barbieri L, Cremonini M, Nocentini A, Supuran CT, Banci L (2020) Drug screening in human cells by NMR spectroscopy allows the early assessment of drug potency. Angew Chem Int Ed 59:6535–6539. https://doi.org/10.1002/anie.201913436

Marino JP, Diener JL, Moore PB, Griesinger C (1997) Multiple-quantum coherence dramatically enhances the sensitivity of CH and CH2 correlations in uniformly 13 C-labeled RNA. J Am Chem Soc 119:7361–7366. https://doi.org/10.1021/ja964379u

Marley J, Lu M, Bracken C (2001) A method for efficient isotopic labeling of recombinant proteins. J Biomol NMR 20:71–75. https://doi.org/10.1023/A:1011254402785

Mayer M, Meyer B (2001) Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. J Am Chem Soc 123:6108–6117. https://doi.org/10.1021/ja0100120

Miclet E, Williams DC Jr, Clore GM, Bryce DL, Boisbouvier J, Bax A (2004) Relaxation-optimized NMR spectroscopy of methylene groups in proteins and nucleic acids. J Am Chem Soc 126(34):10560–10570. https://doi.org/10.1021/ja047904v

Palmer AG, Cavanagh J, Byrd RA, Rance M (1992) Sensitivity improvement in three-dimensional heteronuclear correlation NMR spectroscopy. J Magn Reson 96:416–424. https://doi.org/10.1016/0022-2364(92)90097-Q

Platzer G, Mayer M, Beier A, Brüschweiler S, Fuchs JE, Engelhardt H, Geist L, Bader G, Schörghuber J, Lichtenecker R, Wolkerstorfer B, Kessler D, McConnell DB, Konrat R (2020) PI by NMR: probing CH–π interactions in protein–ligand complexes by NMR spectroscopy. Angew Chem Int Ed 59:14861–14868. https://doi.org/10.1002/anie.202003732

Pople JA (1956) Proton magnetic resonance of hydrocarbons. J Chem Phys 24:1111–1111

Rees DC, Congreve M, Murray CW, Carr R (2004) Fragment-based lead discovery. Nat Rev Drug Discov 3:660–672. https://doi.org/10.1038/nrd1467

RStudio T (2020) RStudio: integrated development for R. Rstudio Team, Boston

Ruschak AM, Velyvis A, Kay LE (2010) A simple strategy for 13 C,1H labeling at the Ile-γ2 methyl position in highly deuterated proteins. J Biomol NMR 48:129–135. https://doi.org/10.1007/s10858-010-9449-1

Schleucher J, Schwendinger M, Sattler M, Schmidt P, Schedletzky O, Glaser SJ, Sörensen OW, Griesinger C (1994) A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. J Biomol NMR 4:301–306. https://doi.org/10.1007/BF00175254

Schörghuber J, Geist L, Bisaccia M, Weber F, Konrat R, Lichtenecker RJ (2017) Anthranilic acid, the new player in the ensemble of aromatic residue labeling precursor compounds. J Biomol NMR 69:13–22. https://doi.org/10.1007/s10858-017-0129-2

Shuker SB, Hajduk PJ, Meadows RP, Fesik SW (1996) Discovering high-affinity ligands for proteins: SAR by NMR. Science (80-) 274:1531–1534. https://doi.org/10.1126/science.274.5292.1531

Tugarinov V, Kay LE (2013) Estimating side-chain order in [U-2H;13CH 3]-labeled high molecular weight proteins from analysis of HMQC/HSQC spectra. J Phys Chem B 117:3571–3577. https://doi.org/10.1021/jp401088c

Tugarinov V, Hwang PM, Ollerenshaw JE, Kay LE (2003) Cross-correlated relaxation enhanced ^{1}H-^{13} C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. J Am Chem Soc 125:10420–10428. https://doi.org/10.1021/ja030153x

Tugarinov V, Sprangers R, Kay LE (2004) Line narrowing in methyl-TROSY using zero-quantum 1H-13 C NMR spectroscopy. J Am Chem Soc 126:4921–4925. https://doi.org/10.1021/ja039732s

Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins Struct Funct Bioinform 59:687–696. https://doi.org/10.1002/prot.20449

Vuister GW, Bax A (1992) Resolution enhancement and spectral editing of uniformly 13 C-enriched proteins by homonuclear broadband 13 C decoupling. J Magn Reson 98:428–435. https://doi.org/10.1016/0022-2364(92)90144-V

Werkhoven TM, van Nispen R, Lugtenburg J (1999) Spcific isotope enrichment of methyl methacrylate. Eur J Org Chem 11:2909–2914

Williamson MP (2013) Using chemical shift perturbation to characterise ligand binding. Prog Nucl Magn Reson Spectrosc 73:1–16. https://doi.org/10.1016/j.pnmrs.2013.02.001

Synthesis of a ^{13}C-Methylene-Labeled Isoleucine Precursor as a Useful Tool for Studying Protein Side-chain Interactions and Dynamics

Theresa Höfurthner[1,3], Giorgia Toscano[2,3], Georg Kontaxis[4], Andreas Beier[1], Moriz Mayer[5], Leonhard Geist[5], Darryl B. McConnell[5], Harald Weinstabl[5], Robert Konrat[1,4*], Roman Lichtenecker[2*]

[1] *Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Department of Structural and Computational Biology, Max Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, 1030 Vienna (Austria)*

[2] *Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Institute of Organic Chemistry, University of Vienna, Währingerstraße 38, 1090 Vienna (Austria)*

[3] *Vienna Doctoral School in Chemistry (DoSChem), University of Vienna, Währingerstraße 42, 1090 Vienna (Austria)*

[4] *Max Perutz Laboratories, Department of Structural and Computational Biology, Campus Vienna Biocenter 5, 1030 Vienna*

[5] *Boehringer Ingelheim RCV GmbH & Co. KG, Dr. Boehringer Gasse 5-11, 1121 Vienna*

* *corresponding authors Dr.Roman Lichtenecker, roman.lichtenecker@univie.ac.at; Univ.Prof. Dr. Robert Konrat, robert.konrat@univie.ac.at*

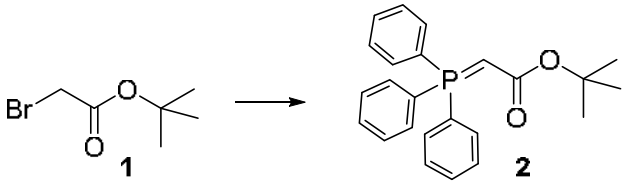
General Information – Organic Synthesis

Unless otherwise stated, all reagents and reactants were purchased from commercial suppliers and used without further purification. All solvents were distilled before use. Iodomethane-^{13}C and Iodomethane-d_3 were purchased by Merck ISOTEC®. Oxygen- and moisture sensitive reactions were carried out under an argon atmosphere and yields refer to pure compounds. The reactions were monitored via thin layer chromatography (TLC) on silica gel 60 with fluorescent indicator UV254 by MACHEREY-NAGEL GmbH & Co. KG. Visualization of the compounds was carried out using an UV-lamp (254 nm) and by application of specific reagents: $H_3PMo_{12}O_{40}$ 10 % in ethanol with subsequent heating using a heat-gun. Flash column chromatography was performed on silica gel 60 (0.040-0.063 mm) from Merck. Freeze-drying was performed by cooling in liquid nitrogen and subsequent application of high vacuum (oil pump). ^{1}H and ^{13}C 1D NMR spectroscopic data were recorded on a Bruker AVANCE-DRX 400 MHz spectrometer. NMR solvent signals were calibrated to 7.27 ppm ($CDCl_3$), 4.79 ppm (D_2O). Chemical shifts (δ) are given in ppm (s = singlet, d = doublet, dd = doublet of doublets, m = multiplet) and reported relative to the residual solvent peaks. Coupling constants (J) are given in Hertz (Hz). High resolution mass spectrometry experiments were performed using electrospray ionization (ESI, 3 keV, in the positive or negative ion mode) or electron ionization (EI, 70 eV).

Abbreviations: PPh_3 …Triphenylphosphine, EtOAc …Ethyl acetate, Et_2O …Diethyl ether, DCM …Dichloromethane, RT …room temperature, THF …Tetrahydrofuran, NaHMDS …Sodium bis(trimethylsilyl)amide, sat …saturated

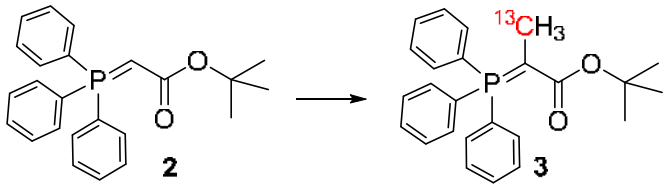
Synthetic Procedures

Synthesis of Compound 2



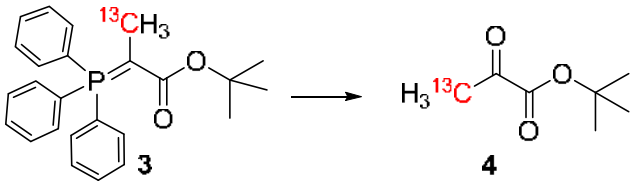
The procedure was adapted from literature (Werkhoven et al. 1999). A solution of 12,3 g (1,05 eq) of PPh_3 in 25 ml of EtOAc, 8,23 g (44,6 mmol) of tert-butyl-2- bromoacetate 1 were added and the solution was stirred at room temperature overnight. The next day, the white precipitate was filtered off and washed three times with Et_2O. After drying, the precipitate was dissolved in DCM and 100 ml of 1 M (2,2 eq) NaOH solution was added. The mixture was stirred vigorously for 15 min, and the two phases separated. The aqueous phase was extracted twice with DCM. All organic phases were combined and dried over $MgSO_4$. Evaporating the solvents in vacuum yielded a yellow solid, which was taken up in a minimum amount of DCM. Precipitation of the product was promoted by adding a layer of heptane and stirring for one hour at 0 °C. The precipitate was filtered and washed with heptane to yield *tert*-butyl 2-(triphenylphosphoranylidene)acetate **2** as a white powder. Yield: 16.78 g (quantitative) ^{1}H NMR ($CDCl_3$, 400 MHz) δ 7,95 (3H, m), 7.77 – 7.49 (12H, m), 5.67 (1H, m), 1.21 & 0.98 (9H, two broad s).

Synthesis of Compound 3



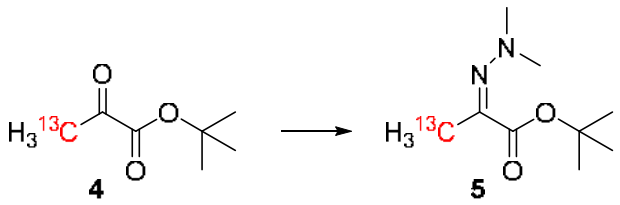
The experimental procedure was conducted as reported in the literature (Werkhoven et al. 1999). A solution of 2,9 g (7,7 mmol) (*tert*-butoxycarbonylmethylene)triphenylphosphorane **2** in 15 ml DCM, 1 g (0,9 eq) $^{13}CH_3I$ was added. After stirring the solution overnight, the solvent was removed under reduced pressure. The yellow residue was taken up in 30 ml of DCM and shaken vigorously with a solution of 560 mg (2 eq) NaOH in 10 ml of H_2O in a separation funnel. The aqueous phase was extracted 2x with DCM. The combined organic phases were dried over $MgSO_4$ to afford a clear, yellow product solution, which was used for the next step without any further purification.

Synthesis of Compound 4



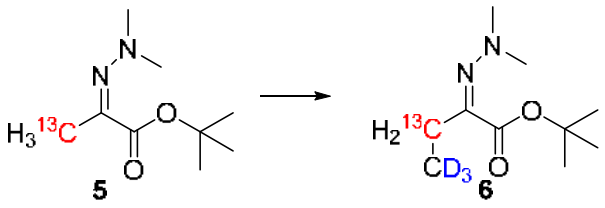
A constant stream of O_3 from an ozone generator was purged through the solution of the ylid **3** in DCM (resulting from the preceding reaction) at -78 °C until the solution turned blue. Subsequently, a stream of argon was then purged through the solution until the blue colour vanished. 2,22 mg (1,2 eq) of PPh_3 were added and the solution was stirred under inert conditions for another 15 min at -78 °C. The mixture was then brought to RT and stirring was continued for two hours. The bulk solvent was carefully removed under reduced pressure. [3-^{13}C] *Tert*-butyl pyruvate **4** was obtained by bulb-to-bulb distillation (100 mbar, up to 90 °C) as a colorless liquid. Yield: 782.3 mg (70 % over 2 steps) ^{1}H NMR ($CDCl_3$, 400 MHz) δ 2.42 (3H, d, $1J_{CH}$ = 129.2 Hz), 1.56 (9H, s). ^{13}C ($CDCl_3$, 100 MHz) δ 84.13, 26.36, 27.93.

Synthesis of Compound 5



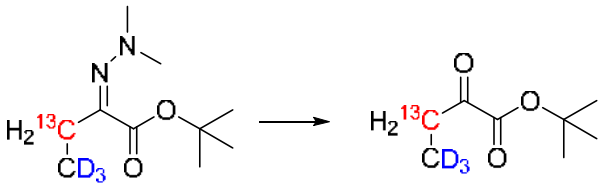
Compound 5 was prepared similar as reported in (Hajduk et al. 2000). To a solution of 311 mg (2,14 mmol) [3-^{13}C] *tert*-butyl pyruvate **4** in 5 ml Et_2O, 200 µl (1,2 eq) N,N-dimethylhydrazine were added under argon atmosphere. The mixture was stirred at RT overnight. The next day, the solution was diluted with Et_2O and then washed 2x with small amounts of H_2O. The aqueous phase was extracted 3x with small amounts of Et_2O. The organic phases were pooled, dried over $MgSO_4$, and concentrated under reduced pressure. [3-^{13}C] *Tert*-butyl 2-(2,2-dimethylhydrazono)propanoate **5** was purified by bulb-to-bulb distillation (10 mbar, up to 110 °C) and obtained as a yellow liquid. Yield: 303 mg (76 %) of two isomers in an approx. ratio of 5:1. ^{1}H NMR ($CDCl_3$, 400 MHz) Isomer 1: δ 2.79 (6H, s), 2.06 (3H, d, $1J_{CH}$ = 129.7 Hz), 1.53 (9H, s); Isomer 2: 2.51 (1.2H, s), 2.04 (0.6H, d, $1J_{CH}$ = 129.7 Hz), 1.53 (1.8H, s). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 15.95, 19.48, 26.61, 28.09, 47.06, 47.52, 81.64, 82.61. HRMS δ (M + Na)+ for $C_8{}^{13}CH_{18}N_2O_2$ calculated 210.1294; found 210.1298.

Synthesis of Compound 6



A solution of 303 mg (1,62 mmol) of hydrazone **5** in 15 ml anhydrous THF was set under argon atmosphere and cooled to -78 °C. 3,25 ml (2 eq) of 1 M NaHMDS in THF/hexanes were added over the course of 30 seconds. Stirring was continued for 2 min and 102 µl (1 eq) CD_3I were quickly added. After 5 min of additional stirring, the reaction was quenched with 5 ml of H_2O and brought to RT. After phase separation, the aqueous phase was extracted 4x with Et_2O. The combined organic phases were washed with H_2O and dried over $MgSO_4$. The solvent and silyl side products were removed under reduced pressure (50 mbar, 40 °C) for several hours until mass consistency was reached. Bulb-to-bulb distillation (10 mbar, up to 110 °C) afforded [3-^{13}C; 4,4,4-$^{2}H_3$] *Tert*-butyl 2-(2,2-dimethylhydrazono)butanoate **6** as a clear, yellow liquid. Yield: 264.65 mg of two isomers in a ratio of approx. 4:1 (80 %) ^{1}H NMR ($CDCl_3$, 400 MHz) Isomer 1: δ 2.77 (6H, s), 2.53 (2H, d, $1J_{CH}$ = 129.7 Hz), 1.52 (9H, s); Isomer 2: 2.50 (1.2H, s), 2.35 (0.5H, d, $1J_{CH}$ = 129.7 Hz), 1.52 (2.2H, s); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 82.59, 81.71, 47.89, 32.40, 28.45, 26.97, 22.12, 19.93, 18.52, 16.13, 10.28; HRMS (M + Na)+ for $C_9{}^{13}CH_{17}D_3N_2O_2$+Na calculated 227.1633; found 227.1638.

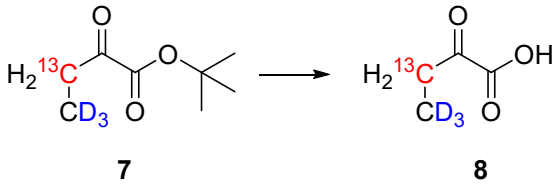
Synthesis of Compound 7



6 **7**

The synthesis of α-ketoester 7 was performed as described in the literature (Lichtenecker et al. 2004). A solution of 190 mg (0,93 mmol) of hydrazone **6** in 25 ml Et_2O was shaken vigorously with 1,12 ml of 1 M HCl in a separation funnel for 1 minute. The phases were separated, and the aqueous phase was extracted 2x with Et_2O. The combined organic phases were washed 1x with sat. $NaHCO_3$ solution and 1x with brine. The solution was then dried over $MgSO_4$ and bulk solvent was carefully removed under reduced pressure. The product was further purified by column chromatography (silica gel, 10 % diethyl ether in pentane), which afforded 230 mg of a clear, colourless solution of [3-^{13}C; 4,4,4-$^{2}H_3$] *tert*-butyl-2-ketobutanoate **7**. Yield: 143 mg (95 %) calculated from ^{1}H-NMR spectrum. ^{1}H NMR ($CDCl_3$, 400 MHz) δ 2.78 (2H, d, $1J_{CH}$ = 127.0 Hz), 1.54 (9H, s) HRMS (M + Na)+ for $C_7{}^{13}CH_{11}D_3O_3$ calculated 185.1062; found 185.1054.

Synthesis of Compound 8

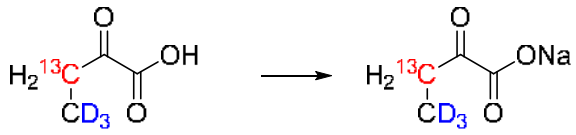


7 **8**

The α-ketoacid 8 was prepared as previously described (Lichtenecker et al. 2004). A solution of 143 mg (0,70 mmol) **7** in 20 ml of a 50:50 mixture of Et_2O and DCM was cooled to 0 °C. HCl gas (generated from

slowly dropping concentrated HCl into concentrated sulphuric acid) was bubbled through this solution for 15 minutes at a moderate rate. Then, the reaction vessel was sealed and stirred at RT for 1 h. The solution was then cooled again to 0 °C, and the process was repeated 3 times, while stirring was continued overnight after the last cycle. The solvent was removed under reduced pressure to afford [3-^{13}C; 4,4,4-$^{2}H_3$] 2-ketobutyric acid as a clear, slightly yellowish viscous liquid. Yield: 66.8 mg (90 %) ^{1}H NMR ($CDCl_3$, 400 MHz) δ 3.00 (2H, d, $1J_{CH}$ = 127.6 Hz). HRMS (M-H)- for $C_3{}^{13}CH_3D_3O_3$ calculated 105.0466, found 105.0459.

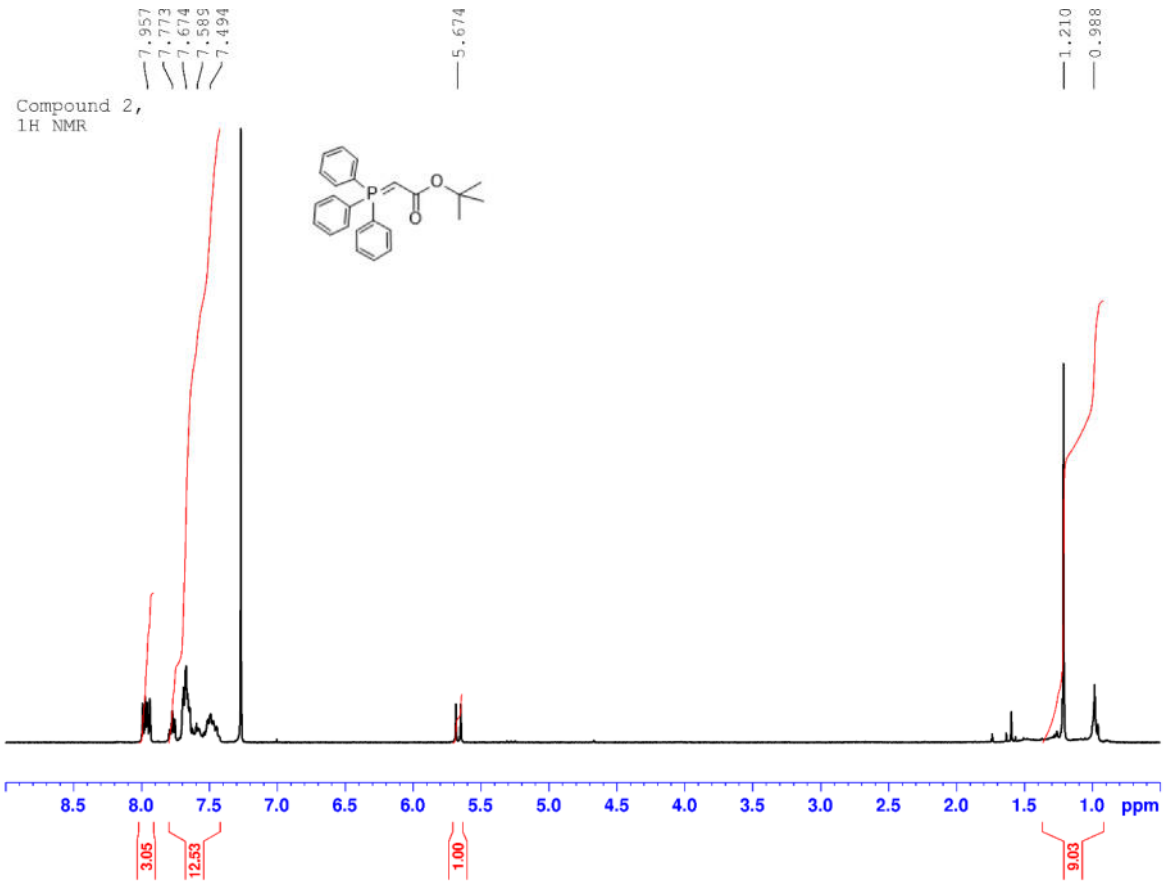
Synthesis of Compound 9

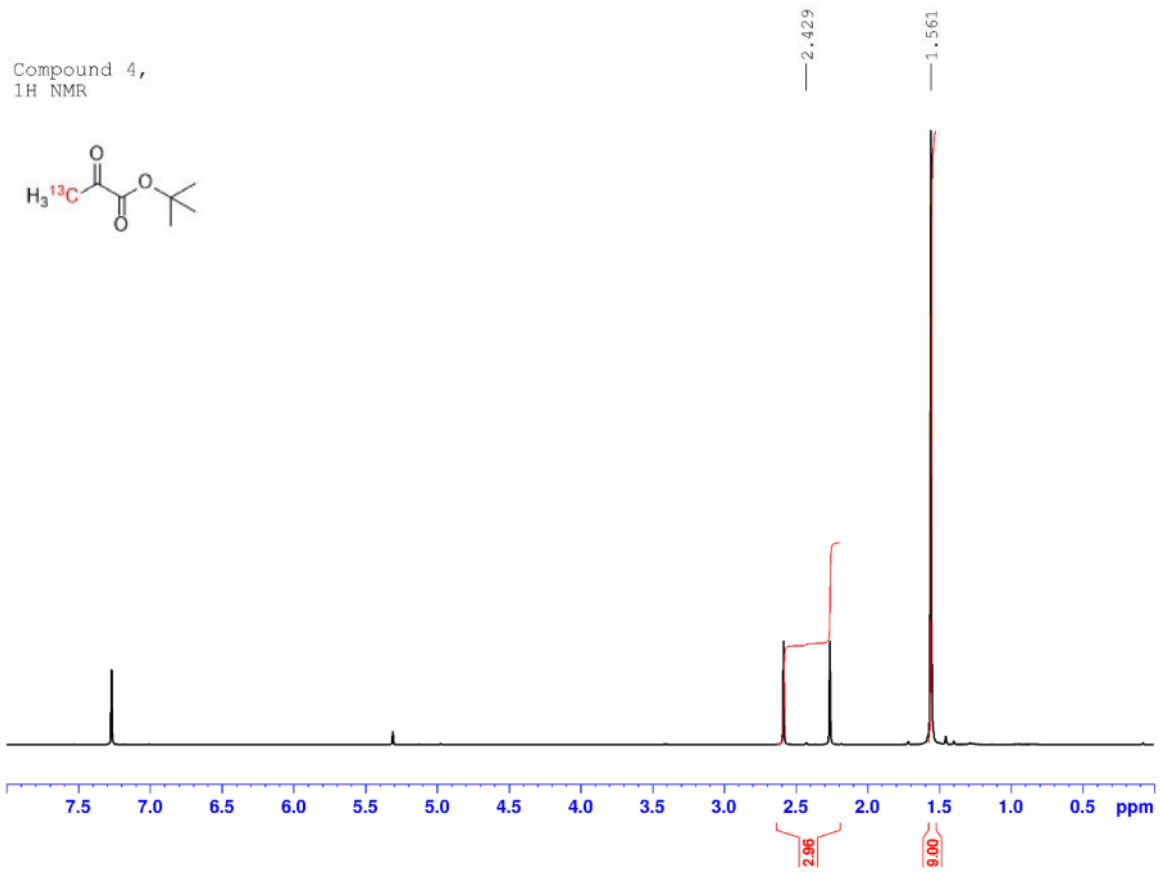


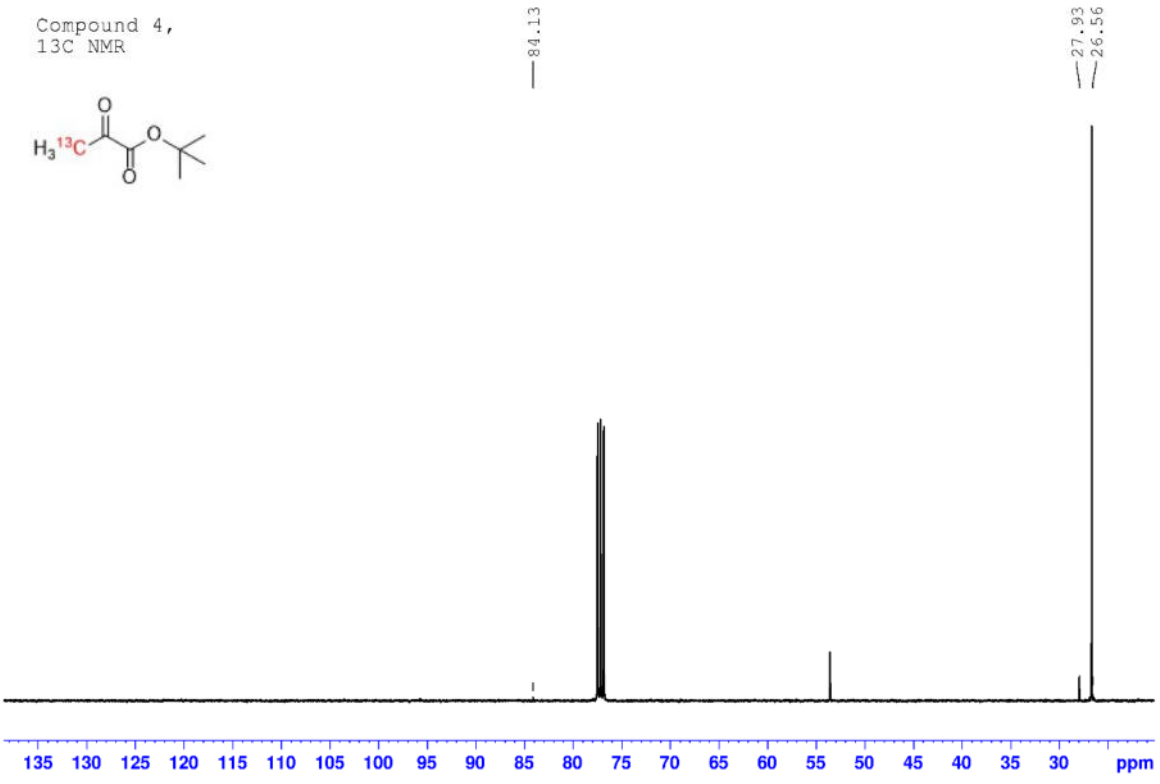
8 **9**

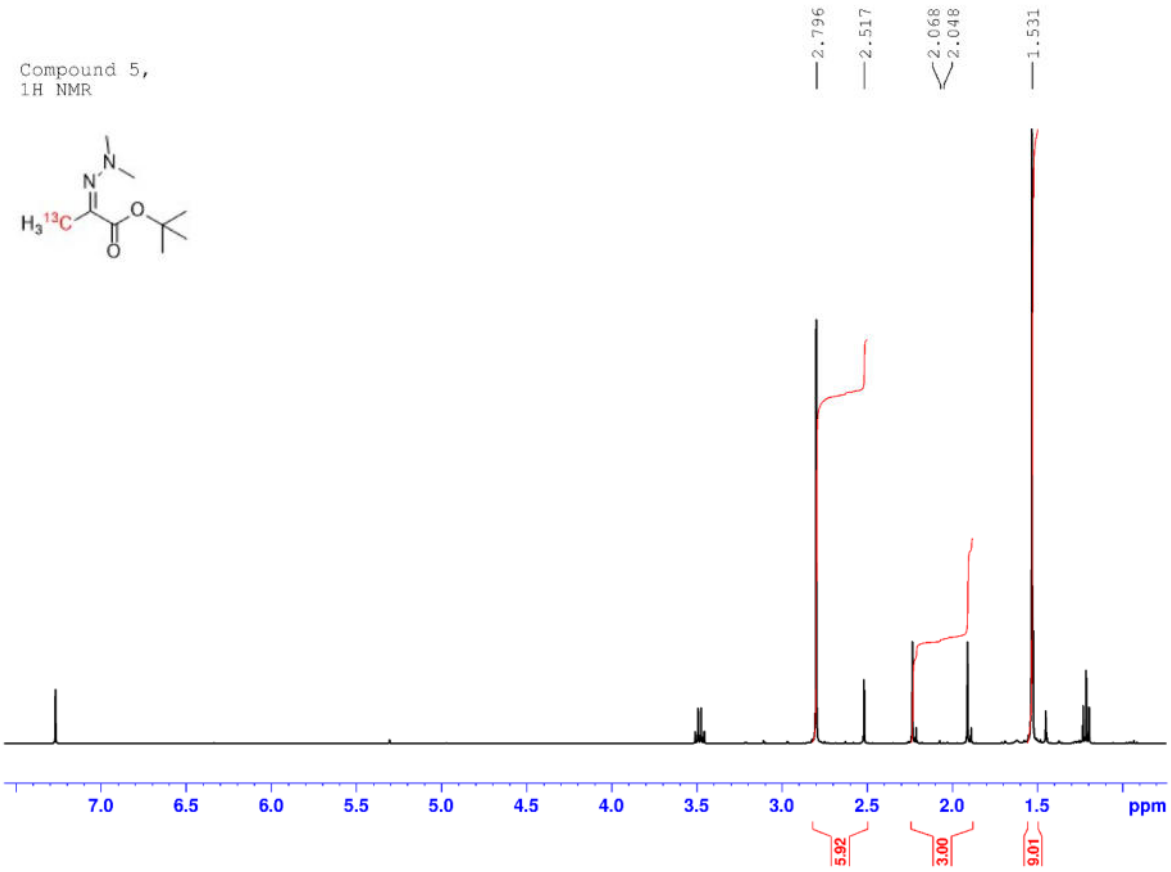
The α-keto-acid **8** was suspended in 5 ml of H_2O and neutralized with 0.1 M NaOH carefully to a pH of 7. The sodium salt was lyophilised to afford [3-^{13}C; 4,4,4-$^{2}H_3$] sodium 2-ketobutyrate **9** as a white, powdered solid. Yield: 72.5 mg (90 %). ^{1}H NMR (D_2O, 400 MHz) δ 2.73 (2H, d, $1J_{CH}$ = 127.1 Hz).

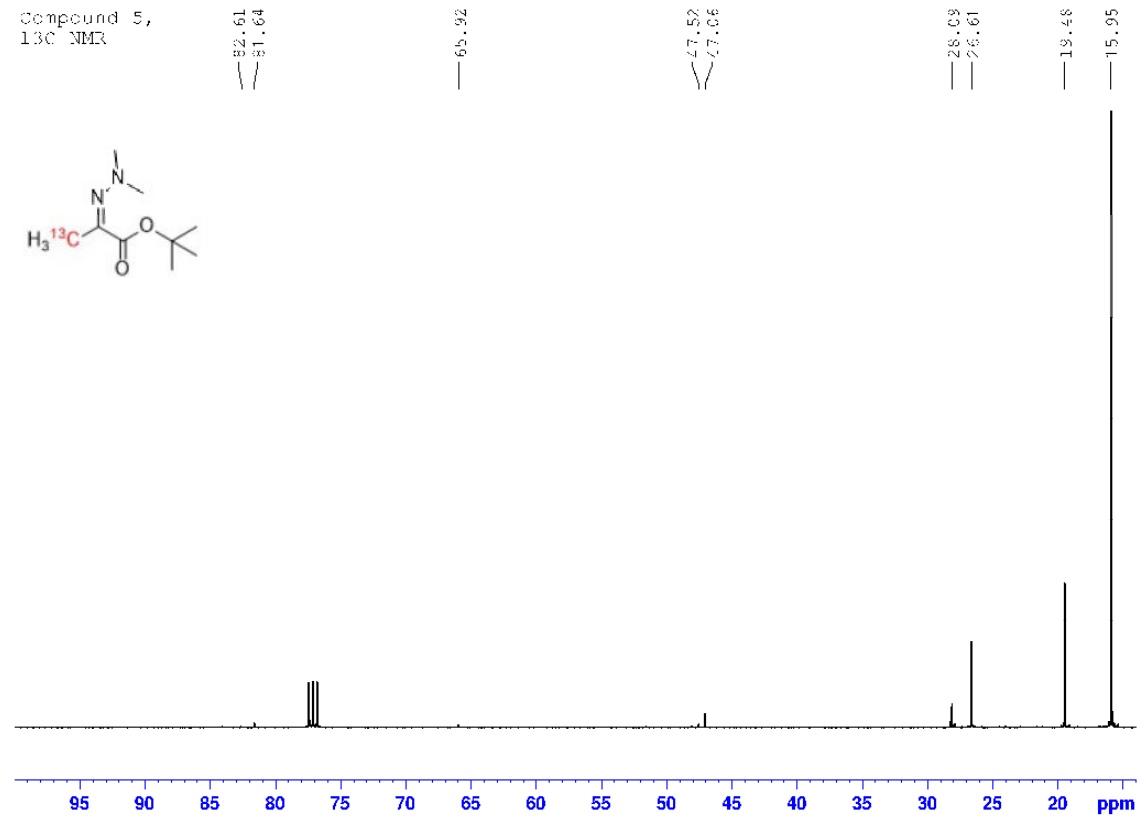
NMR Spectra:

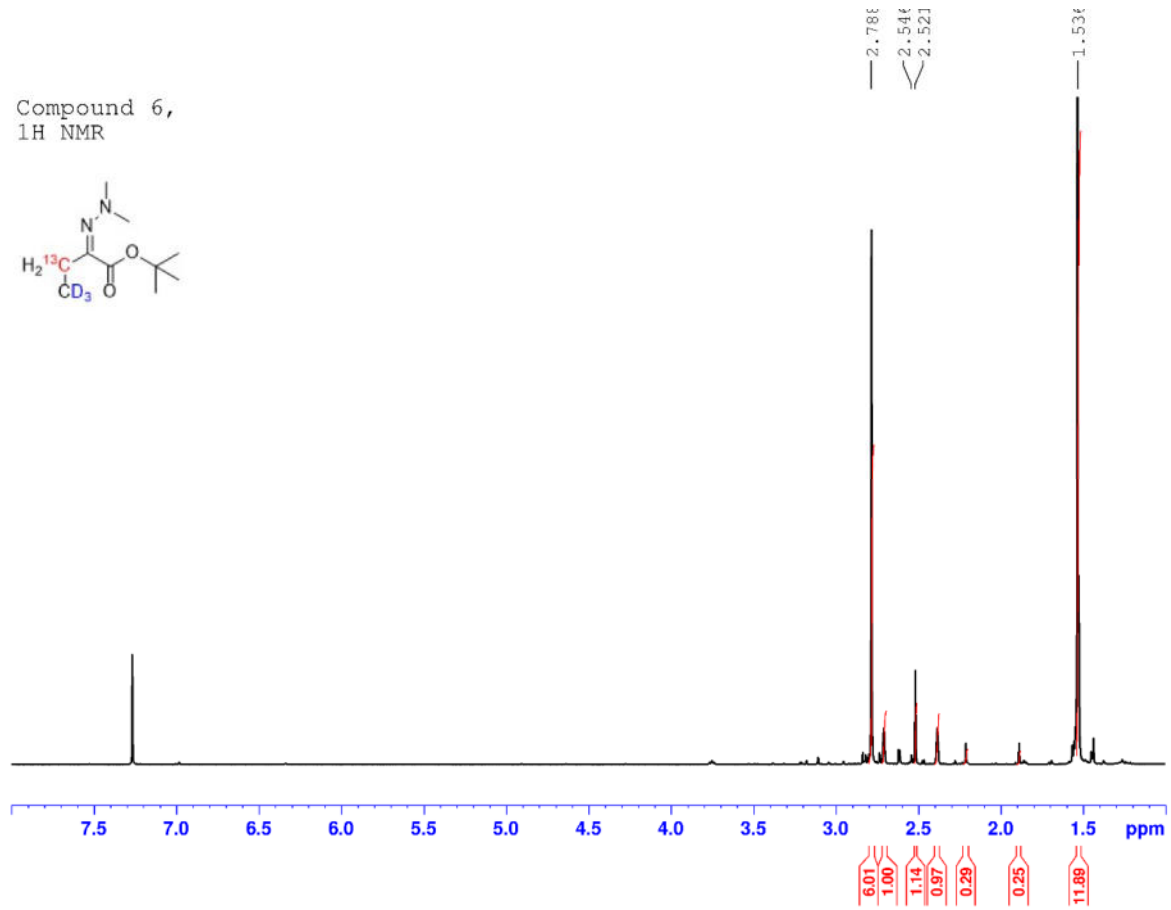


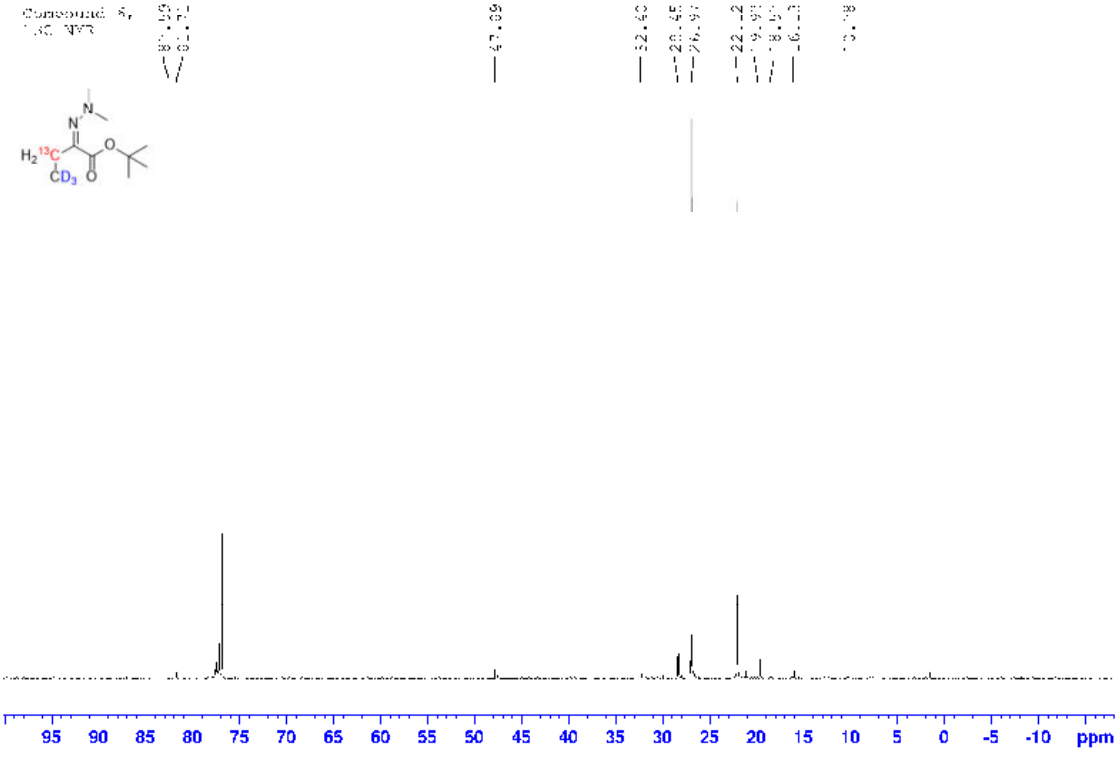

Compound 4,
1H NMR
2.429
1.561
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 ppm
2.96
9.00

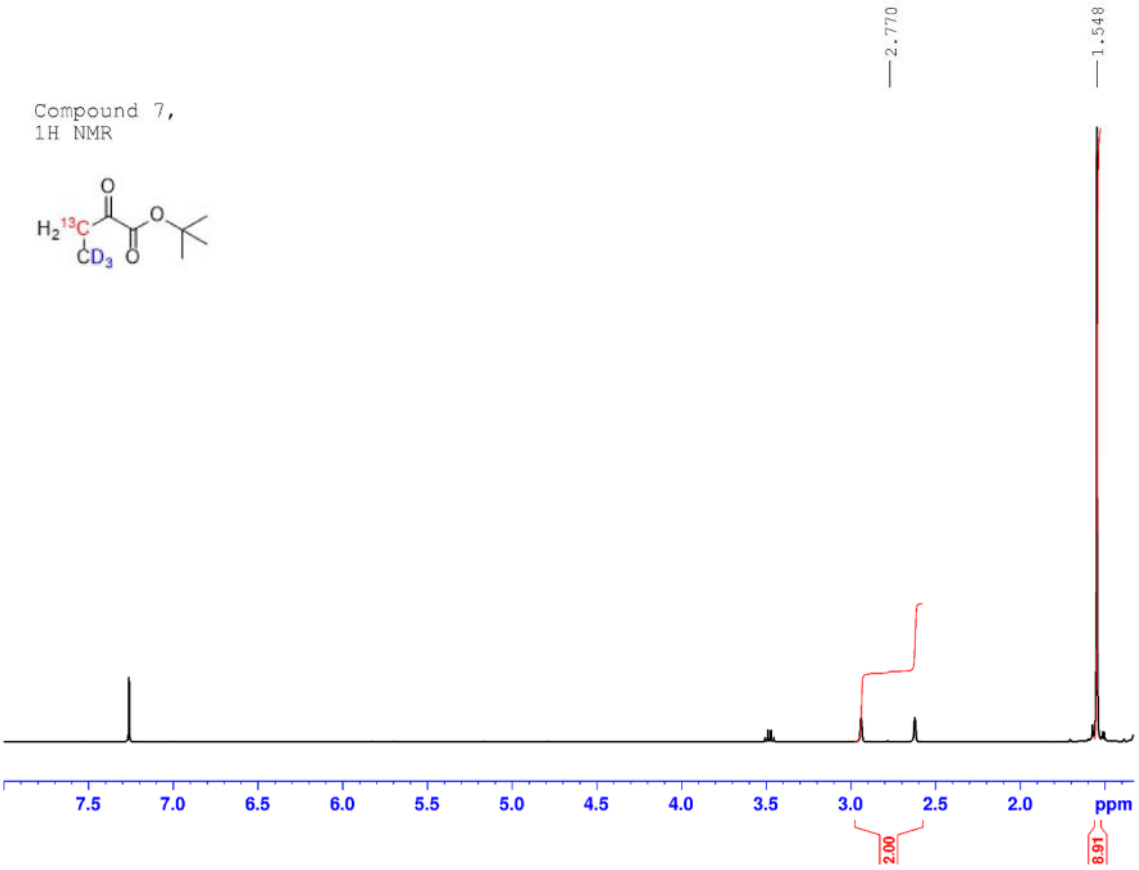

Compound 4,
13C NMR
84.13
27.93
26.56
135 130 125 120 115 110 105 100 95 90 85 80 75 70 65 60 55 50 45 40 35 30 ppm

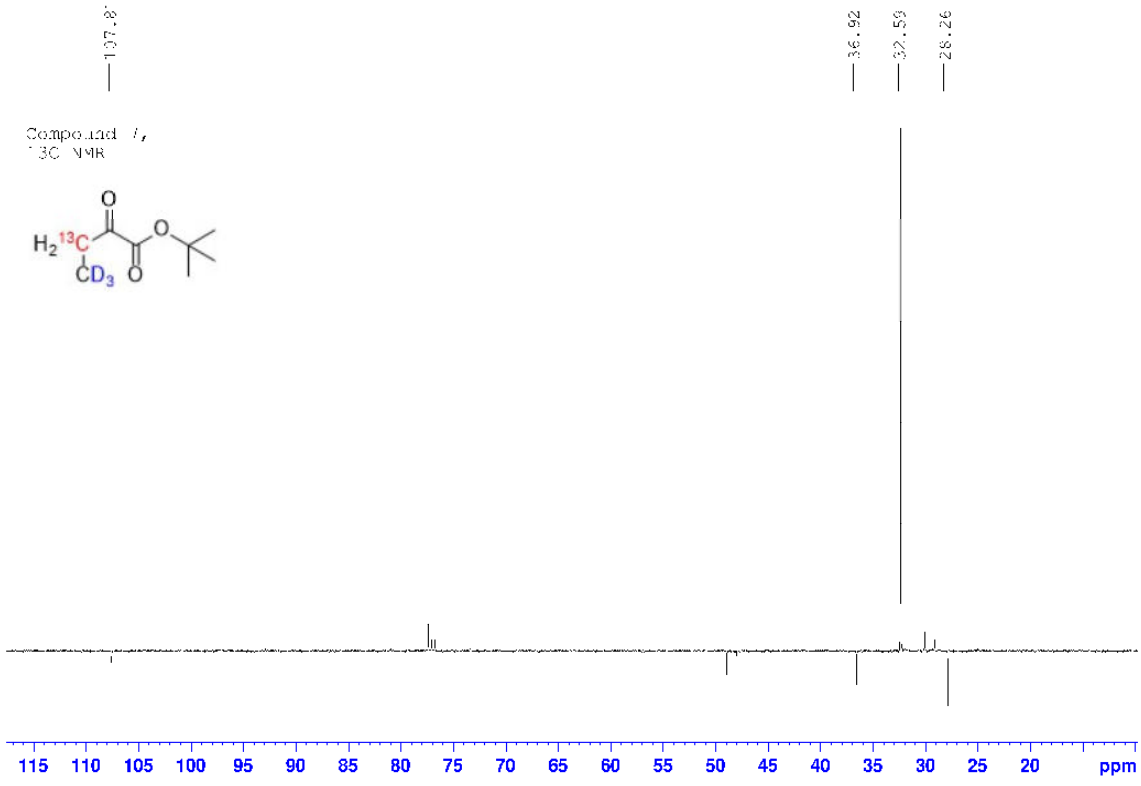

Compound 5,
1H NMR
2.796
2.517
2.068
2.048
1.531
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
ppm
5.92
3.00
9.01

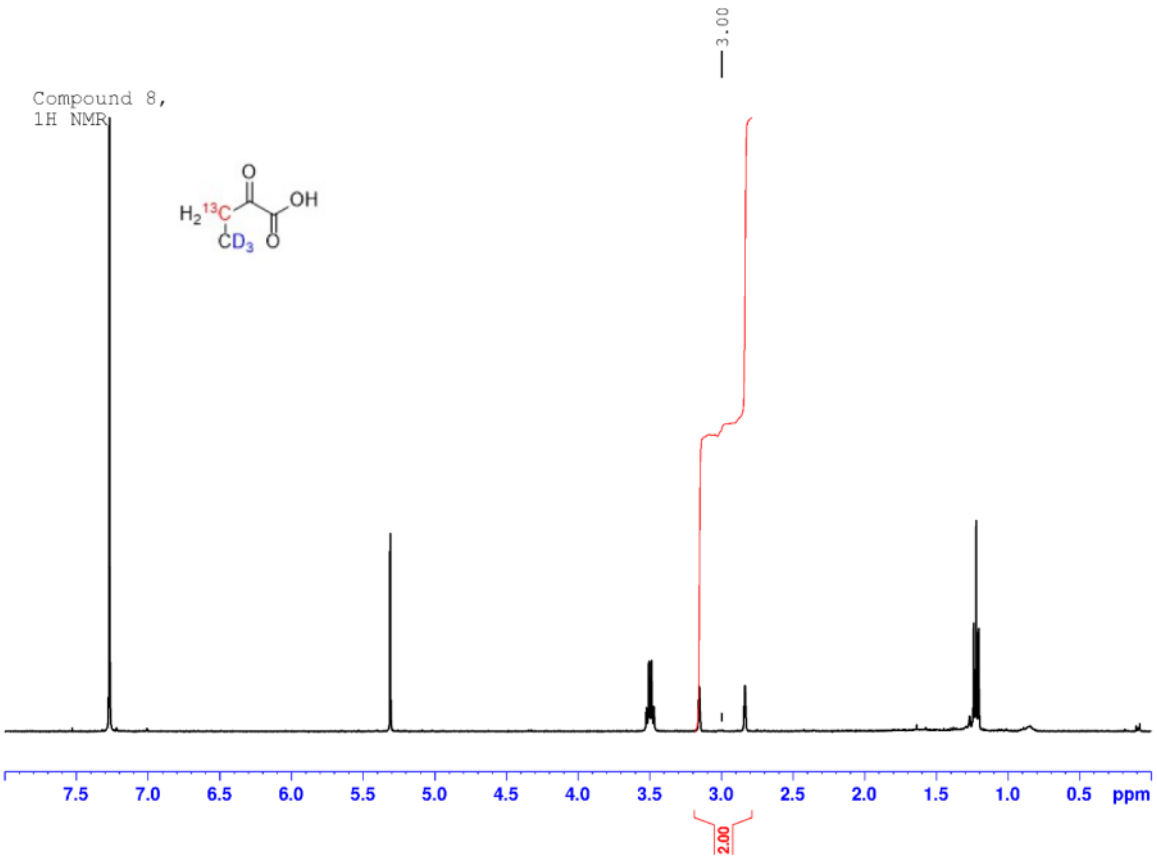

Compound 5,
13C NMR
82.61
81.64
65.92
27.52
27.06
26.09
25.61
19.48
15.95
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
ppm

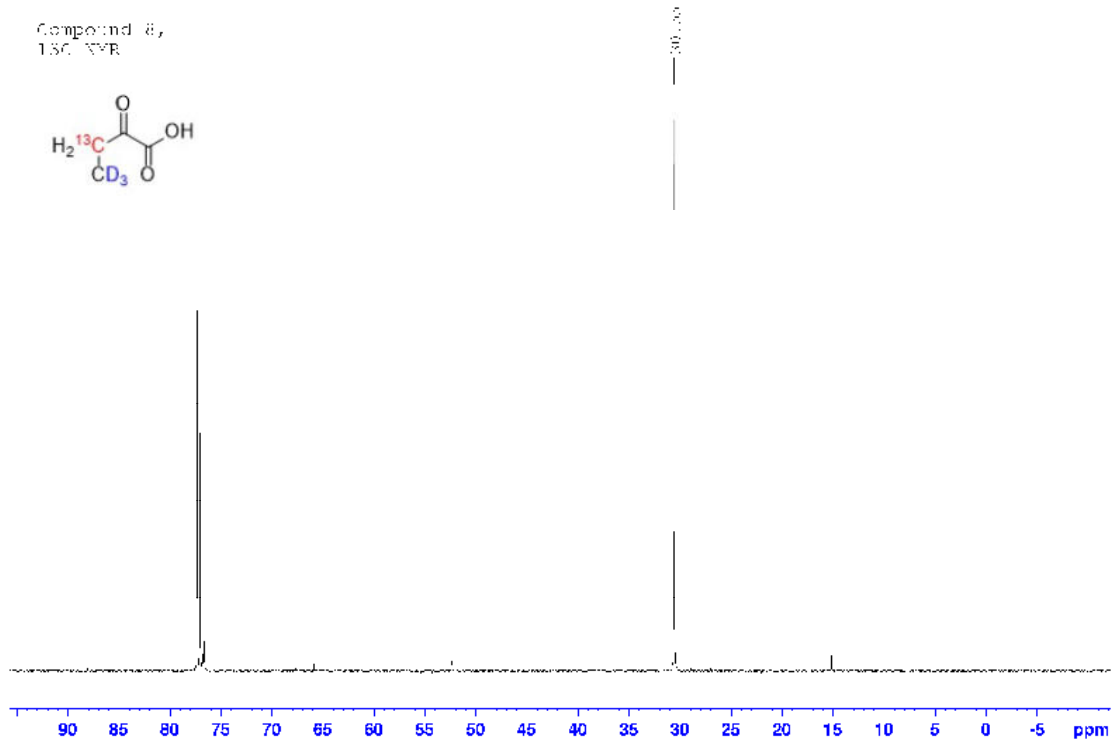

Compound 6,
1H NMR
2.788
2.546
2.521
1.534
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 ppm
6.01
1.00
1.14
0.97
0.29
0.25
11.88

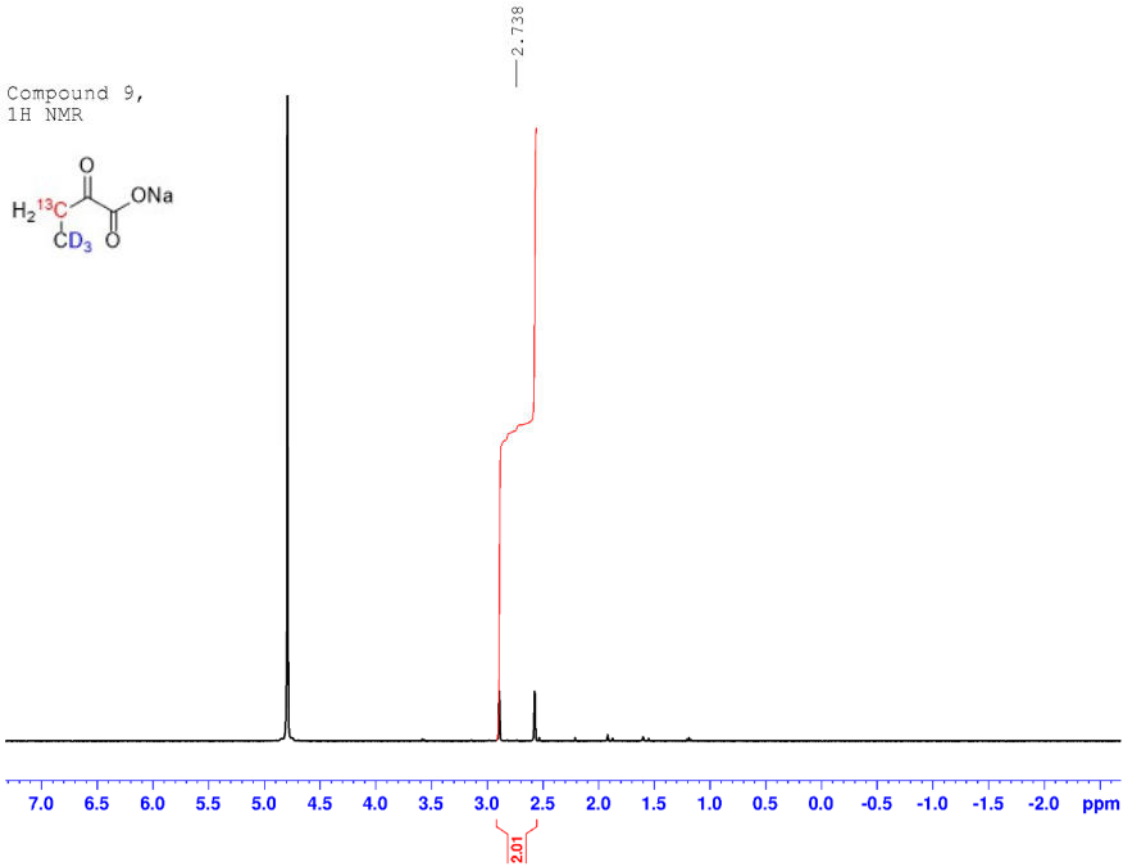

95 90 85 80 75 70 65 60 55 50 45 40 35 30 25 20 15 10 5 0 -5 -10 ppm


Compound 7,
1H NMR
2.770
1.548
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
ppm
2.00
8.91


107.8
36.92
32.59
28.26
Compound 7,
13C NMR
115
110
105
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
ppm

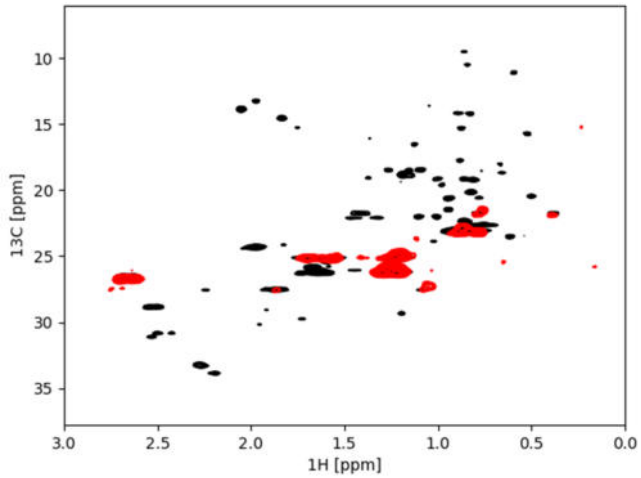

Compound 8,
1H NMR
3.00
OH
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 ppm
2.00


OH
90 85 80 75 70 65 60 55 50 45 40 35 30 25 20 15 10 5 0 -5 ppm



Incorporation of [3-^{13}C; 4,4,4-$^{2}H_3$] α-ketobutyric acid

As mentioned in the main text, spurious cross peaks are found in the methyl group region of the ^{1}H-^{13}C HSQC spectra. In order to verify that these peaks are indeed natural abundance correlations of methyl CH_3 obtained from the expression with D-Glucose (^{12}C) in H_2O medium (Figure 1, in red), we incorporated the [3-^{13}C; 4,4,4-$^{2}H_3$] α-ketobutyric acid also in D_2O minimal media. As shown in SI Figure 1 in red, the natural abundance peaks are suppressed by deuteration, and additionally, no metabolic scrambling of the precursor can be observed.



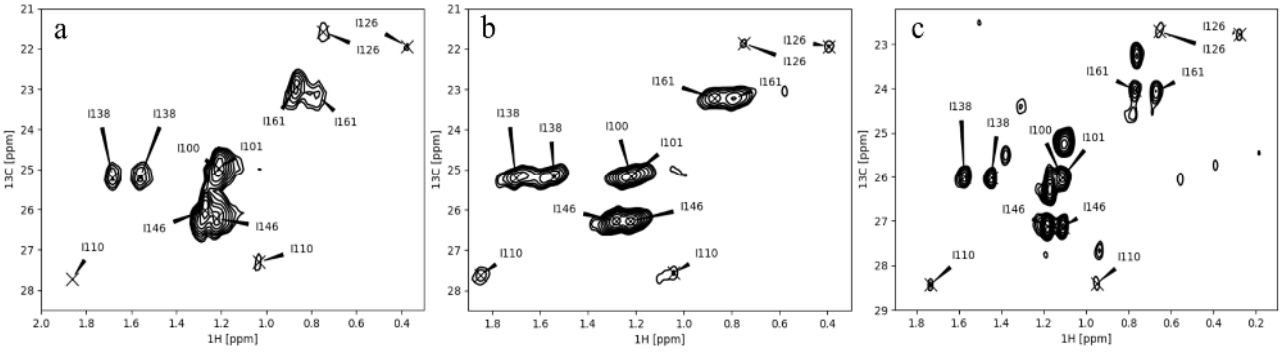
SI Figure 1 Overlay of ^{1}H-^{13}C-HSQC spectra of selectively labeled Brd4-BD1, expressed in H_2O (black) and selectively labeled Brd4-BD1, expressed in D_2O (red)

^{13}C Transverse Relaxation Studies

Analysis of the ^{13}C relaxation studies were performed by fitting the logarithmic intensities of the peaks (SI Figure 1) to the linear regression model in RStudio (RStudio Team 2020). The linearized logarithmic function was taken as $\log[I(t)]=\log(A)-t/T_2$. The peak intensities were normed to the maximum intensity of the same peak in each time step. The extracted T_2 values are displayed in milliseconds, as well as the residual standard error and the multiple R-squared of the fit (arbitrary unit).

	HTQC			HSQC		
Residue	T_2 [ms]	S	R^2	T_2 [ms]	S	R^2
ILE101	80.6	0.10	0.969	14.9	0.06	0.998
ILE126 HG12	144.3	0.13	0.850	16.2	0.21	0.967
ILE138 HG13	78.5	0.08	0.980	27.1	0.04	0.997
ILE146 HG13	66.7	0.13	0.965	28.4	0.01	0.999
ILE161 HG13	73.6	0.11	0.966	16.0	0.01	0.999
ILE161 HG12	72	0.09	0.981	18.1	0.12	0.990
ILE100	84.0	0.13	0.945	14.9	0.06	0.998
ILE146 HG12	67.1	0.13	0.965	21.4	0.05	0.998
ILE138 HG12	69.2	0.14	0.955	28.8	0.18	0.945
ILE126 HG13	83.2	0.23	0.753	-6.4	NaN	1
ILE110 HG13	113.4	0.2	0.726	14.3	NaN	1
ILE110 HG12	84.6	0.13	0.919	11.7	NaN	1

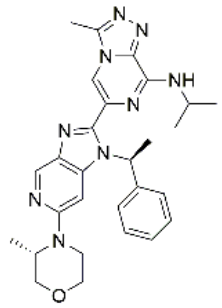
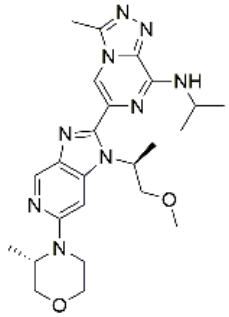
SI Table 1: extracted T_2 relaxation times for each peak observed in the CT-HSQC and CT-HTQC.



SI Figure 2 Comparison of ^{1}H-^{13}C-CT-HSQC (a), ^{1}H-^{13}C-CT-HTQC (b) and CH_2-CT-TROSY (c) spectra of perdeuterated and selectively labeled Brd4-BD1 at a constant time delay of 0.028s sec for each spectra. Signals in (c) are shifted compared to (a) and (b) due to J_{CH} coupling. Residual peaks in the CH_2-TROSY (c) are presumably due to incomplete suppression because of J-mismatch (in the S3E filter), imperfections in the TROSY element and differential ^{1}H relaxation times.

Protein-Ligand Interaction (Pople 1956; Platzer et al. 2020)

For detailed description of the equation, please refer to (Pople 1956; Platzer et al. 2020).

Ligand	Structure	CSP [ppm] pro-R	H—X [Å] pro-R	H—Y [Å] pro-R	θ [rad] pro-R	Δσ [ppm] pro-R
		pro-S	pro-S	pro-S	pro-S	pro-S
A		0.66	0.41	4.17	0.09	0.50
		1.78	0.97	2.91	0.32	1.11
B		0.59	0.28	3.93	0.07	0.61
		1.72	0.96	2.55	0.36	1.50

SI Table 2 Structures, chemical shifts and geometrical parameters used in this study

References

Hajduk PJ, Augeri DJ, Mack J, et al (2000) NMR-Based Screening of Proteins Containing 13 C-Labeled Methyl Groups. J Am Chem Soc 122:7898–7904. https://doi.org/10.1021/ja000350l

Lichtenecker R, Ludwiczek ML, Schmid W, Konrat R (2004) Simplification of Protein NOESY Spectra Using Bioorganic Precursor Synthesis and NMR Spectral Editing. J Am Chem Soc 126:5348–5349. https://doi.org/10.1021/ja049679n

Platzer G, Mayer M, Beier A, et al (2020) PI by NMR: Probing CH–π Interactions in Protein–Ligand Complexes by NMR Spectroscopy. Angew Chemie Int Ed 59:14861–14868. https://doi.org/10.1002/anie.202003732

Pople JA (1956) Proton Magnetic Resonance of Hydrocarbons. J Chem Phys 24:1111–1111

RStudio Team (2020) RStudio: Integrated Development for R

Werkhoven TM, van Nispen R, Lugtenburg J (1999) Specific Isotope Enrichment of Methyl Methacrylate. European J Org Chem 1999:2909–2914. https://doi.org/10.1002/(sici)1099-0690(199911)1999:11<2909::aid-ejoc2909>3.3.co;2-x

3

^{13}C-METHINE-LABELED VALINE AND LEUCINE PRECURSOR

G. Toscano, T. Höfurthner, B. Nagl, A. Beier, M. Mayer, L. Geist, D. B. McConnell, H. Weinstabl, R. J. Lichtenecker, and R. Konrat, "13Cβ-Valine and 13Cγ-Leucine Methine Labeling To Probe Protein-Ligand Interaction," *ChemBioChem*, 2024

DOI: 10.1002/cbic.202300762

PERSONAL CONTRIBUTION

Protein production, data analysis, validation, writing
Shared First Author with G. Toscano

Research Article
doi.org/10.1002/cbic.202300762
Chemistry Europe
European Chemical Societies Publishing
www.chembiochem.org

^{13}Cβ-Valine and ^{13}Cγ-Leucine Methine Labeling To Probe Protein Ligand Interaction.

Giorgia Toscano[+],[a, c] Theresa Höfurthner[+],[b, c] Benjamin Nagl,[a] Andreas Beier,[b] Moriz Mayer,[d] Leonhard Geist,[d] Darryl B. McConnell,[d] Harald Weinstabl,[d] Robert Konrat,*[b, e] and Roman J. Lichtenecker*[a, e]

Precise information regarding the interaction between proteins and ligands at molecular resolution is crucial for effectively guiding the optimization process from initial hits to lead compounds in early stages of drug development. In this study, we introduce a novel aliphatic side chain isotope-labeling scheme to directly probe interactions between ligands and aliphatic sidechains using NMR techniques. To demonstrate the applicability of this method, we selected a set of Brd4-BD1 binders and analyzed ^{1}H chemical shift perturbation resulting from CH-π interaction of H_{β}-Val and H_{γ}-Leu as CH donors with corresponding ligand aromatic moieties as π acceptors.

Introduction

NMR spectroscopy has developed into an important source of information for analyzing protein binding pockets and interaction surfaces at atomic resolution, complementing and extending information obtained by X-ray crystallography. Experimental NMR data encompasses a wide range of binding affinities, spanning from strong to weak, with K_D values varying from low nanomolar (nM) to millimolar (mM). In fragment-based screening approaches, NMR methods are applied to identify weak binders, which serve as starting points for further structure optimization.[1–3] In contrast to X-ray crystallography, NMR can probe the binding surface of a protein closely resembling physiological conditions. Numerous experimental techniques observing either the NMR spectra of the ligand or the protein have been developed.[4–5] The chemical shift of a protein NMR signal is a sensitive indicator for the chemical environment of the nucleus observed. Consequently, alterations in electron density due to ligand interaction lead to chemical shift perturbation (CSP). CSP studies above a certain protein size are challenged by the limited resolution and sensitivity of NMR spectroscopy, which has been addressed in the past by evolution in NMR experimental techniques and progress in hardware development.[6] Moreover, sample preparation, especially methods of selectively introducing carbon-13, nitrogen-15 and deuterium isotopes, constantly expands the potential of NMR based protein analysis and increases the size of attainable biomolecules.[7–10]

Selective isotope labeling has been developed for both aromatic and aliphatic residues.[11–17] α-Ketoisovalerate (2-oxo-3-methylbutyrate) has been introduced as an additive in *E. coli* minimal media to implement certain isotope patterns in leucines and valines, respectively.[18–20] Subsequently, advancements have been made in separating leucine and valine labeling and exploring alternative precursors for achieving stereoselective methyl labeling in these target residues.[21–25] The majority of these strategies primarily emphasize methyl labeling due to the advantageous spectroscopic characteristics of the $-CH_3$ threefold symmetry leading to favorable relaxation properties and maximum signal intensities. In contrast to these commonly used isotope patterns, our study explores the possibilities of the methine ^{13}C–H system located at the branching points of the isopropyl valine and the isobutyl leucine side chains, respectively.

The binding affinity of a ligand towards a protein target arises from various interactions such as Van der Waals forces, electrostatic interactions and hydrogen bonds.[26] Among these interactions, the XH–O (X=N, O) hydrogen bond is widely prevalent and plays a critical role in molecular recognition, serving as a main stabilizer of active protein conformations.[27] In the case of weaker CH–O hydrogen bonds, the dominance of electrostatic forces is reduced, and dispersive components gain influence.[28–29] In addition to lone pair carrying heteroatoms,

[a] G. Toscano,[+] B. Nagl, Dr. R. J. Lichtenecker
Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Institute of Organic Chemistry, University of Vienna, Währingerstraße 38, 1090 Vienna, Austria
E-mail: roman.lichtenecker@univie.ac.at
robert.konrat@univie.ac.at
[b] T. Höfurthner,[+] A. Beier, Prof. Dr. R. Konrat
Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Max Perutz Laboratories, Department of Structural and Computational Biology, Campus Vienna Biocenter 5, 1030 Vienna, Austria
[c] G. Toscano,[+] T. Höfurthner[+]
Vienna Doctoral School of Chemistry, University of Vienna, Währingerstr. 38, 1090 Vienna, Austria
[d] Dr. M. Mayer, Dr. L. Geist, Dr. D. B. McConnell, Dr. H. Weinstabl
Boehringer Ingelheim RCV GmbH & Co. KG, Dr. Boehringer Gasse 5-Wien, 11, 1121 Vienna
[e] Prof. Dr. R. Konrat, Dr. R. J. Lichtenecker
MAG-LAB, Karl-Farkas Gasse 22, 1030 Vienna
[+] These authors contributed equally to this work.
Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202300762

aromatic π-systems are enabled to take on the role of the hydrogen acceptor. Despite being relatively weak (~1,5 kcal/mol), these so-called CH-π interactions are recognized as significant modulators of ligand affinity, and several examples highlighting their impact on K_D values have been reported.[30–33] However, in most hit-to-lead structure optimization programs, the consideration of CH-π bonds is often overlooked, primarily due to the scarcity of experimental methods that can directly quantify these weak interactions.

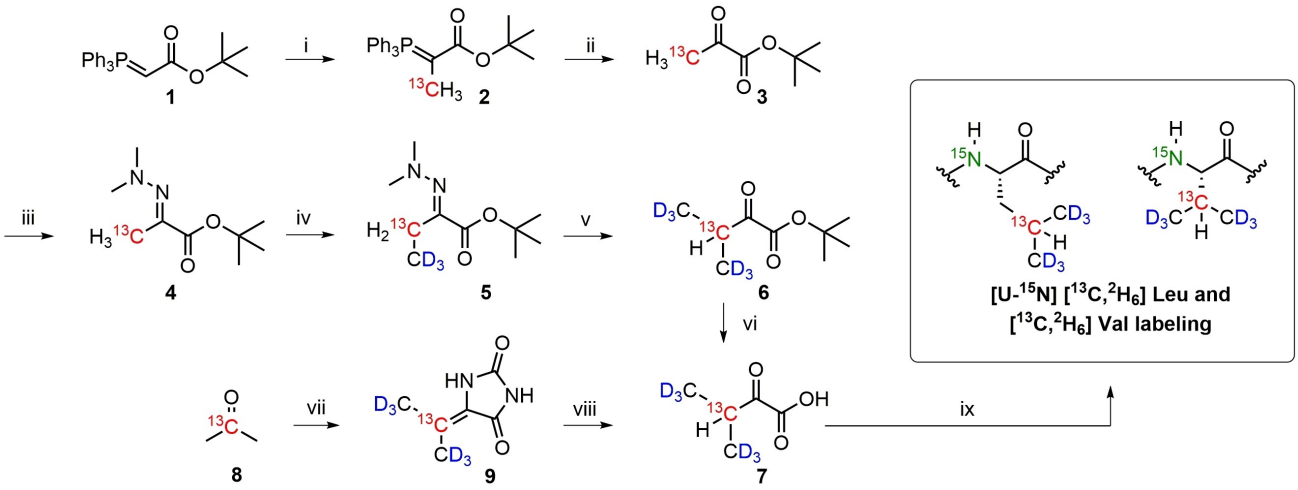
NMR spectroscopy has been used to detect methyl-π interactions via J-couplings in the past.[34] Recently, we presented a novel approach to correlate CSP-values with favorable CH-π interactions.[35–36] Specifically, we investigated the geometric parameters between aromatic ligands and labeled tryptophan residues of the corresponding target protein (PI by NMR).[35] In another application, we could assign the CSP values to *proS* and *proR* hydrogen – ligand interactions in corresponding methylene isoleucine labelled protein samples.[36] We have now pursued this concept further to the aliphatic residues leucine and valine. This extension is motivated by two key factors: Firstly, minimizing the exposure of nonpolar regions to water is the main thermodynamic driving force for a protein to fold into a conformation with a binding pocket enriched in hydrophobic residues. In the case of solvent-exposed surface amino acids Leu or Val are represented to a lesser extent, but if present have a high probability to interact with a molecular binding partner.[37–38] Secondly, therapeutically active compounds frequently feature aromatic systems. Recent statistical analysis of physicochemical and structural properties reveal that 76% of drugs approved within the last decade feature one or more heteroaromatic ring structures.[39]

Taking these factors into account, along with additional supporting evidence from the literature, we propose that $C–H_{Leu/Val}–\pi_{Ligand}$ interactions serve as significant criteria in fine-tuning weak and reversible molecular recognition events.[40–41]

Results and Discussion

A literature-based route to access ^{13}C-pyruvate **3** was modified and adapted to our purpose (scheme 1).[42] Methylation of the ylide **1** was accomplished using ^{13}C-iodomethane as an economic source of carbon-13. Ozonolysis yielded the *tert*-butyl pyruvate **3**, which was subsequently converted to the two isomers of the dimethylhydrazone derivative **4** (=major isomer).[20] Subsequent methylation of this hydrazone was described previously,[43] but following this literature reported procedure using lithium diisopropyl amide (LDA) as a base did not result in the desired product in our hands. Instead, we optimized this key step and identified reaction conditions applying sodium bis(hexamethylsilyl)amide (NaHMDS) to abstract the proton in α-position.[36] Reduced reaction times were essential to efficiently perform the conversion of compound **4** to **5** (=major isomer). The remaining steps have been conducted according to protocols we developed in earlier work, but now applied on different isotopologues.[20] We additionally established an alternative route to access precursor **7** starting from [2-^{13}C] acetone **8** (scheme 1). Condensation with hydantoin in presence of D_2O yielded compound **9**. The target compound **7** was obtained after basic hydrolysis of the hydantoin ring.[44] NaOD was applied in this step to avoid methyl protonation and the α-position was subsequently protonated by treatment with 1 M NaOH at room temperature. This sequence resulted in a methyl deuteration grade of >97%. This second synthetic approach features a different source of carbon-13 (acetone instead of methyl iodide) and is distinguished by a very effective two-step reaction sequence. [3-^{13}C, 4,4′-$^{2}H_6$] Ketoisovaleric acid **7** was added to the $^{15}NH_4Cl$ supplemented minimal medium of a BL21(DE3) *E. coli* based overexpression system for Brd4-BD1.

We chose the binding domain one of the bromodomain-containing protein 4 (Brd4-BD1) as our model system and correlated chemical shift perturbations with the presence of



Scheme 1. Reagents and conditions: i. $^{13}CH_3I$, CH_2Cl_2, overnight, then NaOH/H_2O; ii. Ozone, CH_2Cl_2, −78 C, then PPh_3, RT, 2 hours, 70% over two steps; iii. N,N-dimethylhydrazine $_{asym.}$, diethyl ether, RT, overnight, 76%; iv. NaHMDS, $THF_{anhyd.}$, −78 C, CD_3I, 5 min, 80%; v. LDA, $THF_{anhyd.}$, −78 C, CD_3I, 2 hours; then diethyl ether, 1 N HCl, RT, 71%; vi. CH_2Cl_2/diethyl ether 1:1, HCl_g, 6 h, 90%. vii. Hydantoin, ethanolamine, D_2O, overnight, 70 C; viii. NaOD/D_2O, 120 C, 5 hours; then NaOH (1 M), RT, 3 hours, 72%. ix. *E. coli* overexpression, $^{15}NH_4Cl$.

Research Article
doi.org/10.1002/cbic.202300762
Chemistry Europe
European Chemical Societies Publishing

CH-π interactions. Brd4 affects oncogene expression and corresponding inhibitors have already revealed results in clinical trial tumor treatment.[45–46] An in-house collection of protein-ligand co-crystal structures provides extensive structural data to evaluate our NMR based approach.[47] Our novel precursor enables the introduction of a ^{13}C–H spin system at the β-position of valine and the γ-position of leucine (Scheme 1). The ^{1}H–^{13}C–HSQC of the resulting protein sample revealed well-defined signals, which arise from ^{13}CH methine spin pairs of five valine- and thirteen leucine residues in the Brd4-BD1 sequence (Figure 1). While the valine signals have been assigned using a combination of NMR experiments recorded on a uniformly ^{13}C labeled protein sample (see experimental part), correlations of methine resonances to the leucine residues are currently not available. In order to explicitly identify the signals of two leucine side chains, which are located in the binding site (Leu92 and Leu94; see Figure 1C), corresponding L92A and L94A mutants were overexpressed and studied by 2D NMR (supporting information). Thereby, two peaks (Figure 1D) were assigned to these two residues. Since our precursor labeling scheme does not feature a chiral carbon center, stereoselective synthesis is not required to further reduce protein spectra complexity, as e.g. in the case of single methyl group labeling. However, in the case of high molecular weight proteins it might be beneficial to separate valine from leucine labeling and reduce signal overlap. In this case, leucine isotope labeling can be inhibited by adding unlabeled α-ketoisocaproate to the over-expression medium as described in earlier work.[22]

Applying a fully ^{13}C labeled version of precursor **7** can have the additional advantage of signal assignment using triple resonance NMR experiments. It is of special value that both synthetic routes depicted in scheme 1 may lead to this *all*-^{13}C precursor **7** if corresponding isotope sources are used. On the one hand, the stable ylide **1** can be synthesized from commercially available [1,2-$^{13}C_2$] bromoacetate and be transformed to uniformly ^{13}C α-ketoisovalerate by using $^{13}CH_3I$ as a reagent in the subsequent reaction steps. On the other hand, the *all*-^{13}C isotopologue of **7** is also accessible using [$^{13}C_3$]

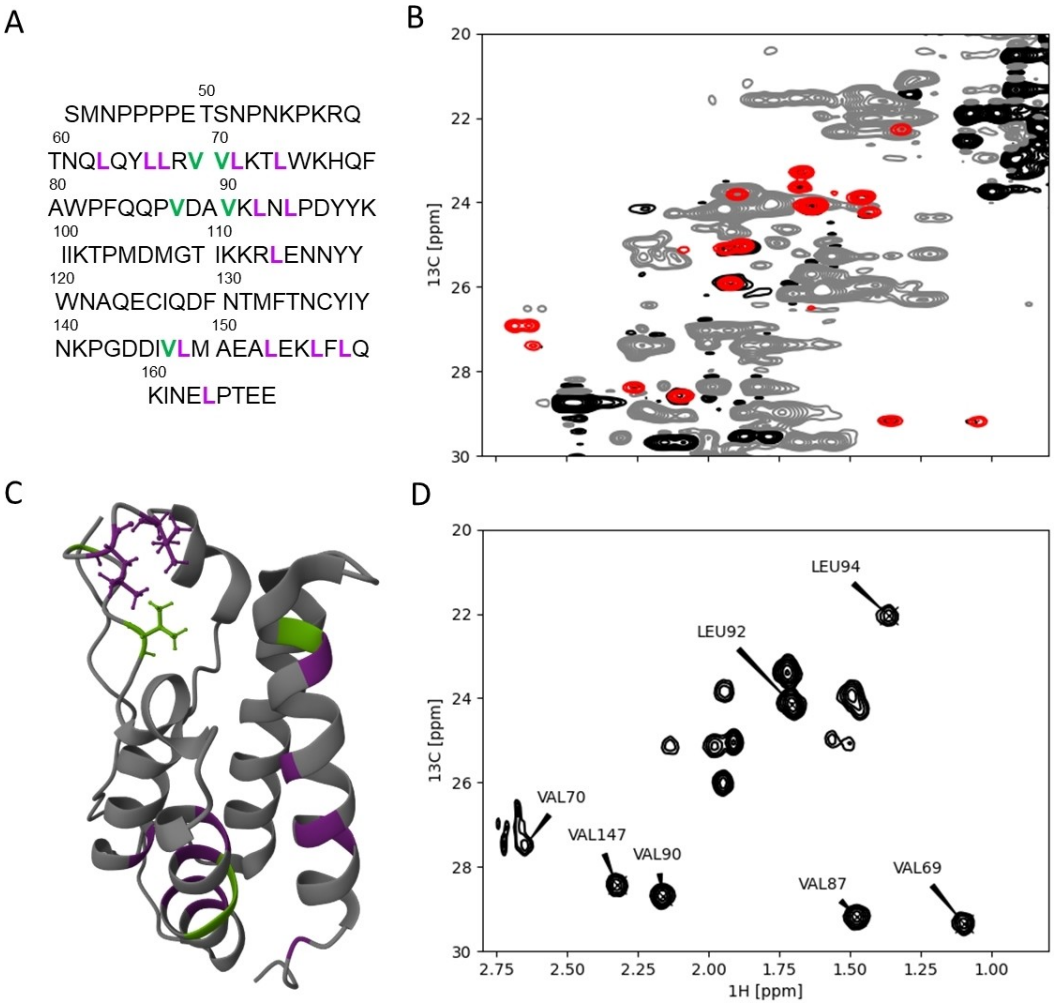


Figure 1. A) Brd4-BD1 sequence featuring 13 leucines (purple) and 5 valines (green). B) Overlay of ^{1}H-^{13}C-HSQC spectra of selectively Leu and Val ^{13}CH labeled Brd4-BD1 (red) onto the ^{1}H-^{13}C-HSQC of ^{13}C-uniformly labeled Brd4-BD1 (^{13}CH and $^{13}CH_3$ signals in black and $^{13}CH_2$ signals in grey), zoomed into the CH region. C) Brd4-BD1 structure (PDB entry 6XUZ shown without ligand for clarity) highlighting leucine- (purple) and valine residues (green). The side-chains of Leu- and Val residues within the binding site (Leu92, Leu94 and Val87) are depicted top left. D) ^{1}H-^{13}C-SOFAST-HMQC of labeled Brd4-BD1 using BL21(DE3) *E. coli* in M9 minimal medium supplemented with compound **7**.

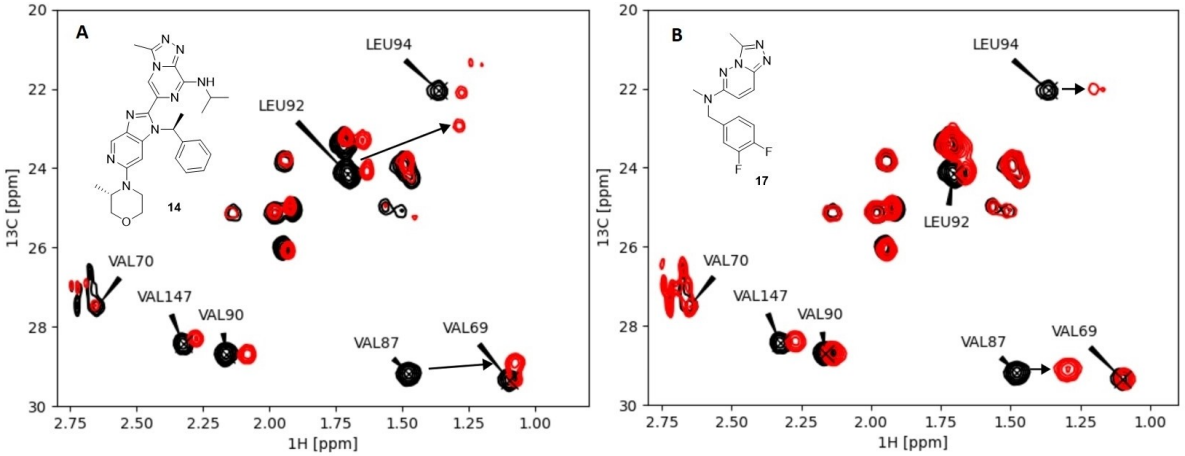


Figure 2. ^{1}H-^{13}C-SOFAST-HMQC of selectively Leu/Val labeled Brd4-BD1 in absence (black) and in presence (red) of ligand **14** (A) or ligand **17** (B). Chemical shift perturbations > 0,25 ppm are indicated with black arrows.

acetone and [1,2-$^{13}C_2$] hydantoin, with the latter being synthesized from commercially available [$^{13}C_2$] glycine as described in literature.[21]

Ten different ligands (compounds **10**–**19**, see supporting information for corresponding structures) were added to the selectively Val/Leu labeled protein samples and the resulting CSPs recorded by acquisition of ^{1}H-^{13}C-HMQC SOFAST spectra. The protein-ligand affinities had been analyzed in previous studies and the ligands were synthesized following literature procedures.[35,47] The ligands contain strong (nM) and weak (μM) binders and corresponding structural details are given in the supporting information. For each of these compounds, the binding mode to Brd4-BD1 is known and some of them are available in the PDB database (6XV7, 6XUZ and 6XV3).

In the case of high affinity binders (e.g. the triazolo-pyrazine/5-azabenzimidazole derivative **14**; K_D = 38 nM), overlay of the ^{1}H-^{13}C-SOFAST-HMQC of the protein in the apo-state and in presence of the ligand displays significant chemical shift perturbation of two leucine and one valine methine signal (Figure 2A). This finding corresponds to the reported binding mode, indicating two leucines (Leu92 and 94) and one valine (Val87) side chains within the binding site (PDB 6XV3). Taking a closer look at the binding mode, the distance between the Leu92 methine proton and the 5-azabenzimidazole ring system in ligand **14** displays a CH-π interaction, which is detected by the large CSP observed (Figure 3 top). In the case of the Val87 methine proton, the structural data again corroborates the large CSP observed. Worth to notice here is the already moderate upfield position of the Val87 methine resonance, which is due to the proximal aromatic side chain of Tyr97 (Figure 3 bottom). An additional upfield shift upon ligand interaction is due to the positioning of the triazolo-pyrazine system of the ligand, which leads to additional shielding of the methine group.

In case of weaker binders (e.g. compound **17**; K_D 8,9 μM), the absence of the large bicyclic π-system impedes a significant interaction with Leu92 and a CSP is consequently missing in the corresponding ^{1}H-^{13}C-SOFAST-HMQC spectrum (Figure 2B). Spectra of Brd4-BD1 in presence of ligands **10**–**19** confirm the finding that a Leu92 CSP is only observed in cases where the ligand structure features a rigid bicyclic aromatic system as corresponding CH-π acceptor (supporting information). Hence, our method provides an effective and direct tool to answer the question if certain ligands serve as appropriate acceptors in noncovalent CH-π leucine/valine interaction.

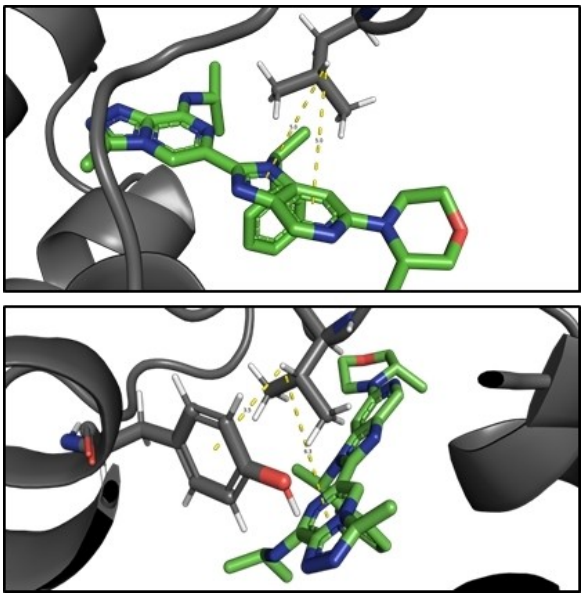


Figure 3. Orientation of ligand **14** to Leu92 (top) and Val87 (bottom) according to PDB entry 6XV3. Leu92 is in proximity to the 5-azabenzimidazole of the ligand, whereas Val87 is oriented close to the triazolo-pyrazine system, as well as the Tyr97 side-chain.

The extent of chemical shift perturbations caused by CH-π interactions is determined by the proximity between the proton donor and the acceptor aryl rings, as well as their spatial orientation. The effect of an aromatic ring current on a proton signal CSP can be described by the point-dipole model introduced by Pople (Figure 4A).[48] The corresponding structural parameters in the Pople equation for Leu92 and Leu94 (r and θ,

Research Article
doi.org/10.1002/cbic.202300762

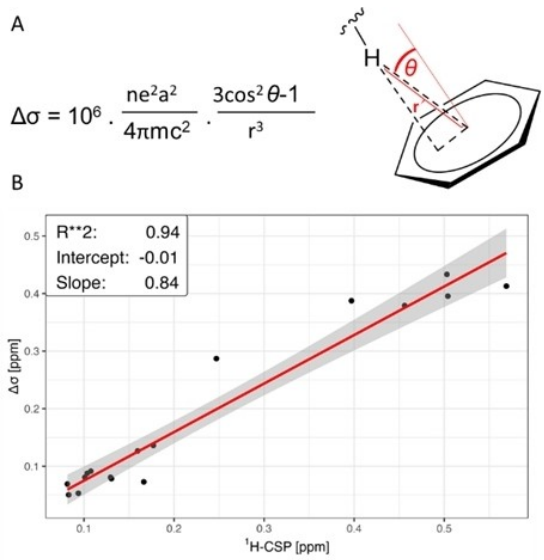


Figure 4. Pople equation illustrating the change in the isotropic nuclear shielding constant (Δσ) of a proton depending on the relative orientation of a proton to an aromatic π-system (A). Parameters: Δσ…isotropic nuclear shielding constant [ppm]; n…number of circulating electrons (=6); e…elementary charge [Franklin] (=4.803207×10^{-10}); a…radius of the aromatic ring [cm] (=1.39×10^{-8} for 6-membered ring and 1.18×10^{-8} for 5-membered ring); m…electron mass [g] (=$9{,}1094\times10^{-28}$); c…speed of light [cm s^{-1}] (=2.998×10^{10}); Θ…angle between the ring normal through the aromatic center and the proton to ring center vector [rad] (see SI for corresponding values); r…distance from the proton to the ring center [cm] (see SI for corresponding values). (B) Correlation between the changes in chemical shift obtained from the Pople equation and the experimentally observed CSP values for Leu92 and Leu94 residues of Brd4-BD1 in presence of different ligands.

see table SI1 in the supporting information) have been extracted from x-ray data.

The correlation of experimental CSP values for Leu92 and Leu94 with predictions made using the Pople equation for different ligands is shown in Figure 4B. Overall, the positive correlation of Δσ derived from structural data and the observed CSP suggests that the observation of methine proton signals has the capability to serve as an additional tool for evaluating binding modes within protein-ligand complexes. Interestingly, the slightly smaller slope of 0.84 for linear regression is comparable to the value (0.92) for aromatic CH-π interactions.[35] Pople's approximation describes the secondary magnetic field induced by the electric ring current via a magnetic dipole positioned at the ring center. The implementation of this model offers high simplicity and decreased calculation efforts. However, more advanced models to represent the ring current effect considering the current circulating above and below the ring plane, as well as the Hückel molecular orbital theory can be included in future studies to further expand the methodological possibilities.[49–51]

Conclusions

We developed a method to introduce ^{13}C-^{1}H spin systems into the methine positions of valines and leucines in presence of deuterated methyl groups. A corresponding isotopologue of the metabolic precursor α-ketoisovaleric acid **7** was synthetized in two different approaches using the economic isotope sources $^{13}CH_3I/CD_3I$ and [2-^{13}C] acetone/D_2O, respectively. The compound was applied in the *E.coli*-based overexpression of selectively labeled Brd4-BD1 resulting in well-resolved ^{13}C-HSQC NMR spectra. CSPs were observed upon addition of ten different ligand compounds with known binding modes. In this study, we could show, that the magnitudes of the observed CSPs correlate with the geometrical arrangement of the interacting moieties via Pople's equation and can be exploited as a valuable data source to characterize protein ligand interaction. Methine-based CSP experiments can thus complement methods focusing on methyl resonances, especially in the case of small to medium-sized proteins and serve as an additional guiding principle in ligand optimization processes.

Experimental Section

General synthetic procedures: Unless otherwise stated, all reagents and reactants were purchased from commercial suppliers and used without further purification. All solvents were distilled before use. Iodomethane-^{13}C and Iodomethane-d_3 were purchased from Sigma Aldrich®. Oxygen- and moisture sensitive reactions were carried out under an argon atmosphere and yields refer to pure compounds. The reactions were monitored via thin layer chromatography (TLC) on silica gel 60 with fluorescent indicator UV254. Visualization of the compounds was carried out using an UV-lamp (254 nm) and by application of specific reagents: $H_3PMo_{12}O_{40}$ (10% in ethanol) with subsequent heating using a heat-gun. Flash column chromatography was performed on silica gel 60 (0.040–0.063 mm) from Merck. Freeze-drying was done with liquid nitrogen and subsequent lyophilization by inducing high vacuum on an oil pump. ^{1}H and ^{13}C NMR spectroscopic data were recorded on a Bruker AVANCE-DRX 400 MHz spectrometer. NMR solvent signals were calibrated to 7.27 ppm ($CDCl_3$) and 4.79 ppm (D_2O). Chemical shifts (δ) are given in ppm (s=singlet, d=doublet, dd=doublet of doublets, m=multiplet) and reported relative to the residual solvent peaks. Coupling constants (*J*) are given in Hertz (Hz). High resolution mass spectrometry experiments were performed using electrospray ionization (ESI, 3 keV, in the positive or negative ion mode) or electron ionization (EI, 70 eV).

Synthesis of precursor (7) starting from triphenylphosphorane (1)

[3-^{13}C] 1,1-Dimethylethyl 2-(triphenylphosphoranylidene)propanoate (2): 2,9 g (7,7 mmol) of (*tert*-butoxycarbonylmethylene)triphenyl-phosphorane **1** was dissolved in 15 ml of dichloromethane (DCM). To this solution, 1 g (0,9 eq) of $^{13}CH_3I$ was added. The mixture was stirred overnight, and the solvent removed under reduced pressure. The resulting yellow residue was dissolved in 30 ml of DCM and vigorously shaken with a solution of 560 mg (2 eq) of NaOH in 10 ml of water using a separatory funnel. The aqueous phase was extracted twice with DCM. The combined organic layers were dried over $MgSO_4$, resulting in a clear, yellow product solution, which was used for the next step without further purification.

[3-^{13}C] *tert*-Butyl pyruvate (3): A constant stream of O_3 from an ozone generator was purged through the solution of the ylid **2** in DCM (resulting from the preceding reaction) at −78 C until the

solution turned blue. Subsequently, a stream of argon was purged through the solution until the blue colour vanished. 2,2 g (1,2 eq) of triphenylphosphine (PPh_3) were added and the solution was stirred under inert conditions for another 15 min. at −78 C. The mixture was then brought to room temperature (RT) and stirring was continued for two hours. The bulk solvent was carefully removed under reduced pressure. [3-^{13}C] *tert*-butyl pyruvate **3** was obtained by bulb-to-bulb distillation (100 mbar, up to 90 C) as a colorless liquid. Yield: 782 mg (70% over 2 steps) ^{1}H NMR (400 MHz, $CDCl_3$, ppm): δ = 2.42 (d, $^1J_{C,H}$ = 129.2 Hz, 3H; $^{13}CH_3$), 1.56 (s, 9H; $C(CH_3)_3$). ^{13}C NMR (100 MHz, $CDCl_3$, ppm) δ = 84.1 ($C(CH_3)_3$), 27.9 ($C(CH_3)_3$), 26.4 ($^{13}CH_3$).

[3-^{13}C] *tert*-Butyl 2-(2,2-dimethylhydrazono)propanoate (4): 200 μl (1.2 equivalents) of dimethylhydrazine were added to a solution of 311 mg (2.1 mmol) of [3-^{13}C] *tert*-butyl pyruvate **3** in 5 ml of diethyl ether (Et_2O) under an argon atmosphere. The resulting mixture was stirred at RT overnight. The following day, the solution was diluted with additional Et_2O and subjected to two washes with small amounts of water. The aqueous phases where extracted three times with small portions of Et_2O. The organic phases were combined, dried over magnesium sulfate ($MgSO_4$) and concentrated under reduced pressure. The resulting yellow liquid, [3-^{13}C] *tert*-butyl 2-(2,2-dimethylhydrazono)propanoate **4**, was purified by bulb-to-bulb distillation (10 mbar, up to 110 C). The yield was 303 mg (76%), consisting of two isomers in an approximate ratio of 5 : 1. ^{1}H NMR (400 MHz, $CDCl_3$), Isomer 1: δ = 2.79 (s, 6H; $N(CH_3)_2$), 2.06 (d, $^1J_{C,H}$ = 129.7 Hz, 3H; $^{13}CH_3$), 1.53 (s, 9H; $C(CH_3)_3$); Isomer 2: δ = 2.51 (s, 6H; $N(CH_3)_2$), 2.04 (d, $^1J_{C,H}$ = 129.7 Hz, 3H; $^{13}CH_3$), 1.53 (s, 9H; $C(CH_3)_3$). ^{13}C NMR (100 MHz, $CDCl_3$, ppm) Isomer 1 δ = 15.95 ($^{13}CH_3$), 26.61 ($C(\mathbf{C}H_3)_3$, 47.06 ($N(CH_3)_2$, 81.64 $\mathbf{C}(CH_3)_3$); Isomer 2 δ = 13.48 ($^{13}CH_3$), 28.09 ($C(\mathbf{C}H_3)_3$, 47.52 ($N(CH_3)_2$, 82.61 $\mathbf{C}(CH_3)_3$).

[3-^{13}C; 4,4,4-$^2H_3$] *tert*-Butyl 2-(2,2-dimethylhydrazono)butanoate (5): A solution containing 303 mg (1.62 mmol) of compound **4** in 15 ml of anhydrous tetrahydrofuran (THF) was prepared under an argon atmosphere. The solution was then cooled to −78 C. Next, 3.25 ml (2 equivalents) of a 1 M solution of sodium hexamethyldisilazide (NaHMDS) in a mixture of THF and hexanes was slowly added to the reaction over a period of 30 seconds. Stirring was continued for 2 minutes, after which 102 μl (1 equivalent) of CD_3I was rapidly introduced. The reaction mixture was stirred for additional 5 minutes. To stop the reaction, 5 ml of water were added, and the mixture was brought to RT. After the separation of the phases, the aqueous layer was extracted four times with diethyl ether (Et_2O). The combined organic phases were washed with water and dried over magnesium sulfate. The solvent and unwanted silyl by-products were removed by applying reduced pressure (50 mbar) and heating at 40 C for several hours until mass consistency. The resulting product, [3-^{13}C; 4,4,4-$^2H_3$] *tert*-butyl 2-(2,2-dimethylhydrazono)butanoate **5**, was obtained as a clear, yellow liquid by bulb-to-bulb distillation under vacuum (10 mbar) up to a temperature of 110 C. Yield: 264 mg of two isomers in a ratio of approx. 4 : 1 (80%) ^{1}H NMR (400 MHz, $CDCl_3$, ppm) Isomer 1: δ = 2.77 (s, 6H; $N(CH_3)_2$), 2.53 (d, 2H; $^1J_{C,H}$ = 129.7 Hz), 1.52 (s, 9H; $C(CH_3)_3$); Isomer 2: δ = 2.50 (s, 6H; $N(CH_3)_2$), 2.35 (d, $^1J_{C,H}$ = 129.7 Hz, 2H; $^{13}CH_2$), 1.52 (s, 9H; $C(CH_3)_3$). ^{13}C NMR (100 MHz, $CDCl_3$, ppm) Isomer 1 δ = 26.62 ($^{13}CH_2$); Isomer 2 δ = 21.77 ($^{13}CH_2$). HRMS (ESI): m/z calcd for $C_9{}^{13}CH_{17}D_3N_2O_2 + Na^+$:227.1633 [$M + Na^+$]; found: 227.1638.

[3-^{13}C; 4,4,4-$^2H_3$] *tert*-Butyl 2-keto-3-(methyl-$^2H_3$)-butanoate (6): 128 mg (0.68 mmol) of compound **5** were dissolved in 10 mL of THF under an argon atmosphere. The solution was cooled to −78 C, and then 0.7 mL of commercial Lithium diisopropylamide (LDA) (2 M in THF, 1.36 mmol) was added. The mixture was stirred at −78 C for 30 minutes, followed by the rapid addition of 124 μl CD_3I (1.36 mmol). The progress of the reaction was monitored using NMR spectroscopy, which showed completion after 30 minutes. The solvent was evaporated at reduced pressure at 40 C; the resulting residue was taken up in ethyl acetate and washed with 10 mL of water. The water phases were discarded while the organic phase was vigorously shaken with 3 mL of aqueous HCl (1 M). The organic phase was dried over $MgSO_4$, and the solvent was evaporated under reduced pressure. The resulting pale-yellow oil was purified using bulb-to-bulb distillation to a vacuum pressure of 20 mbar and a temperature of 90 C, which yielded 87 mg (71%) of [3-^{13}C; 4,4,4-$^2H_3$] *tert*-butyl 2-keto-3-(methyl-$^2H_3$)-butanoate **6** as a colorless oil. ^{1}H NMR (400 MHz, $CDCl_3$, ppm): δ = 3.14 (d, $^1J_{C,H}$ = 129.8 Hz, 1H; ^{13}CH), 1.56 (s, 9H; $C(CH_3)_3$). ^{13}C NMR (100 MHz, $CDCl_3$, ppm): δ = 83.92 ($C(CH_3)_3$), 36.53 (^{13}CH), 27.90 ($C(CH_3)_3$).

[3-^{13}C; 4,4,4-$^2H_3$] 2-Keto-3-(methyl-$^2H_3$)-butanoic acid (7). A solution of 143 mg (0.47 mmol) compound **6** in a mixture of 10 mL diethyl ether (Et_2O) and 10 mL dichloromethane (DCM) was cooled to 0 C. Gaseous HCl was purged into this solution by adding concentrated HCl to concentrated sulfuric acid dropwise. The resulting gas was bubbled through the solution for 15 minutes at a moderate pace. After that, the reaction vessel was tightly closed and stirred at room temperature for 1 hour. The solution was then cooled again to 0 C, and this process was repeated three times. Stirring was continued overnight after the final cycle. The following day, the solvent was removed under reduced pressure, resulting in the formation of 58 mg (99%) of [3-^{13}C; 4,4,4-$^2H_3$] 2-keto-3-(methyl-$^2H_3$)-butanoic acid **7** as a clear, colorless viscous liquid. ^{1}H NMR (400 MHz, $CDCl_3$, ppm): δ = 3.45 (d, $^1J_{C,H}$ = 131.0 Hz, 1H; ^{13}CH). ^{13}C NMR (100 MHz, $CDCl_3$, ppm): δ = 34.76 (^{13}C), HRMS: *m/z* calcd for $C_4{}^{13}CH_2D_6O_3$: 122.0811 [*M*-H]$^-$ (100%) and 123.0845 [*M*-H]$^-$ (4.3%); found: 122.0812 (100%) and 123.0851 (4,3%).

Synthesis of precursor (7) starting from [2-^{13}C] acetone (8)

5-([2-^{13}C; 3,3,3,3′,3′,3′-$^2H_6$] Propan-2-ylidene)imidazolidine-2,4-dione (9). 150 μL (1 eq) of [2-^{13}C] acetone and 120 μL (1 eq) of ethanolamine were added to 306 mg (1.5 eq) hydantoin in 2 mL of D_2O and the resulting mixture stirred at 70 C over-night. After product precipitation, the mixture was cooled to 0 C for 30 min. and the solid separated off by filtration. This product was redissolved in methanol, which was then removed by rotary evaporation. The product was dried in *vacuo* yielding 201 mg (67%) of product **9**. ^{1}H NMR (400 MHz, $CDCl_3$, ppm): δ = 10.79 (s, 1H; NH), 9.76 (s, 1H; NH), 2.05 (s, 0.31H, residual protonation; CH_3), 1.17 (s, 0.28H, residual protonation; CH_3) ^{13}C NMR (100 MHz, $CDCl_3$, ppm): δ = 123.90 (^{13}C).

[3-^{13}C; 4,4,4-$^2H_3$] 2-Keto-3-(methyl-$^2H_3$)-butanoic acid (7). 108 mg of the hydantoin derivative **9** were stirred at 100 C in 4 mL of aqueous NaOD 20% (w/w) under an argon atmosphere for 5 hrs. After addition of 10 ml of water, 10 ml of ethyl acetate were added, and the phases separated. The aqueous phase was washed one more time with 10 ml ethyl acetate and the combined organic phases discarded. The aqueous phase was acidified by addition of 1 ml of $HCl_{conc.}$ at 0 C. The resulting solution was extracted five times with 10 ml ethyl acetate and the combined organic phases carefully evaporated via rotary evaporation > 200 mbar. Then, 4 mL of a 1 M aqueous solution of NaOH were added to the flask. The mixture was stirred for 3 hours at room temperature. After that, 10 mL of water and 10 mL of ethyl acetate were added, and the two layers were separated. The organic layer was discarded, and the remaining aqueous layer was acidified with HCl_{conc} at 0 C. This mixture was then extracted five times with 20 mL portions of ethyl acetate. After that, the combined organic phases were dried over magnesium sulfate. Evaporation of the solvents in *vacuo* at > 200 mbar resulted in a pale-yellow oil of α-ketoacid **7** in a yield of 65 mg (72%). Proton NMR spectroscopy indicates a methyl deuteration grade of ~97%. ^{1}H NMR (400 MHz, $CDCl_3$, ppm): δ =

Research Article
doi.org/10.1002/cbic.202300762
Chemistry Europe
European Chemical Societies Publishing

3.44 (d, $^{1}J_{C,H}$ = 131.1 Hz, 1H; ^{13}CH), 1.23 (s, 0.21H, residual protonation of the two methyl groups; CH_3). ^{13}C NMR (100 MHz, $CDCl_3$, ppm): δ = 34.84 (^{13}C).

Expression and purification of Brd4-BD1. Protein overexpression was performed as described previously.[35] Recombinant human Brd4-BD1 was expressed in *E. coli* BL21(DE3), which contains a N-terminal TEV-cleavable His6-Tag (plasmid was kindly provided by Boehringer Ingelheim GmbH & Co KG). Uniformly ^{15}N and selective Leu/Val ^{13}C/^{2}H labeled Brd4-BD1 (wild-type, L92A, L94A) was expressed following the expression protocol for efficient isotopic labeling of recombinant proteins using a fourfold cell concentration in M9 minimal medium, supplemented with 1 g/L $^{15}NH_4Cl$, 3 g/L D-glucose and 100 mg/L [3-^{13}C; 4,4,4-$^{2}H_3$] 2-keto-3-(methyl-$^{2}H_3$)-butanoic acid **7**. Uniformly ^{15}N and ^{13}C-labeled Brd4-BD1 for triple resonance signal assignment was expressed following the same protocol, M9 minimal medium was supplemented with 1 g/L $^{15}NH_4Cl$ and 3 g/L $^{13}C_6$-D-glucose. Cells were grown until an OD_{600nm} of 0.7 and protein expression was induced with 0.4 mM IPTG final concentration. Cells were harvested by centrifugation, lysed by sonication and the lysates were subsequently centrifuged. Proteins were purified from the supernatant by Ni^{2+} affinity chromatography. The purified protein was treated with TEV protease and again loaded onto a Ni^{2+} column to bind the cleaved His6-Tag and the His6-Tagged TEV protease. The flow-through containing Brd4-BD1 was concentrated. Purity was analyzed by SDS-PAGE. NMR samples of Brd4-BD1 were prepared in 10 mM sodium phosphate buffer containing 0.1 mM protein, 100 mM sodium chloride, 10% D_2O and 1 mM DTT.

NMR experiments. All protein NMR experiments were conducted at 298 K on a Bruker Avance 600 MHz or Bruker Avance 700 MHz spectrometer equipped with a TCI cryoprobe with Brd4-BD1 sample concentration of 100 μM. Ligand concentrations were kept at 2 mM. 2D ^{1}H-^{13}C-SOFAST-HMQC spectra were acquired using the pulse sequence "Me_sfhmqcf2qpph.2" of the Bruker library. Spectra were recorded using 128 (t1)×2048 (t2) complex points with acquisition times of 106 ms (t_2) and 14.1 ms (t_1). A total of 48 scans were recorded per t_1 increment. NMR spectra were processed and analyzed with NMRPipe,[52] SPARKY[53] and CCPNmr.[54] Unambiguous assignment of the Valine Cβ residues were performed with a series of three-dimensional NMR experiments (HNCACB, HN(CO)CACB, hCC-TOCSY-coNH and Hcc-TOCSY-coNH).

Acknowledgements

G. Toscano, T. Höfurthner and A. Beier were funded by the Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Austria. The financial support by the Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association is gratefully acknowledged.

Conflict of Interests

R. Lichtenecker and R. Konrat are shareholders in the company Mag-Lab, which is mentioned in the affiliations.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: CH–Pi interaction · methine isotope labeling · NMR spectroscopy · protein ligand interaction · protein modifications

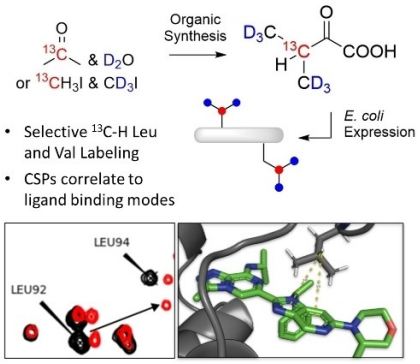
[1] D. C. Rees, M. Congreve, C. W. Murray, R. Carr, *Nat. Rev. Drug Discovery* **2004**, *3(8)*, 660–672.
[2] L. G. Mureddu, G. W. Vuister, *Front. Mol. Biosci.* **2022**, *18(9)*, 834453; DOI: 10.3389/fmolb.2022.834453.
[3] M. J. Harner, A. O. Frank, S. W. Fesik, *J. Biomol. NMR* **2013**, *56(2)*, 65–75.
[4] C. Di Carluccio, M. C. Forgione, S. Martini, F. Berti, A. Molinaro, R. Marchetti, A. Silipo, *Carbohydr. Res.* **2021**, *503*, 108313.
[5] C. Aguirre, O. Cala, I. Krimm, *Curr. Protoc. Protein Sci.* **2015**, 17.18.1–17.18.24. ■■■ Dear Author, if the journal has volumes, please add the journal number■■■
[6] W. Becker, K. C. Bhattiprolu, N. Gubensäk, K. Zangger, *ChemPhysChem* **2018**, *19(8)*, 895–906.
[7] D. M. LeMaster, *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26(4)*, 371–419.
[8] S. Ohki, M. Kainosho, *Prog. Nucl. Magn. Reson. Spectrosc.* **2008**, *53(4)*, 208–226.
[9] B. Rowlinson, E. Crublet, R. Kerfah, M. J. Plevin, *Biochem. Soc. Trans.* **2022**, *50(6)*, 1555–1567.
[10] R. J. Lichtenecker, J. Schörghuber, M. Bisaccia, *Synlett* **2015**, *26*, 2611–2616.
[11] J. Schörghuber, T. Sara, M. Bisacchia, W. Schmid, R. Konrat, R. J. Lichtenecker, *ChemBioChem* **2015**, *16(5)*, 746–751.
[12] J. Schörghuber, L. Geist, G. Platzer, R. Konrat, R. J. Lichtenecker, *ChemBioChem* **2017**, *18*, 1487–1491.
[13] J. Schörghuber, L. Geist, M. Bisaccia, F. Weber, R. Konrat, R. J. Lichtenecker, *J. Biomol. NMR* **2017**, *69*, 13–22.
[14] J. Schörghuber, L. Geist, G. Platzer, M. Feichtinger, M. Bisaccia, L. Scheibelberger, F. Weber, R. Konrat, R. J. Lichtenecker, *J. Biomol. NMR* **2018**, *71*, 129–140.
[15] M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. M. Ono, P. Güntert, *Nature* **2006**, *440*, 52–57.
[16] P. Lundström, K. Teilum, T. Carstensen, I. Bezsonova, S. Wiesner, D. Flemming Hansen, T. L. Religa, M. Akke, L. E. Kay, *J. Biomol. NMR* **2007**, *38*, 199–212.
[17] A. J. Wand, R. J. Bieber, J. L. Urbauer, R. P. McEvoy, Z. H. Gan, *J. Magn. Reson. Ser. B* **1995**, *108*, 173–175.
[18] N. K. Goto, K. H. Gardner, G. A. Mueller, R. C. Willis, L. E. Kay, *J. Biomol. NMR* **1999**, *13(4)*, 369–374.
[19] V. Tugarinov, V. Kanelis, L. E. Kay, *Nat. Protoc.* **2006**, *1(2)*, 749–754.
[20] R. Lichtenecker, M. L. Ludwiczek, W. Schmid, R. Konrat, *J. Am. Chem. Soc.* **2004**, *126(17)*, 5348–5349.
[21] R. J. Lichtenecker, N. Coudevylle, R. Konrat, W. Schmid, *ChemBioChem* **2013**, *14*, 818–821.
[22] R. J. Lichtenecker, K. Weinhäupl, L. Reuther, J. Schörghuber, W. Schmid, R. Konrat, *J. Biomol. NMR* **2013**, *57(3)*, 205–209.
[23] P. Gans, O. Hamelin, R. Sounier, I. Ayala, M. A. Durà, C. D. Amero, M. Noirclerc-Savoye, B. Franzetti, M. J. Plevin, J. Boisbouvier, *Angew. Chem. Int. Ed.* **2010**, *49(11)*, 1958–1962.
[24] G. Mas, E. Crublet, O. Hamelin, P. Gans, J. Boisbouvier, *J. Biomol. NMR* **2013**, *57*, 251–262.
[25] Y. R. Monneau, Y. Ishida, P. Rossi, T. Saio, S.-R. Tzeng, M. Inouye, C. G. Kalodimos, *J. Biomol. NMR* **2016**, *65(2)*, 99–108.
[26] E. R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-Garcia, A. J. Cohen, W. Yang, *J. Am. Chem. Soc.* **2010**, *132(18)*, 6498–6506.
[27] R. Ferreira de Freitas, M. Schapira, *MedChemComm* **2017**, *8(10)*, 1970–1981.
[28] T. Steiner, *Chem. Commun.* **1997**, 727–734. ■■■ Dear Author, if the journal has volumes, please add the journal number■■■
[29] S. Sarkhel, G. R. Desiraju, *Proteins Struct. Funct. Bioinf.* **2004**, *54*, 247–259.
[30] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, *J. Am. Chem. Soc.* **2000**, *122(15)*, 3746–3753.
[31] L. M. Salonen, M. Ellermann, F. Diederich, *Angew. Chem. Int. Ed.* **2011**, *50(21)*, 4808–4842.
[32] M. Ozawa, T. Ozawa, M. Nishio, K. Ueda, *J. Mol. Graphics Modell.* **2017**, *75*, 117–124.

[33] D. B. McConnell, *J. Med. Chem.* **2021**, *64(22)*, 16319–16327.
[34] M. J. Plevin, D. L. Bryce, J. Boisbouvier, *Nat. Chem.* **2010**, *2(6)*, 466–471.
[35] G. Platzer, M. Mayer, A. Beier, S. Brüschweiler, J. E. Fuchs, H. Engelhardt, L. Geist, G. Bader, J. Schörghuber, R. Lichtenecker, B. Wolkerstorfer, D. Kessler, D. B. McConnell, R. Konrat, *Angew. Chem. Int. Ed.* **2020**, *59(35)*, 14861–14868.
[36] T. Höfurthner, G. Toscano, G. Kontaxis, A. Beier, M. Mayer, L. Geist, D. B. McConnell, H. Weinstabl, R. Lichtenecker, R. Konrat, *J. Biomol. NMR* **2023**, DOI: 10.1007/s10858-023-00427-2.
[37] M. H. Al Mughram, C. Catalano, N. B. Herrington, M. K. Safo, G. E. Kellog, *Front. Mol. Biosci.* **2023**, *10*, DOI: 10.3389/fmolb.2023.1116868.
[38] F. Desantis, M. Miotto, L. Di Rienzo, E. Milanetti, G. Ruocco, *Sci. Rep.* **2022**, *12*, 12087.
[39] P. D. Leeson, A. P. Bento, A. Gaulton, A. Hersey, E. J. Manners, C. J. Radoux, A. R. Leach, *J. Med. Chem.* **2021**, *64(11)*, 7210–7230.
[40] L. L. Kiessling, R. C. Diehl, *ACS Chem. Biol.* **2021**, *16(10)*, 1884–1893.
[41] M. Brandl, M. S. Weiss, A. Jabs, J. Sühnel, R. Hilgenfeld, *J. Mol. Biol.* **2001**, 357–377. ■■■ Dear Author, if the journal has volumes, please add the journal number■■■
[42] T. M. Werkhoven, R. van Nispen, J. Lugtenburg, *Eur. J. Org. Chem.* **1999**, *11*, 2909–2914.
[43] P. J. Hajduk, D. J. Augeri, J. Mack, R. Mendoza, J. Yang, S. F. Betz, S. W. Fesik, *J. Am. Chem. Soc.* **2000**, *122(33)*, 7898–7904.
[44] L. Konnert, F. Lamaty, J. Martinez, E. Colacino, *Chem. Rev.* **2017**, *117(23)*, 13757–13809.
[45] Y. Duan, Y. Guan, W. Qin, X. Zhai, B. Yu, H. Liu, *MedChemComm* **2018**, *9(11)*, 1779–1802.
[46] M. E. White, J. M. Fenger, W. E. Carson III, *Cell. Immunol.* **2019**, *337*, 48–53.
[47] H. Engelhardt, D. Gianni, C. Smethurst (Boehringer Ingelheim International GmbH), WO2014076237 A1, **2013**.
[48] J. A. Pople, *J. Chem. Phys.* **1956**, *24*, 1111.
[49] C. E. Johnson, F. A. Bovey, *J. Chem. Phys.* **1958**, *29*, 1012–1014.
[50] C. W. Haigh, R. B. Mallion, *Prog. Nucl. Magn. Reson. Spectrosc.* **1979**, *13*, 303–344.
[51] A. B. Sahakyan, M. Vendruscolo, *J. Phys. Chem. B* **2013**, *117*, 1989–1998.
[52] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*, 277–293.
[53] T. D. Goddard, D. G. Kneller **2008** SPARKY 3. University of California, San Francisco.
[54] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins* **2005**, *59*, 687–696.

Manuscript received: November 6, 2023
Revised manuscript received: January 26, 2024
Accepted manuscript online: January 31, 2024
Version of record online: ■■, ■■

RESEARCH ARTICLE

A novel labeling approach introduces isolated ^{13}C–H spin systems to the methyl-deuterated side chains of leucine and valine residues. Ligands with known binding geometries have been analyzed in combination with corresponding protein samples, showing that chemical shift perturbation data correlate well with the distance- and angle parameters of CH-π hydrogen bonds.



G. Toscano, T. Höfurthner, B. Nagl, A. Beier, Dr. M. Mayer, Dr. L. Geist, Dr. D. B. McConnell, Dr. H. Weinstabl, Prof. Dr. R. Konrat, Dr. R. J. Lichtenecker**

1 – 9

^{13}Cβ-Valine and ^{13}Cγ-Leucine Methine Labeling To Probe Protein Ligand Interaction.

SPACE RESERVED FOR IMAGE AND LINK

Share your work on social media! *ChemBioChem* has added Twitter as a means to promote your article. Twitter is an online microblogging service that enables its users to send and read short messages and media, known as tweets. Please check the pre-written tweet in the galley proofs for accuracy. If you, your team, or institution have a Twitter account, please include its handle @username. Please use hashtags only for the most important keywords, such as #catalysis, #nanoparticles, or #proteindesign. The ToC picture and a link to your article will be added automatically, so the **tweet text must not exceed 250 characters**. This tweet will be posted on the journal's Twitter account (follow us @ChemBioChem) upon publication of your article in its final form. We recommend you to re-tweet it to alert more researchers about your publication, or to point it out to your institution's social media team.

ORCID (Open Researcher and Contributor ID)

Please check that the ORCID identifiers listed below are correct. We encourage all authors to provide an ORCID identifier for each coauthor. ORCID is a registry that provides researchers with a unique digital identifier. Some funding agencies recommend or even require the inclusion of ORCID IDs in all published articles, and authors should consult their funding agency guidelines for details. Registration is easy and free; for further information, see http://orcid.org/.

Giorgia Toscano
Theresa Höfurthner
Benjamin Nagl
Andreas Beier
Dr. Moriz Mayer
Dr. Leonhard Geist
Dr. Darryl B. McConnell
Dr. Harald Weinstabl
Prof. Dr. Robert Konrat http://orcid.org/0000-0001-6489-4080
Dr. Roman J. Lichtenecker http://orcid.org/0000-0003-2168-5216

^{13}Cβ-Valine and ^{13}Cγ-Leucine Methine Labeling To Probe Protein Ligand Interaction.

Giorgia Toscano[+],[a,c] Theresa Höfurthner[+],[b,c] Benjamin Nagl,[a] Andreas Beier,[b] Moriz Mayer,[d] Leonhard Geist,[d] Darryl B. McConnell,[d] Harald Weinstabl,[d] Robert Konrat*[b,e] and Roman J. Lichtenecker*[a,e]

[a] G. Toscano, B. Nagl, Dr. R.J. Lichtenecker
Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Institute of Organic Chemistry, University of Vienna, Währingerstraße 38, 1090 Vienna, Austria
E-mail: roman.lichtenecker@univie.ac.at, robert.konrat@univie.ac.at

[b] T. Höfurthner, Prof. Dr. R. Konrat, A. Beier
Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Max Perutz Laboratories, Department of Structural and Computational Biology, Campus Vienna Biocenter 5, 1030 Vienna, Austria

[c] G. Toscano, T. Höfurthner
Vienna Doctoral School of Chemistry, University of Vienna, Währingerstr. 38, 1090 Vienna, Austria.

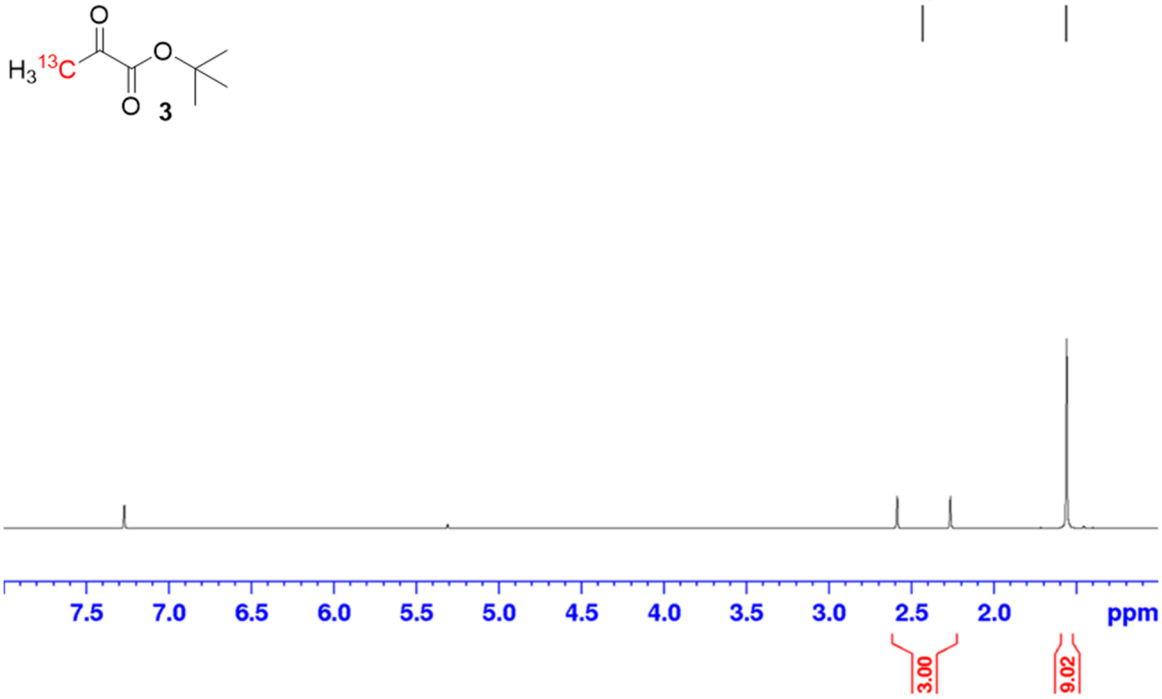
[d] Dr. M. Mayer, Dr. L. Geist, Dr. D.B. McConnell, Dr. H. Weinstabl
Boehringer Ingelheim RCV GmbH & Co. KG, Dr. Boehringer Gasse 5-11, 1121 Vienna

[e] MAG-LAB, Karl-Farkas Gasse 22, 1030 Vienna.

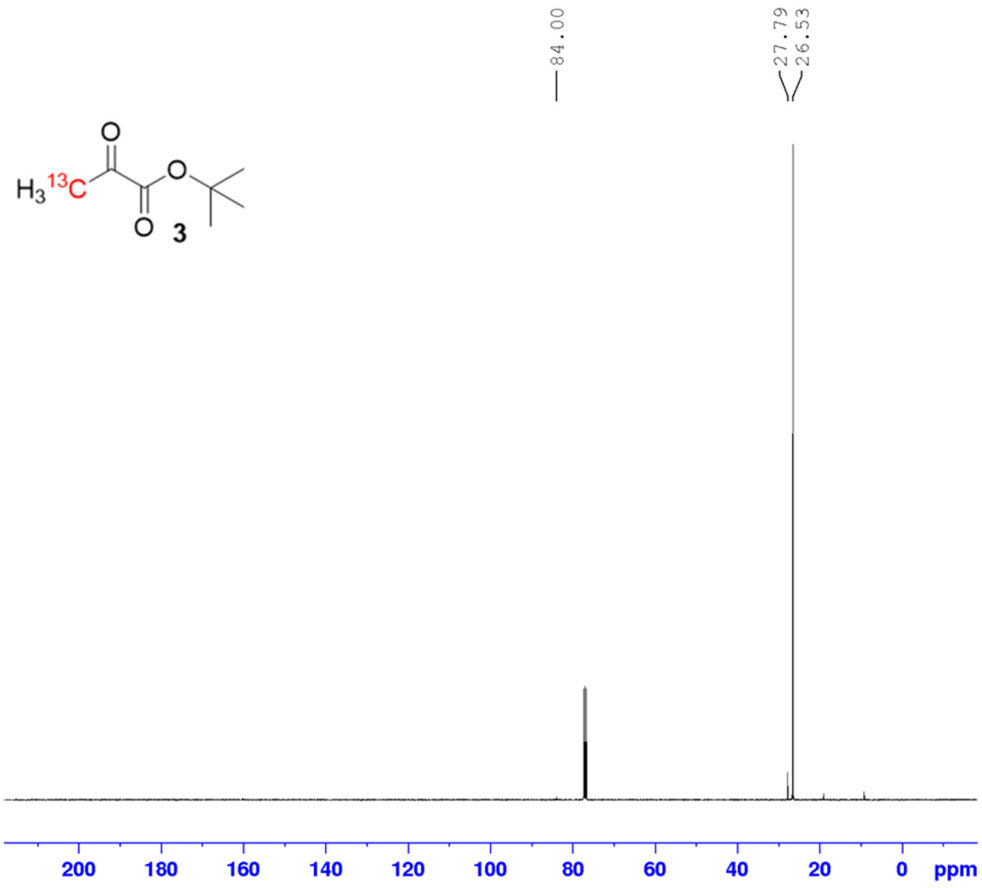
Contents:

1. NMR data of isotope labeled precursors and intermediates

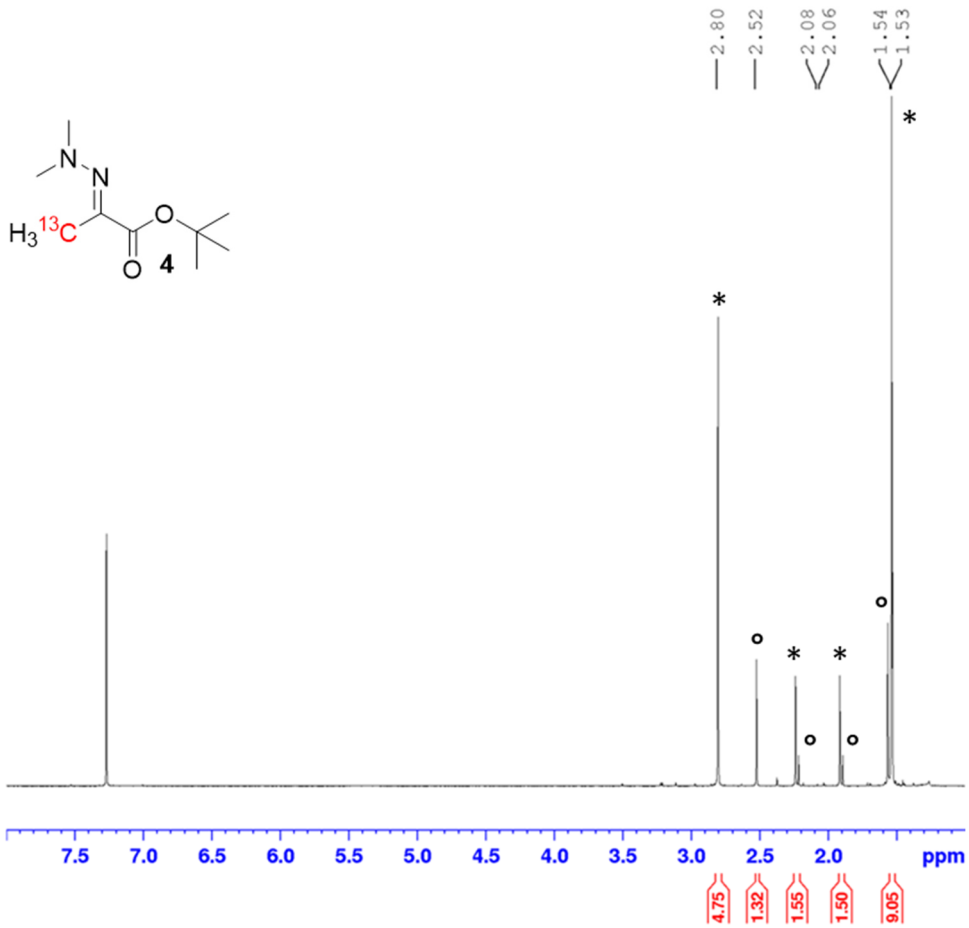
Compound 3, ^{1}H NMR, $CDCl_3$, 400 MHz



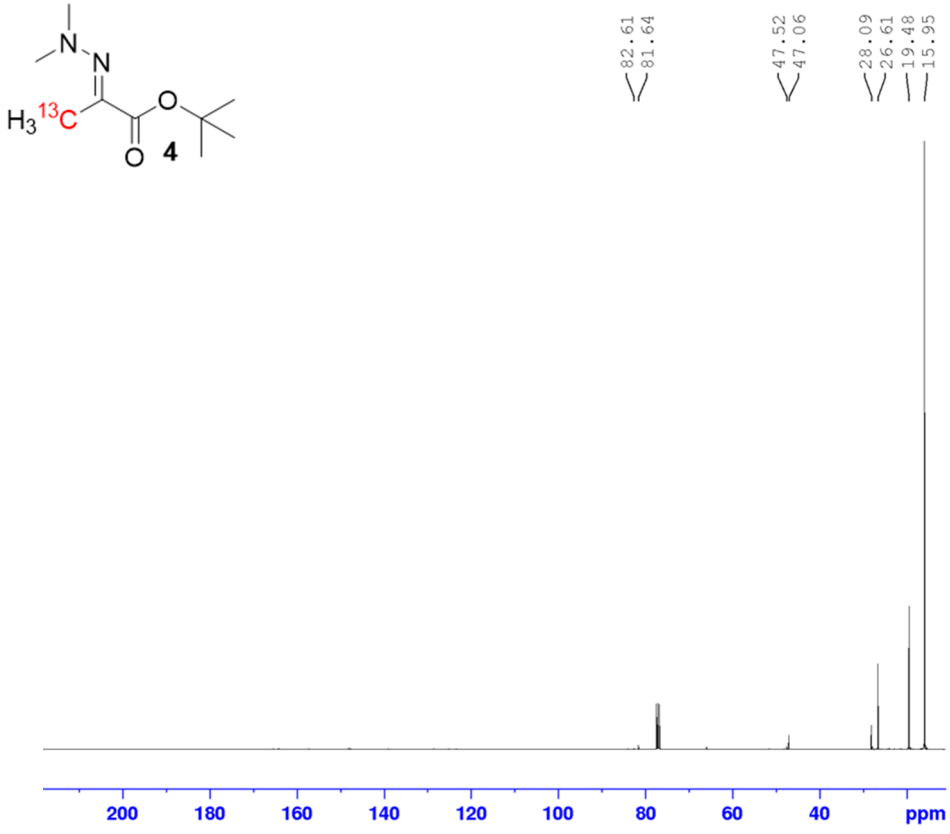
Compound **3**, ^{13}C NMR, $CDCl_3$, 100 MHz



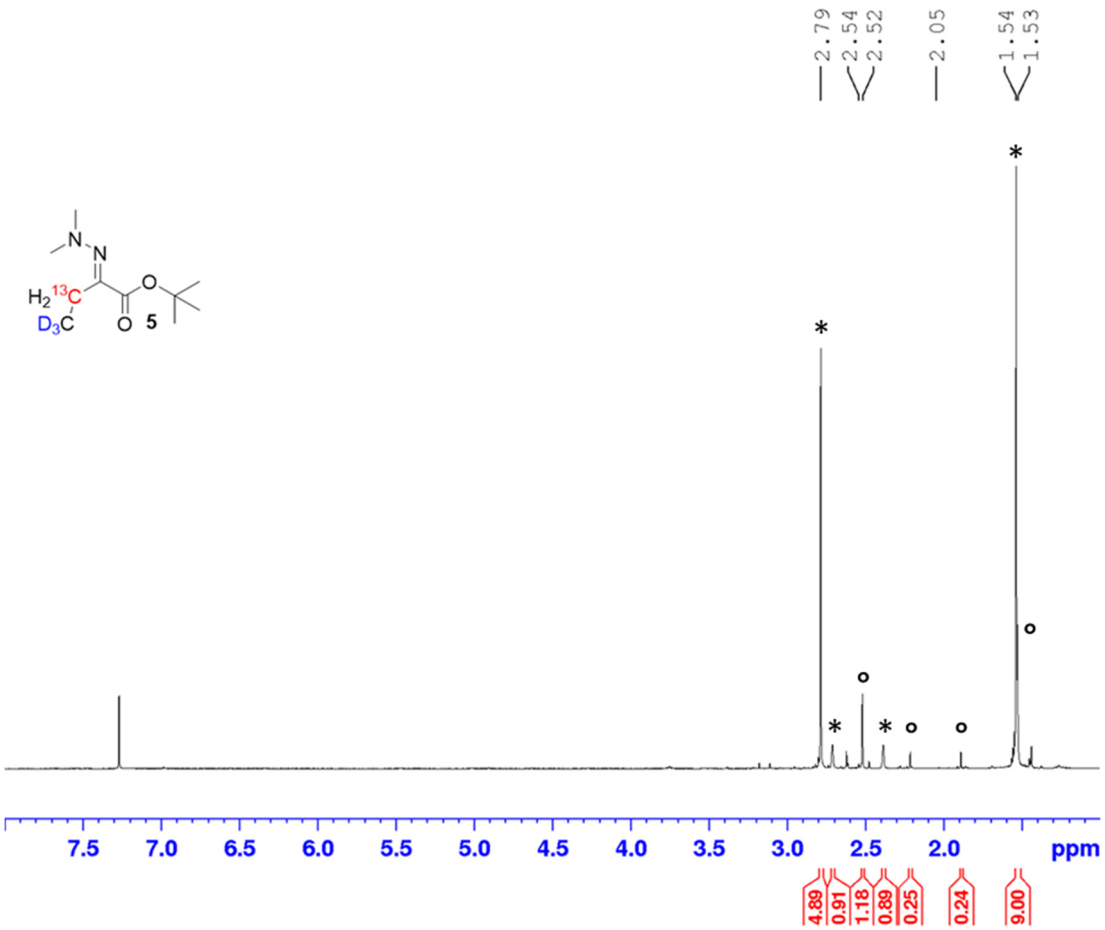
Compound **4**, ^{1}H NMR, $CDCl_3$, 400 MHz, Major isomer (*) and minor isomer (°)



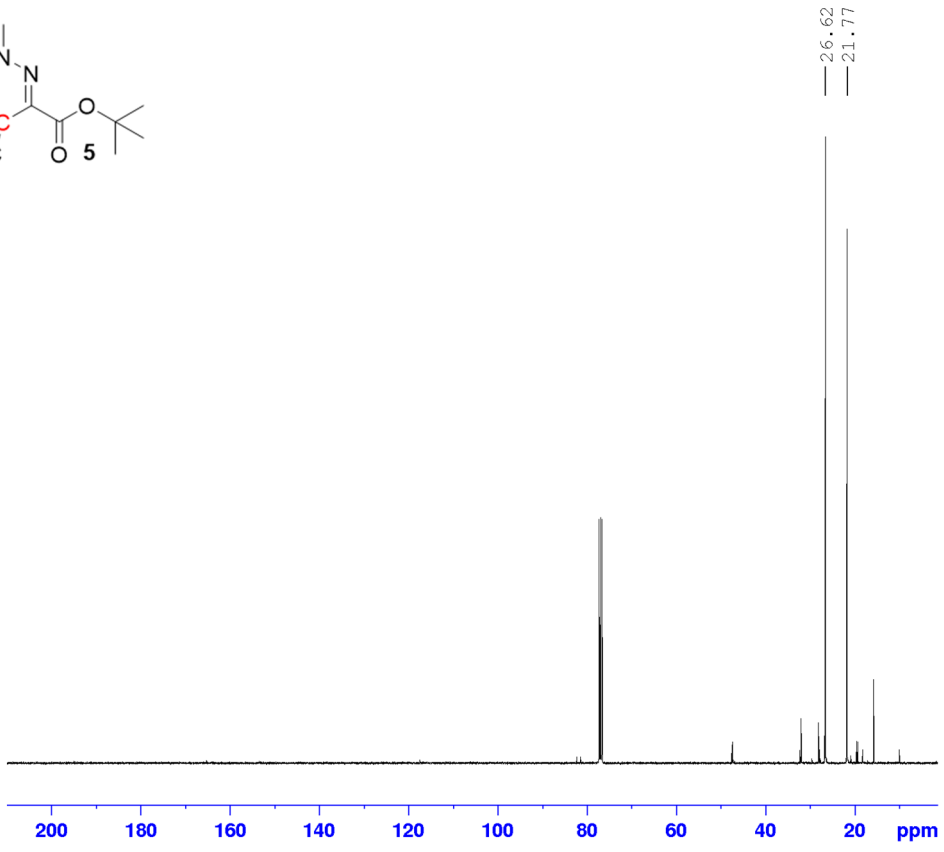
Compound **4**, ^{13}C NMR, $CDCl_3$, 100 MHz



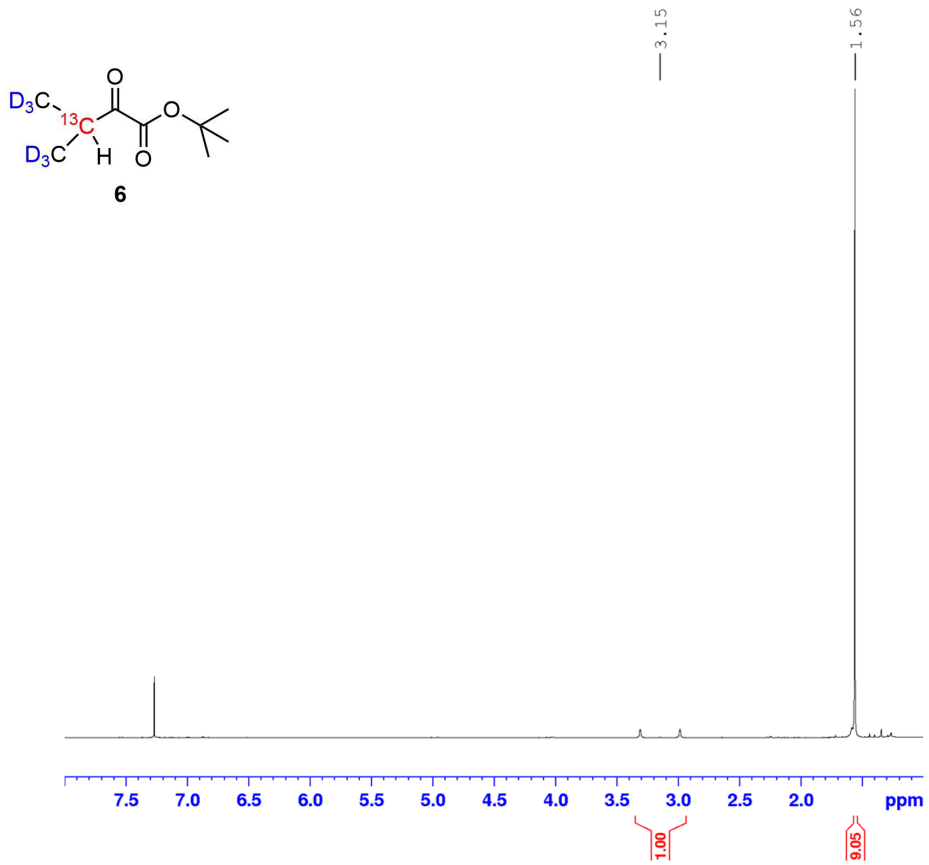
Compound **5**, ^{1}H NMR, $CDCl_3$, 400 MHz, Major isomer (*) and minor isomer (°)



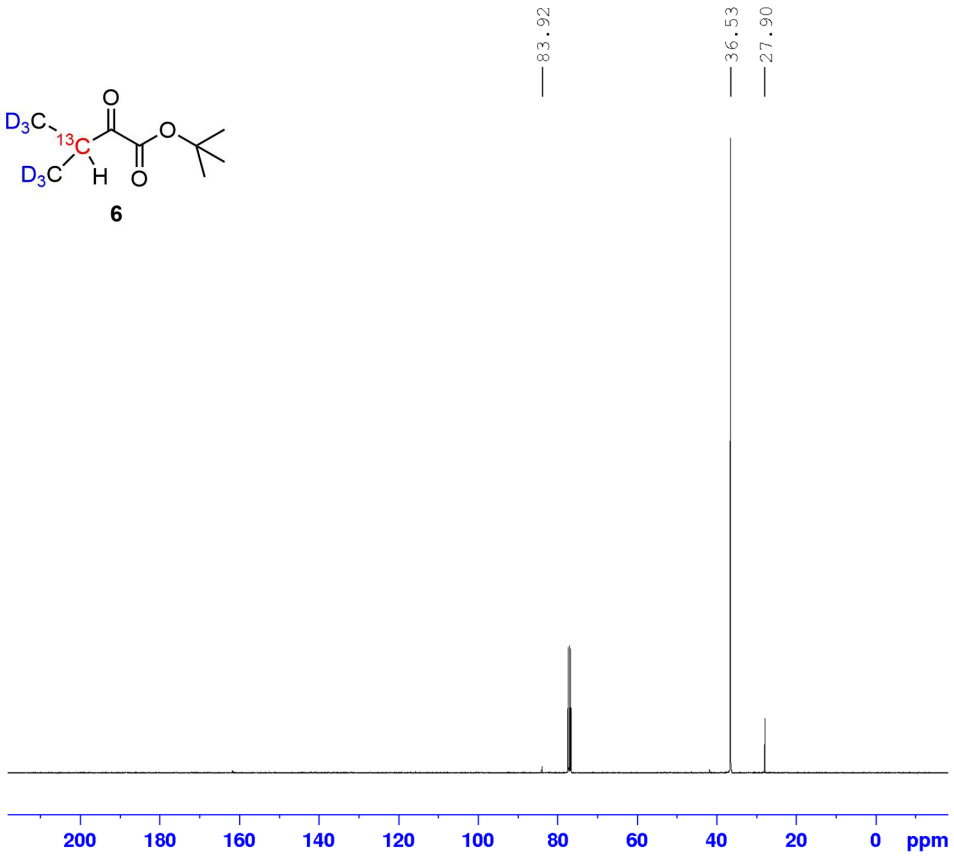
Compound **5**, ^{13}C NMR, $CDCl_3$, 100 MHz



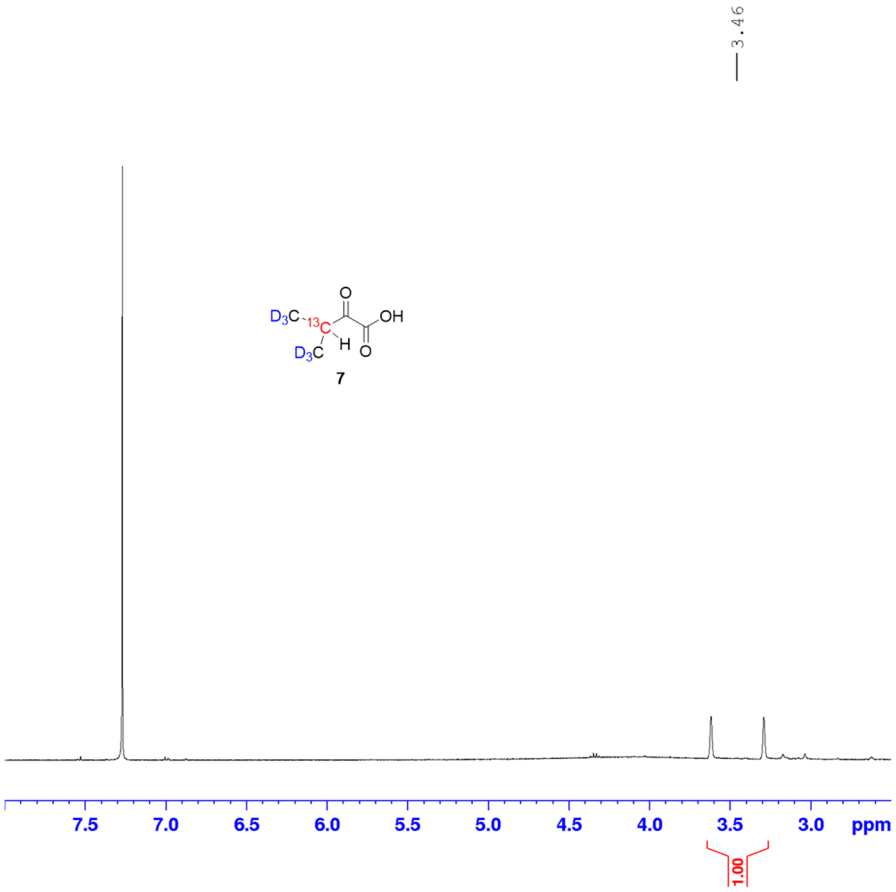
Compound **6**, ^{1}H NMR, $CDCl_3$, 400 MHz



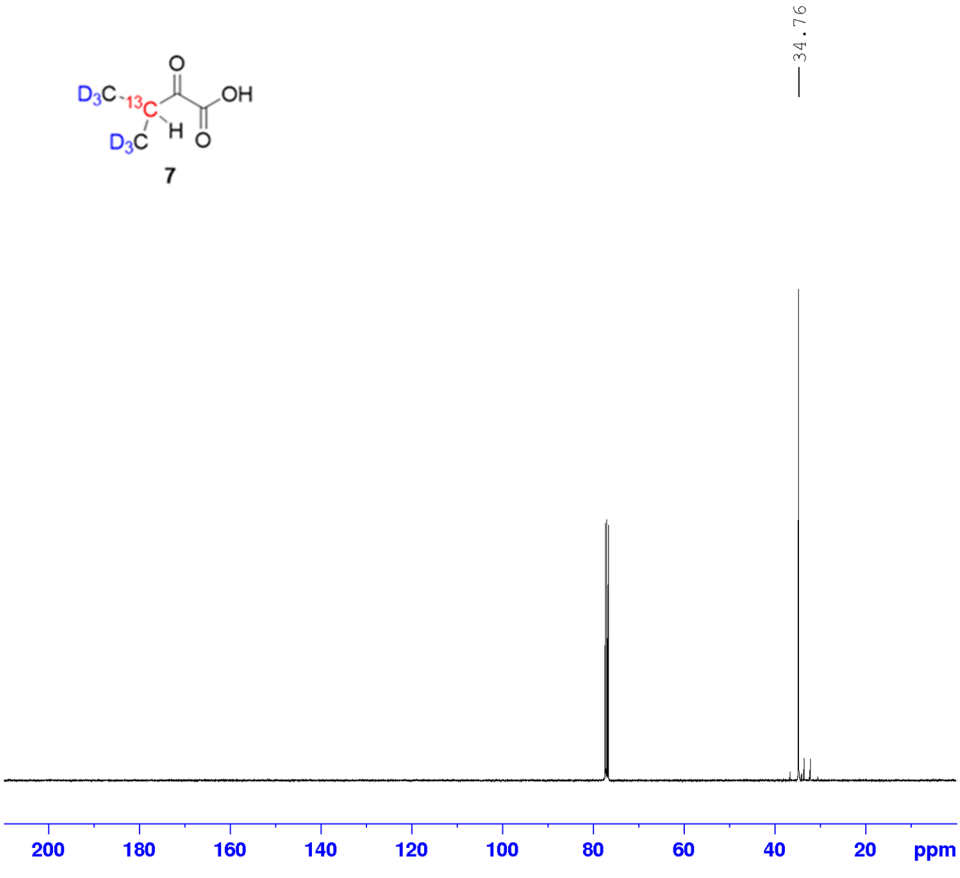
Compound **6**, ^{13}C NMR, $CDCl_3$, 100 MHz



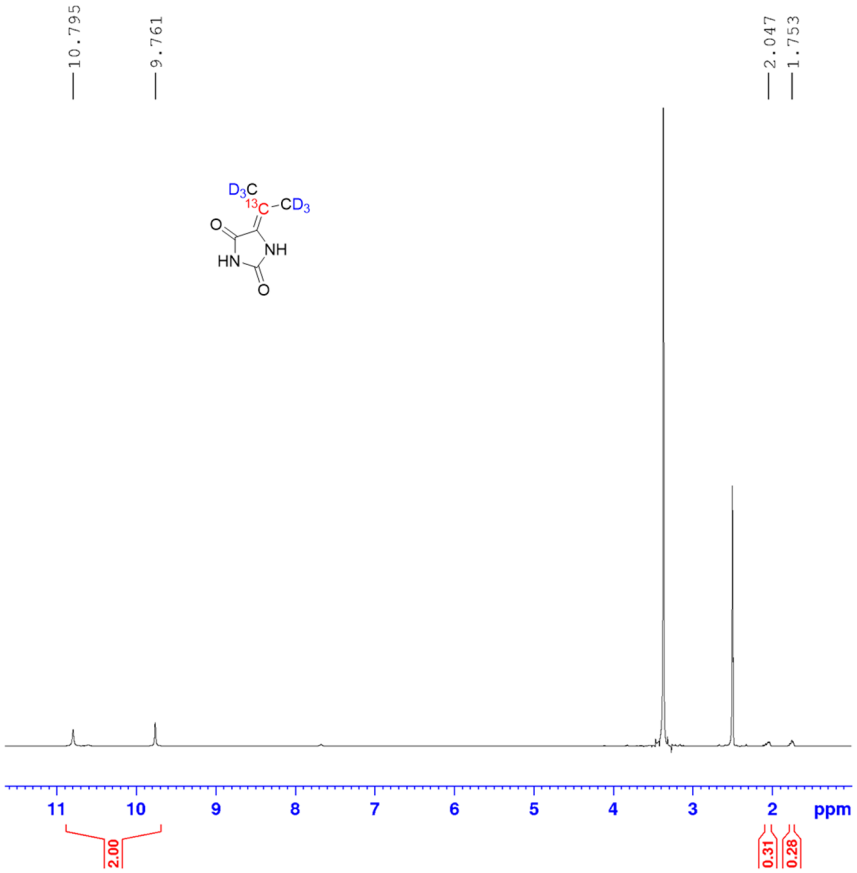
Compound **7**, ^{1}H NMR, $CDCl_3$, 400 MHz



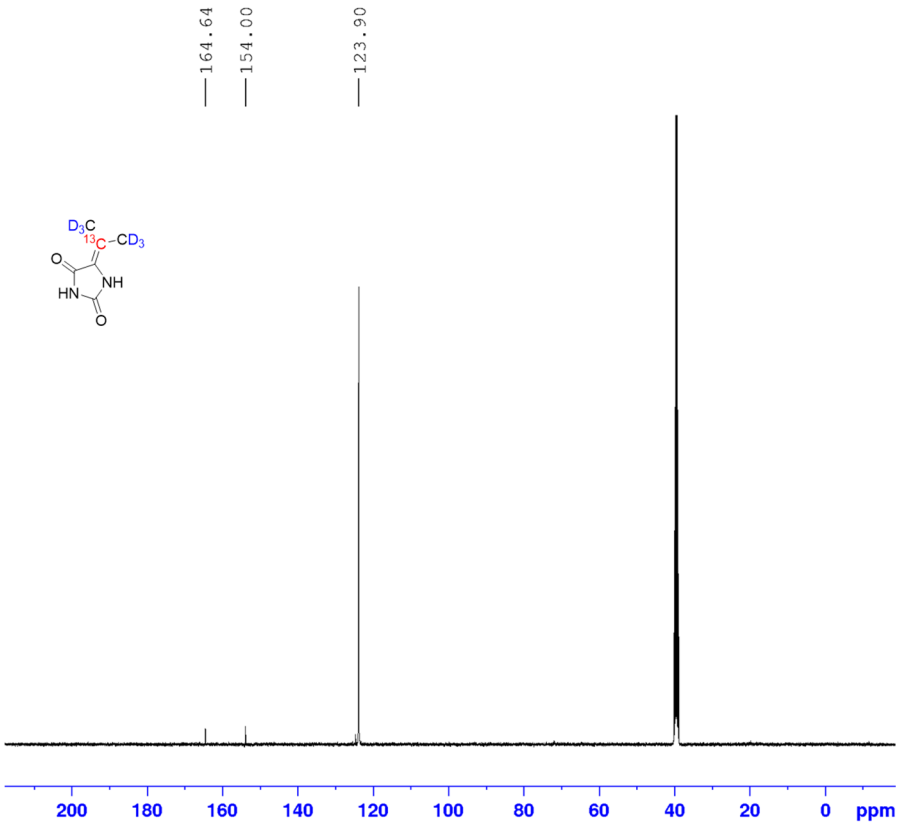
Compound **7**, ^{13}C NMR, $CDCl_3$, 100 MHz



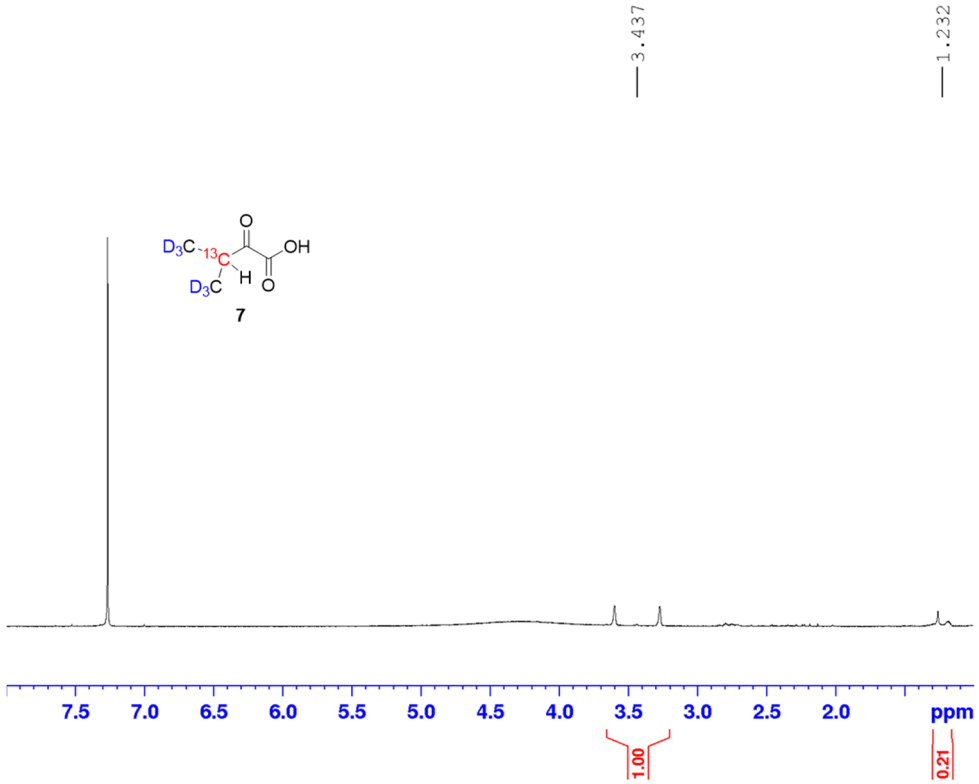
Compound **9,** ^{1}H NMR, 6d-DMSO, 400 MHz



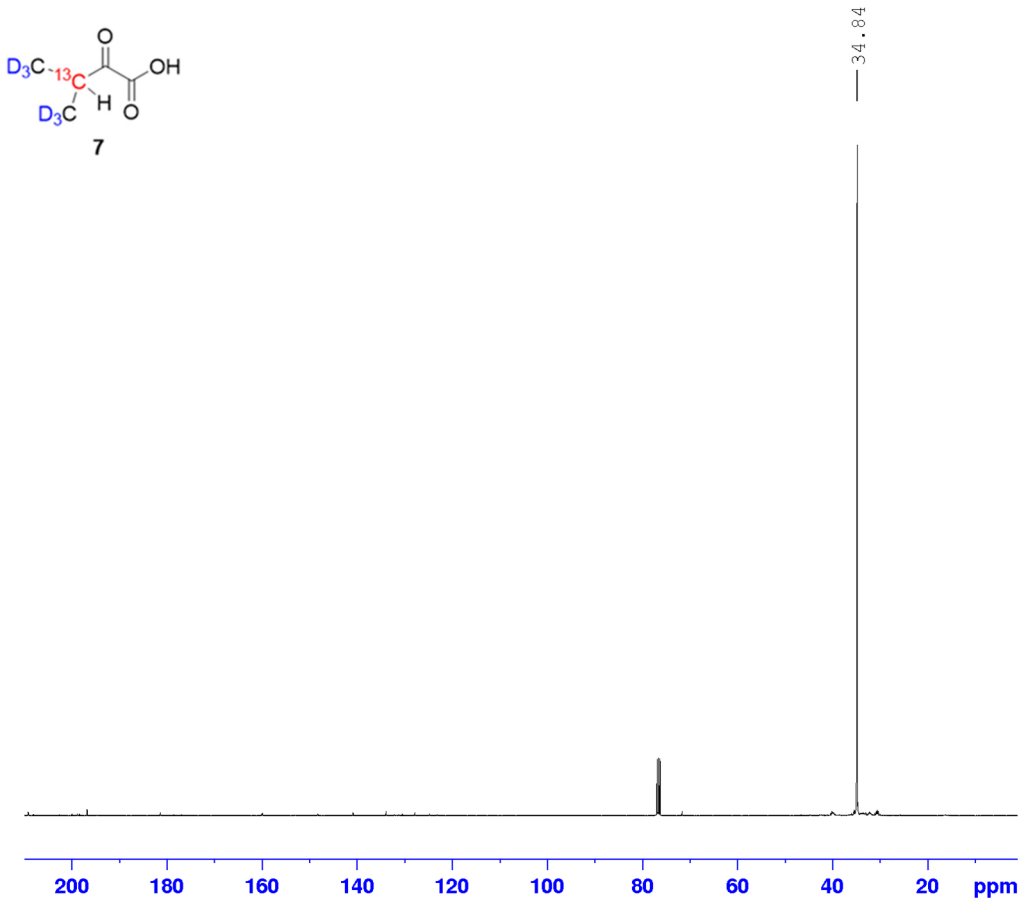
Compound **9**, ^{13}C NMR, 6d-DMSO, 100 MHz



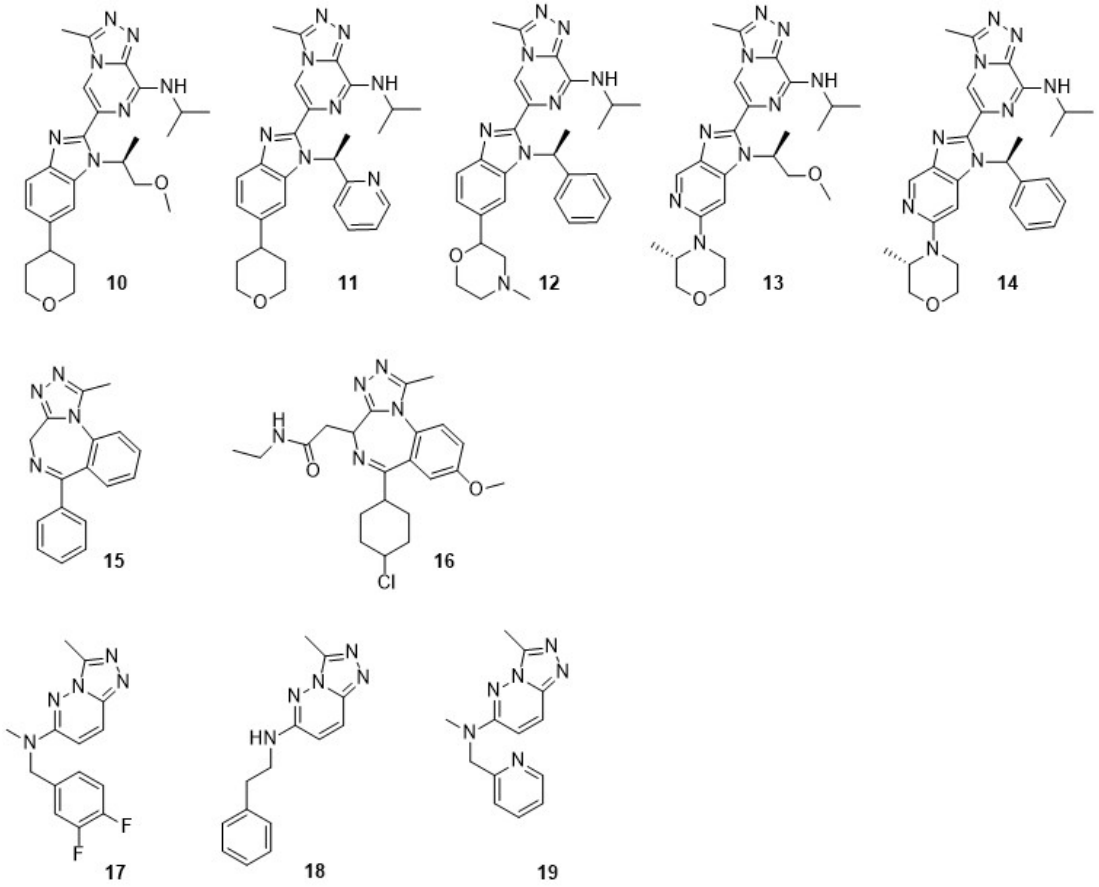
Compound **7** (from [2-^{13}C] acetone **8**), ^{1}H NMR, $CDCl_3$, 400 MHz



Compound **7** (from [2-^{13}C] acetone **8**), ^{13}C NMR, $CDCl_3$, 100 MHz



2. Structures of Brd4-BD1 ligands



3. Table of experimental chemical shifts, geometrical parameters and calculated chemical shifts (Table SI1).

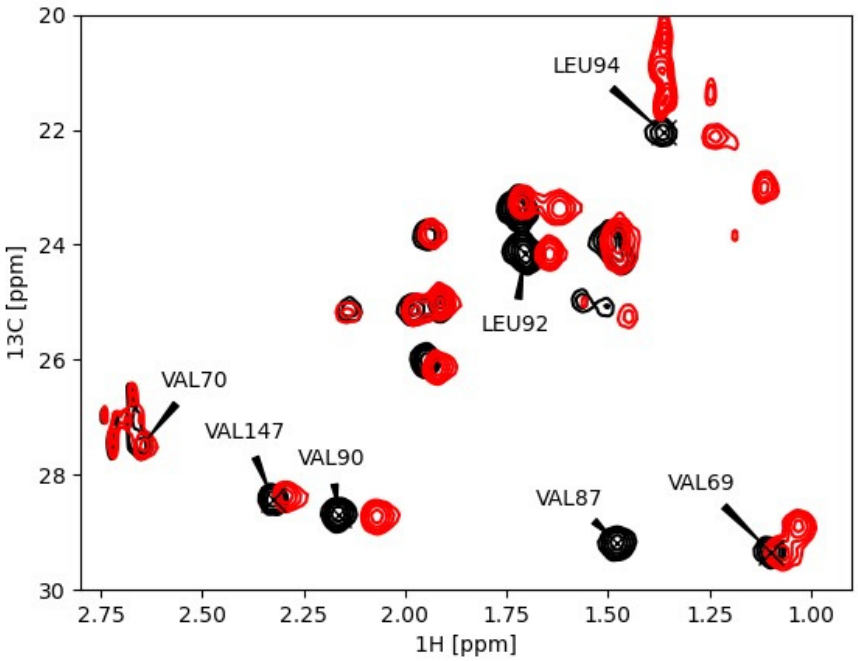
Residue	Ligand Nr	CSP_{exp}	CSP_{pople}	CSP_{exp}-CSP_{pople}	distance [A]	Theta
L92	**10**	0.569	0.4128	0,1562	4.932	0.179
L94	**10**	0.1304	0.0786	0,0518	7.331	0.451
L92	**11**	0.504	0.3955	0,1085	5.000	0.182
L94	**11**	0.1009	0.0808	0,0201	7.171	0.473
L92	**12**	0.503	0.4332	0,0698	4.824	0.208
L94	**12**	0.1036	0.0878	0,0158	7.140	0.431
L92	**13**	0.456	0.3791	0,0769	5.011	0.236
L94	**13**	0.1294	0.0808	0,0486	7.116	0.486
L92	**14**	0.397	0.3875	0,0095	5.005	0.210
L94	**14**	0.0815	0.0691	0,0124	6.782	0.674
L92	**15**	0.247	0.28699	-0,03999	4.9709	0.4816
L94	**15**	0.1076	0.09142	0,01618	7.6060	0.3988
L94	**16**	0.1595	0.1266	0,0329	7.1967	0.2506
L94	**17**	0.1664	0.0729	0,0935	6.526	0.692
L94	**18**	0.1770	0.1361	0,0409	5.7350	0.6197
L92	**19**	0.083	0.05038	0,03262	6.642	0.763
L94	**19**	0.0937	0.0531	0,0406	6.490	0.766

Table SI1: Experimental chemical shift perturbation (CSP_{exp}) of methine ^{13}C-H signals in ^{1}H-^{13}C-SOFAST-HMQC of Brd4-BD1 leucine 92 (L92) and leucine 94 (L94) residues in presence of ligands **10**-**19**. CSP_{Pople} is the chemical shift perturbation calculated from the equation shown in Figure 4 of the main text using r (distance in Ångstroem) and the angle theta in rad. R and theta have been deduced from x-ray structures, either published in the RCSB protein data bank (entry 6XUZ for **13**; 6XV3 for **14**; 6XV7 for **17**), or unpublished (for ligands **10**, **11**, **12**, **15**, **16**, **18** and **19**).

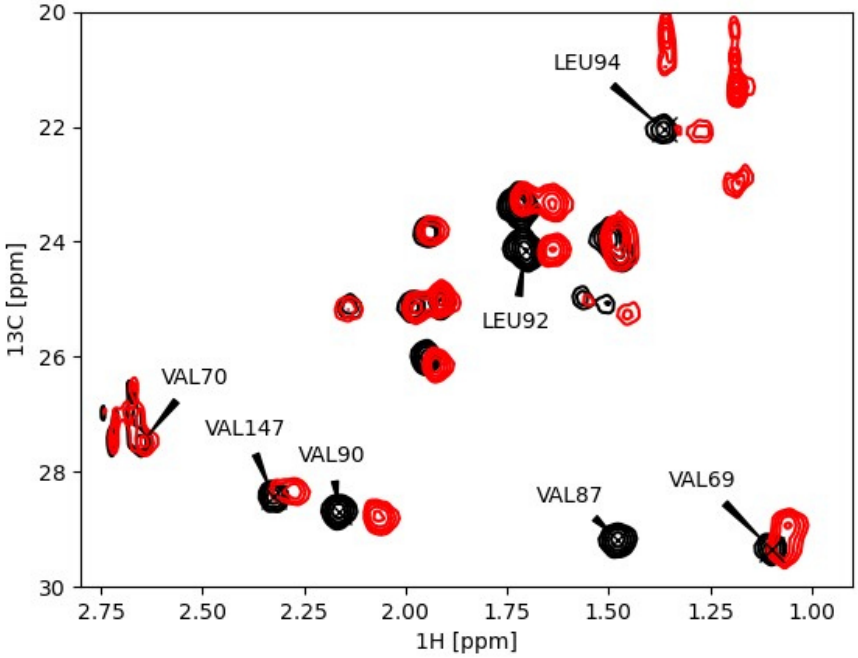
4. ^{1}H-^{13}C-SOFAST-HMQC of Brd4-BD1 in absence (black) and presence (red) of ligands.

Additional signals in red spectra (20-23ppm and 1-1.5ppm) are due to alkyl groups of the ligands, especially the isopropyl groups of ligands **10**-**14**.

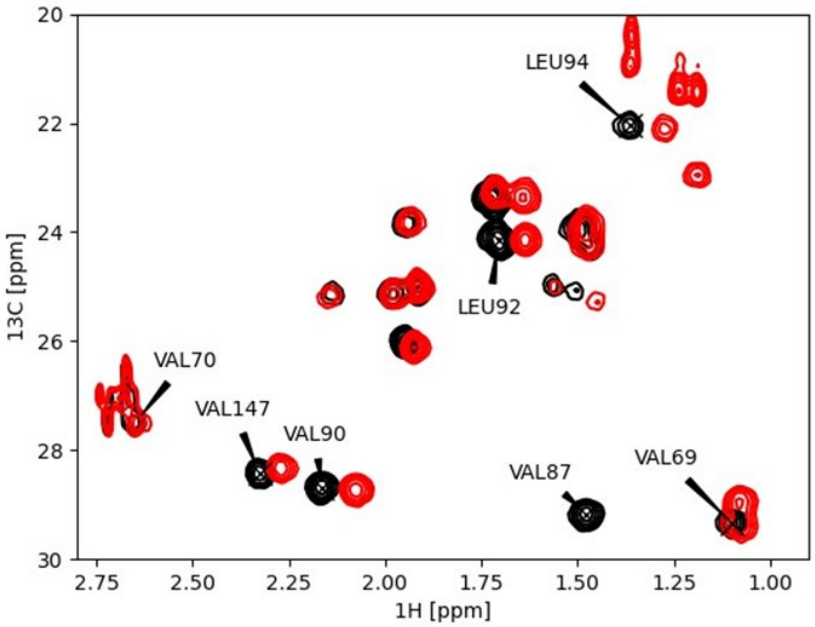
Ligand 10



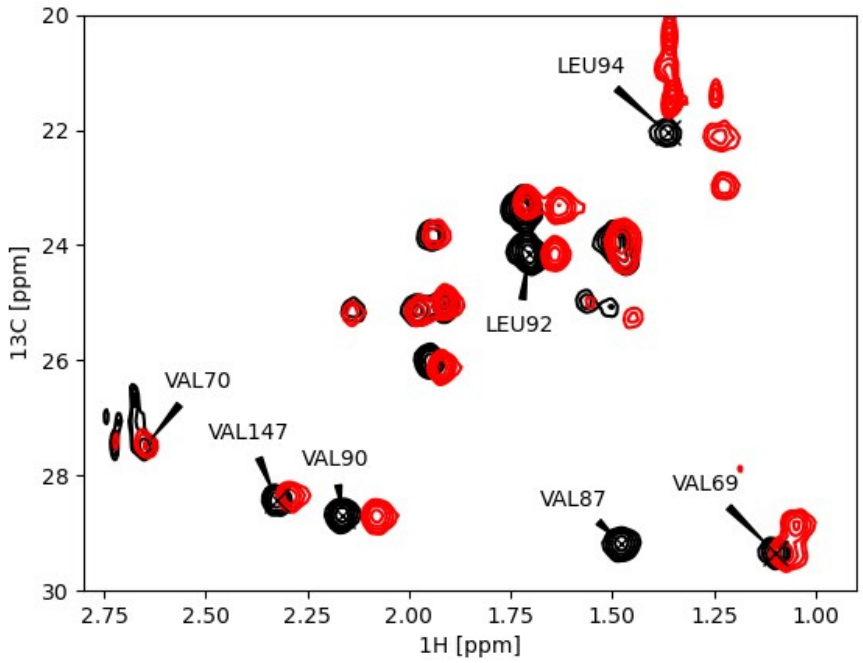
Ligand 11



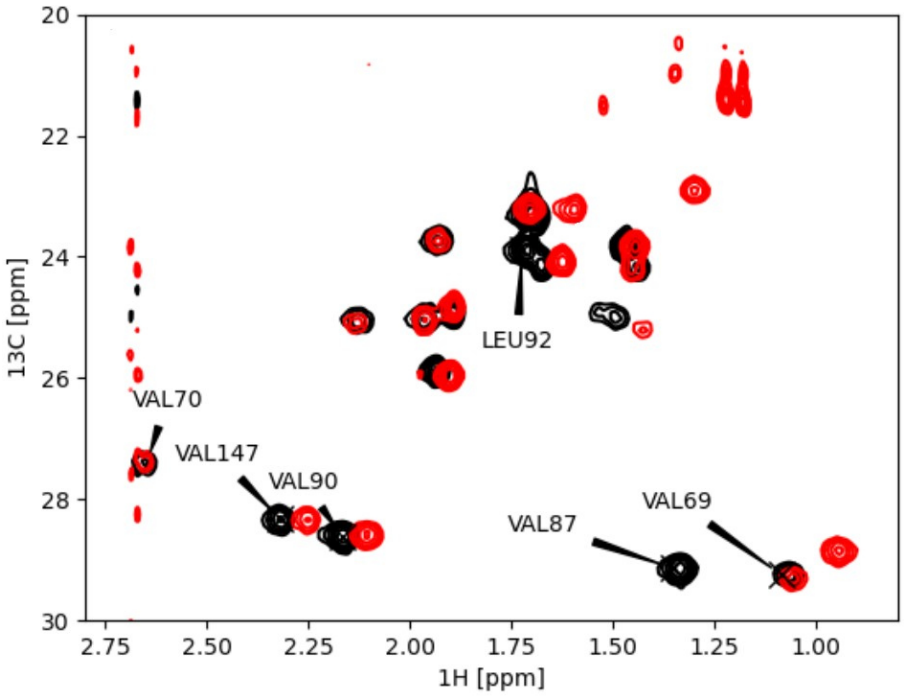
Ligand 12



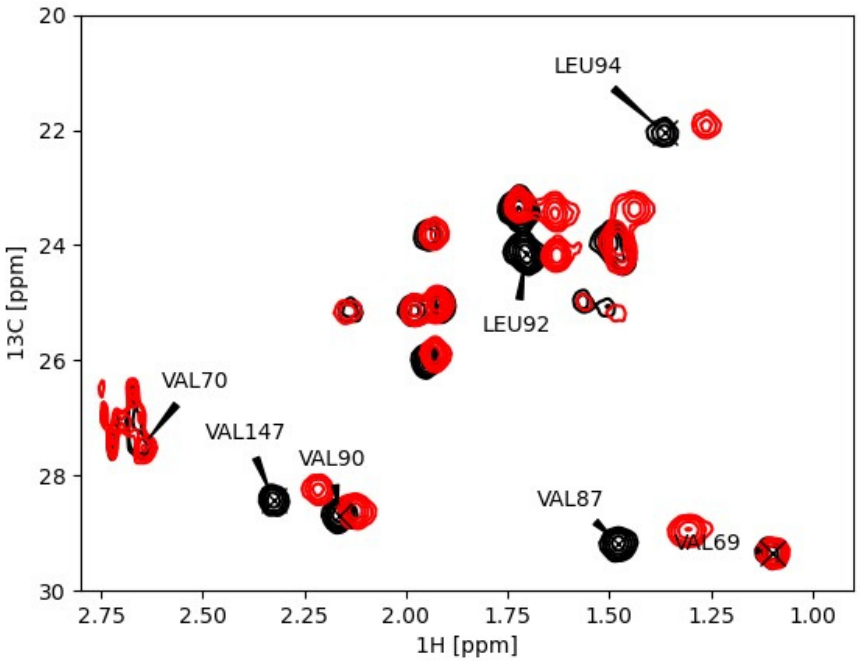
Ligand 13



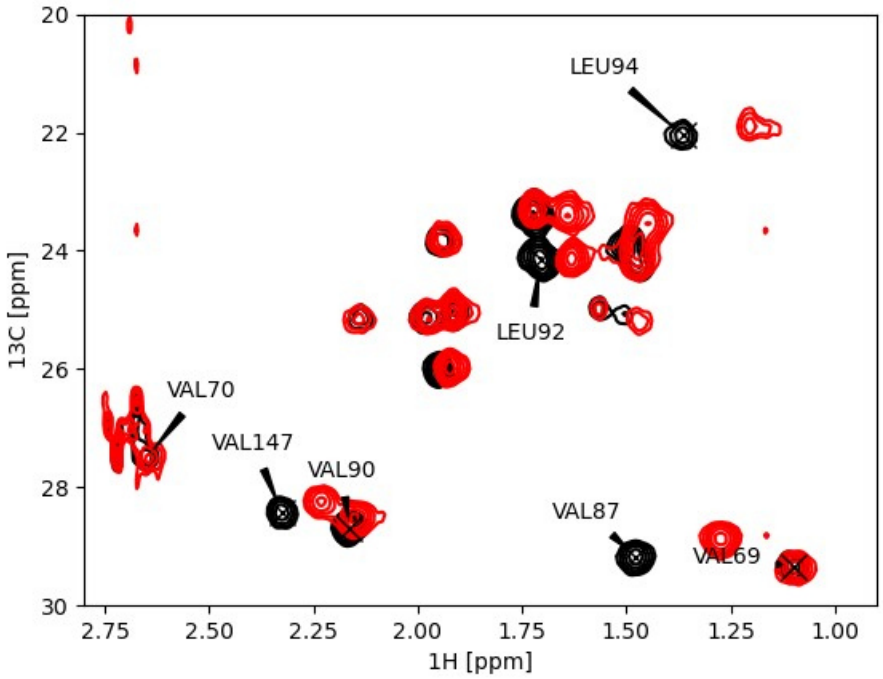
Ligand 14



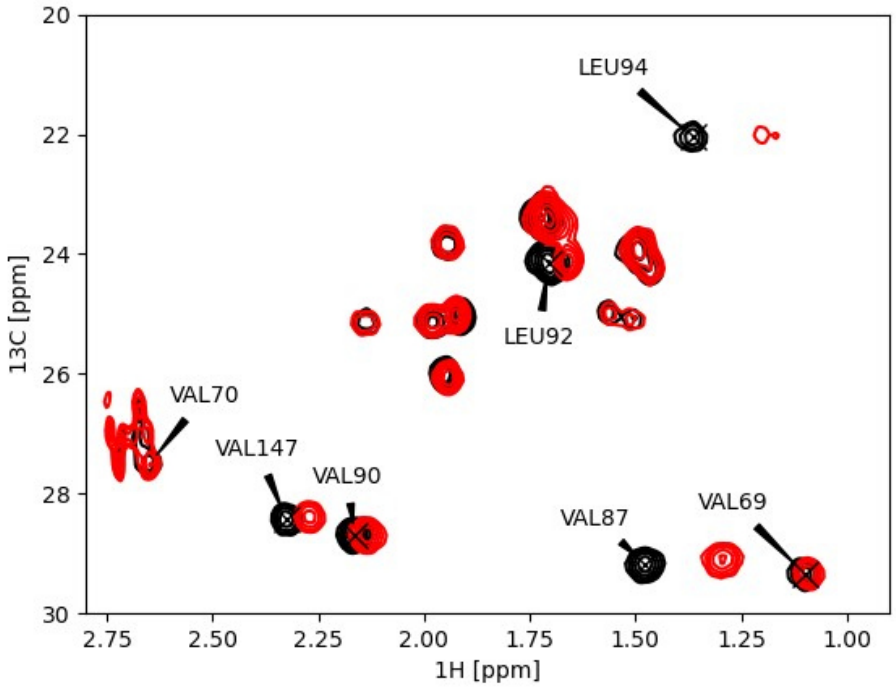
Ligand 15



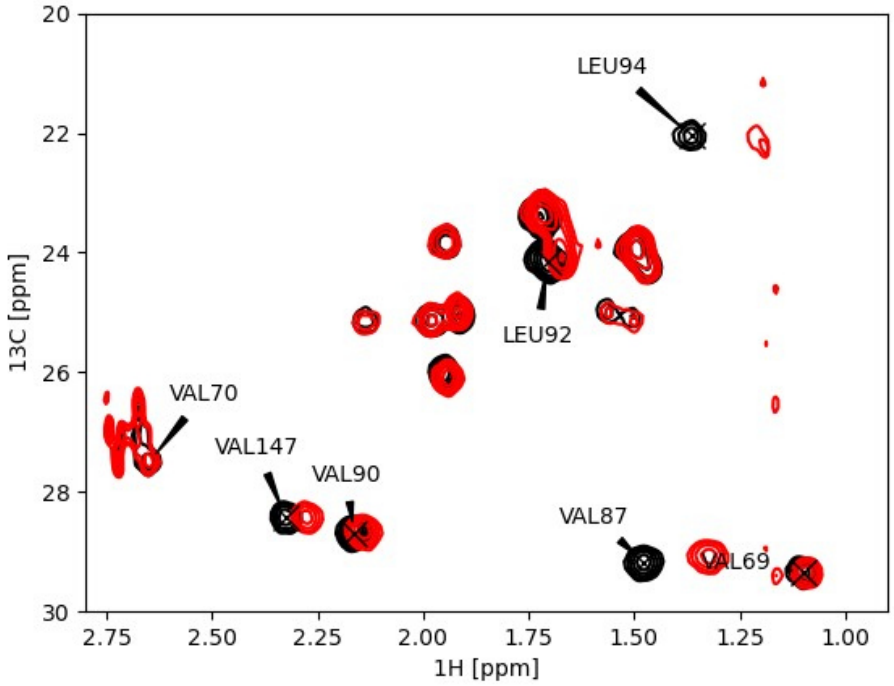
Ligand 16



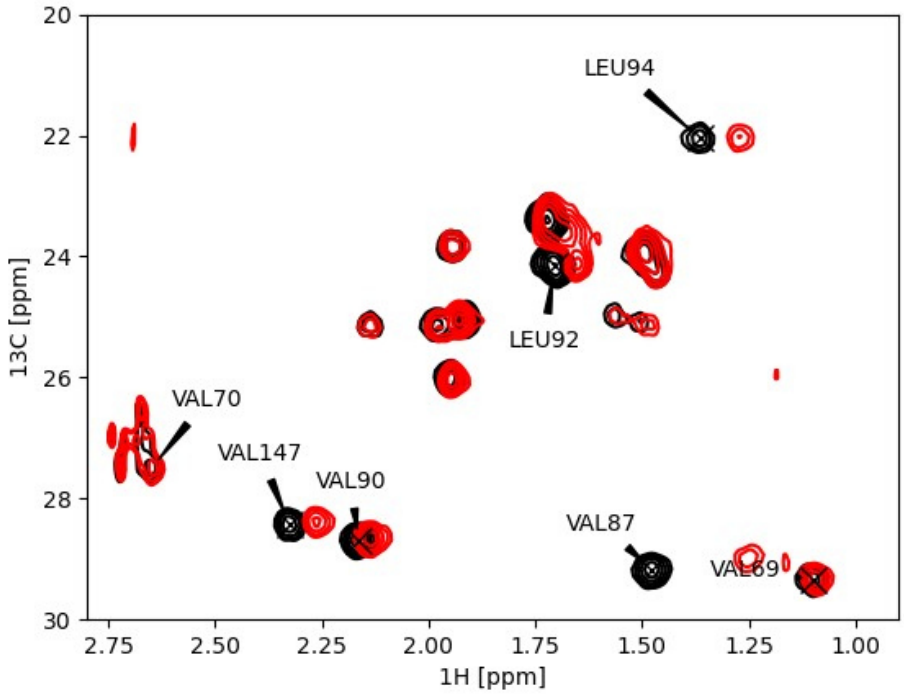
Ligand 17



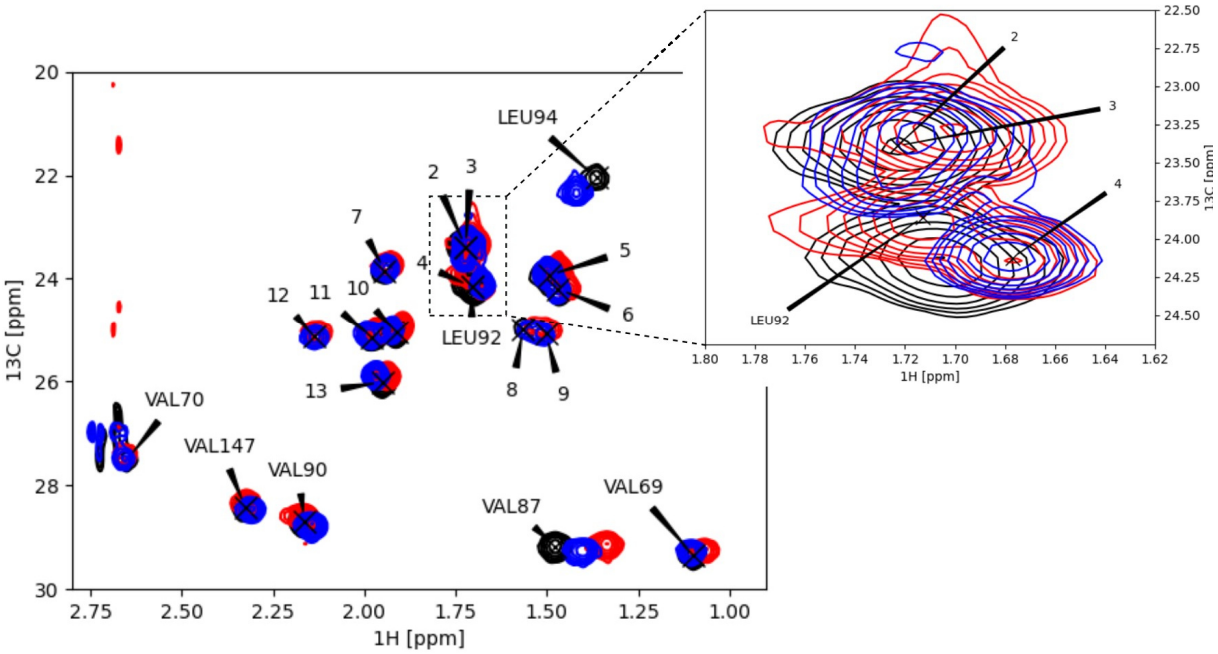
Ligand 18



Ligand 19



5. Overlay ^{1}H-^{13}C-HSQC Brd4-BD1 WT/L92A/L94A



Overlay of ^{1}H-^{13}C HSQCs of WT Brd4-BD1 (black), L94A Brd4-BD1 (red) and L92A Brd4-BD1 (blue). The Leu94 signal was identified by the absence of the corresponding peak in the spectrum of L94A Brd4-BD1 in red. The Leu92 signal was identified by interpreting the spectrum of L92A Brd4-BD1 in blue, which shows the absence of the Leu92 signal in a region of overlapping signals (zoom window). All three protein samples were labeled by adding compound **7** to the minimal medium as described in the experimental part of the main text.

Part III

DISCUSSION AND OUTLOOK

4

DISCUSSION AND OUTLOOK

NMR Spectroscopy has been widely used in structure determination and dynamical investigations of biomacromolecules. Due to its ability to detect weak and transient interactions, NMR Spectroscopy became a powerful tool in drug design programs, especially in the FBDD field. The papers presented in this thesis clearly show that the chemical shifts obtained with NMR Spectroscopy are very sensitive reporters to even little structural changes. Moreover, the chemical shifts reported here serve as a direct probe for specific directional non-covalent interactions, the CH-π interactions. Having the feasibility to unwind different types of non-covalent interactions and their impact on the complex formation, this type of information is predestined to impact medicinal chemistry.

Traditional drug design programs are structure- or ligand-based and require therefore a well-resolved structure of the target protein bound to ligands. These structures are mainly derived from X-ray crystallography data and are certainly a good starting point. Nevertheless, there are some drawbacks in structure-based drug design processes. One major point is that proteins exist as conformational ensembles that are not depicted in the crystal structures. The analysis of non-covalent interactions in the bound state are indirectly estimated according to the crystal structure, rather than directly probed. Moreover, the contribution to the binding entropy resulting from residual dynamical movement in the ligand-bound state is an additional important information in the design and optimization process, which can also not be derived from a single static structure. Another major point is the importance of hydration water molecules in the binding interface and their contribution to the binding energy. An adequate and precise description of the complex is therefore crucial in drug design programs, especially in the hit-to-lead optimization stages.

Structure-based drug discovery programs, and in particular FBDD, were able to find suitable drugs or hit ligands for so-called undruggable protein targets, nevertheless, its success rate in drug discovery processes has not reached its expectations [33]. One major bottleneck in structure-based drug design programs is still the generation of structures of the target proteins. Either the protein does not crystallize or the resolution of the generated structure is too small. Moreover, the output relies on one stoic and simplified state of the complex formation, which is not appropriate, especially in the case of intrinsically disordered proteins (IDPs). IDPs play crucial roles in biological processes, which makes them plausible drug targets. Contrary to folded

proteins, IDPs do not have well-defined binding pockets that lead fragment selection, alternative approaches in drug design programs based on their conformational ensembles are therefore needed.

The interplay of NMR Spectroscopy and computational biology is a fruitful strategy to enhance drug discovery programs. With selective labeling strategies, fundamental NMR experiments, simple chemical shift calculation approaches and AI-generated structure predictions (e.g. AlphaFold2), it is possible to determine the relevance of each contribution to its binding affinity.

Ensemble-averaged structure generation with NMR Spectroscopy and MD simulations leads to a more adequate and accurate description of the formed complex and its dynamical properties. Moreover, as shown in the two papers, the simplicity of the Pople model is sufficient to approximate the effect of the ring current effects onto ^{1}H chemical shifts of the protein to directly probe CH-π interactions. The usage of more advanced QM/MM methods would even improve the accuracy and is also suitable to calculate ^{1}H ligand chemical shifts in order to map the binding interactions from the ligand's perspective [76].

Taking all these things together, it is possible to form a better understanding between the affinity, its non-covalent interactions and structures of protein-ligand binding complexes, thereby accelerating drug discovery processes. Ultimately, by combining NMR-derived data and computational methods, we may even be able to target previously inaccessible protein systems.

BIBLIOGRAPHY

[1] A. Sekhar and L. E. Kay, "An NMR View of Protein Dynamics in Health and Disease," *Annual Review of Biophysics*, vol. 48, pp. 297–319, 2019.

[2] A. Kawale and B. Burmann, "Characterization of backbone dynamics using solution NMR spectroscopy to discern the functional plasticity of structurally analogous proteins," *Cell Press*, vol. 2, no. 4, p. 100 919, 2021.

[3] C. A. Waudby and I. Alfonso, "An introduction to one- and two-dimensional lineshape analysis of chemically exchanging systems," *Journal of Magnetic Resonance Open*, vol. 16-17, no. March, p. 100 102, 2023.

[4] A. M. Gronenborn, "Small, but powerful and attractive: 19f in biomolecular nmr," *Structure*, vol. 30, no. 1, pp. 6–14, 2022.

[5] G. Parigi, E. Ravera, M. Piccioli, and C. Luchinat, "Paramagnetic nmr restraints for the characterization of protein structural rearrangements," *Current Opinion in Structural Biology*, vol. 80, p. 102 595, 2023.

[6] V. K. Shukla, G. T. Heller, and D. F. Hansen, "Biomolecular NMR spectroscopy in the era of artificial intelligence," *Structure*, vol. 31, no. 11, pp. 1360–1374, 2023.

[7] D. B. McConnell, "Biotin's Lessons in Drug Design," *Journal of Medicinal Chemistry*, vol. 64, no. 22, pp. 16 319–16 327, 2021.

[8] A. D. Gossert and W. Jahnke, "NMR in drug discovery: A practical guide to identification and validation of ligands interacting with biological macromolecules," *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 97, pp. 82–125, 2016.

[9] M. J. Harner, L. Mueller, K. J. Robbins, and M. D. Reily, "NMR in drug design," *Archives of Biochemistry and Biophysics*, vol. 628, pp. 132–147, 2017.

[10] A. Boeszoermenyi, S. Chhabra, A. Dubey, D. L. Radeva, N. T. Burdzhiev, C. D. Chanev, O. I. Petrov, V. M. Gelev, M. Zhang, C. Anklin, *et al.*, "Aromatic 19f-13c trosy: A background-free approach to probe biomolecular structure, function, and dynamics," *Nature methods*, vol. 16, no. 4, pp. 333–340, 2019.

[11] C. Dalvit, I. Gmür, P. Rößler, and A. D. Gossert, "Affinity measurement of strong ligands with NMR spectroscopy: Limitations and ways to overcome them," *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 138-139, no. April, pp. 52–69, 2023.

[12] M. J. Harner, A. O. Frank, and S. W. Fesik, "Fragment-Based Drug Discovery Using NMR Spectroscopy," *J. Biomol. NMR*, vol. 56, no. 2, pp. 65–75, 2013.

[13] M. Pellecchia, D. S. Sem, and K. Wüthrich, "Nmr in drug discovery," *Nature Reviews Drug Discovery*, vol. 1, no. 3, pp. 211–219, 2002.

[14] C. Di Carluccio, M. C. Forgione, S. Martini, F. Berti, A. Molinaro, R. Marchetti, and A. Silipo, "Investigation of protein-ligand complexes by ligand-based NMR methods," *Carbohydrate Research*, vol. 503, no. January, p. 108 313, 2021.

[15] C. Bissantz, B. Kuhn, and M. Stahl, "A medicinal chemist's guide to molecular interactions," *Journal of Medicinal Chemistry*, vol. 53, no. 14, pp. 5061–5084, 2010.

[16] G Otting, E Liepinsh, and K Wüthrich, "Protein hydration in aqeuous solution," *Science*, vol. 254, no. 1985, pp. 974–980, 1991.

[17] N. Y. Meredith, S. Borsley, I. V. Smolyar, G. S. Nichol, C. M. Baker, K. B. Ling, and S. L. Cockroft, "Dissecting Solvent Effects on Hydrogen Bonding," *Angewandte Chemie - International Edition*, vol. 61, no. 30, 2022.

[18] J. F. Darby, A. P. Hopkins, S. Shimizu, S. M. Roberts, J. A. Brannigan, J. P. Turkenburg, G. H. Thomas, R. E. Hubbard, and M. Fischer, "Water networks can determine the affinity of ligand binding to proteins," *Journal of the American Chemical Society*, vol. 141, no. 40, pp. 15 818–15 826, 2019.

[19] L. M. Amzel, "Loss of translational entropy in binding, folding, and catalysis.," eng, *Proteins*, vol. 28, no. 2, pp. 144–149, 1997.

[20] L. M. Amzel, "Calculation of entropy changes in biological processes: folding, binding, and oligomerization.," eng, *Methods in enzymology*, vol. 323, pp. 167–177, 2000.

[21] A Cooper and C. M. Johnson, "Introduction to microcalorimetry and biomolecular energetics.," eng, *Methods in molecular biology (Clifton, N.J.)*, vol. 22, pp. 109–124, 1994.

[22] J. D. Dunitz, "Win some, lose some: Enthalpy-entropy compensation in weak intermolecular interactions.," *Chemistry & biology*, vol. 2, no. 11, pp. 709–712, 1995.

[23] U. Ryde, "A fundamental view of enthalpy–entropy compensation," *MedChemComm*, vol. 5, no. 9, pp. 1324–1336, 2014.

[24] M. C. Chervenak and E. J. Toone, "A Direct Measure of the Contribution of Solvent Reorganization to the Enthalpy of Binding," *Journal of the American Chemical Society*, vol. 116, no. 23, pp. 10 533–10 539, 1994.

[25] X. Du, Y. Li, Y. L. Xia, S. M. Ai, J. Liang, P. Sang, X. L. Ji, and S. Q. Liu, "Insights into protein–ligand interactions: Mechanisms, models, and methods," *International Journal of Molecular Sciences*, vol. 17, no. 2, pp. 1–34, 2016.

[26] E. R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, A. J. Cohen, and W. Yang, "Revealing noncovalent interactions," *Journal of the American Chemical Society*, vol. 132, no. 18, pp. 6498–6506, 2010.

[27] E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg, P. Hobza, H. G. Kjaergaard, A. C. Legon, B. Mennucci, and D. J. Nesbitt, "Definition of the hydrogen bond (IUPAC Recommendations 2011)," *Pure Appl. Chem*, vol. 83, no. 8, pp. 1637–1641, 2011.

[28] R. F. D. Freitas and M. Schapira, "A systematic analysis of atomic protein – ligand interactions in the PDB," *MedChemComm*, pp. 1970–1981, 2017.

[29] M. Nishio, Y. Umezawa, M. Hirota, and Y. Takeuchi, "The CH/π interaction: Significance in molecular recognition," *Tetrahedron*, vol. 51, no. 32, pp. 8665–8701, 1995.

[30] C. A. Hunter, "Quantifying intermolecular interactions: Guidelines for the molecular recognition toolbox," *Angewandte Chemie - International Edition*, vol. 43, no. 40, pp. 5310–5324, 2004.

[31] Y. Xiao and R. J. Woods, "Protein-Ligand CH-π Interactions: Structural Informatics, Energy Function Development, and Docking Implementation," *Journal of Chemical Theory and Computation*, vol. 19, no. 16, pp. 5503–5515, 2023.

[32] L. G. Mureddu and G. W. Vuister, "Fragment-Based Drug Discovery by NMR. Where Are the Successes and Where can It Be Improved?" *Frontiers in Molecular Biosciences*, vol. 9, no. February, pp. 1–14, 2022.

[33] L. Wang, J. Gao, R. Ma, Y. Liu, M. Liu, F. Zhong, J. Hu, S. Li, J. Wu, H. Jiang, J. Zhang, and K. Ruan, "Recent progress in fragment-based drug discovery facilitated by NMR spectroscopy," *Magnetic Resonance Letters*, vol. 2, no. 2, pp. 107–118, 2022.

[34] D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, and H. Jhoti, "Twenty years on: The impact of fragments on drug discovery," *Nature reviews Drug discovery*, vol. 15, no. 9, pp. 605–619, 2016.

[35] A. Dubey, K. Takeuchi, M. Reibarkh, and H. Arthanari, "The role of NMR in leveraging dynamics and entropy in drug design," *J. Biomol. NMR*, vol. 74, no. 10-11, pp. 479–498, 2020.

[36] C. W. Murray and D. C. Rees, "The rise of fragment-based drug discovery," *Nature chemistry*, vol. 1, no. 3, pp. 187–192, 2009.

[37] D. A. Erlanson, "Introduction to fragment-based drug discovery," *Fragment-based drug discovery and X-ray crystallography*, pp. 1–32, 2012.

[38] L. Shi and N. Zhang, "Applications of solution nmr in drug discovery," *Molecules*, vol. 26, no. 3, pp. 1–21, 2021.

[39] M. P. Williamson, "Using chemical shift perturbation to characterise ligand binding," *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 73, pp. 1–16, 2013.

[40] V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw, and L. E. Kay, "Cross-correlated relaxation enhanced 1H-13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes," *Journal of the American Chemical Society*, vol. 125, no. 34, pp. 10 420–10 428, 2003.

[41] V. Tugarinov, R. Sprangers, and L. E. Kay, "Line narrowing in methyl-TROSY using zero-quantum 1H-13C NMR spectroscopy.," eng, *Journal of the American Chemical Society*, vol. 126, no. 15, pp. 4921–4925, 2004.

[42] A. G. Palmer, J. Cavanagh, R. A. Byrd, and M. Rance, "Sensitivity improvement in three-dimensional heteronuclear correlation NMR spectroscopy," *Journal of Magnetic Resonance*, vol. 96, no. 2, pp. 416–424, 1992.

[43] V. Tugarinov, V. Kanelis, and L. E. Kay, "Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy.," eng, *Nature protocols*, vol. 1, no. 2, pp. 749–754, 2006.

[44] K. Pervushin, R. Riek, G. Wider, and K. Wüthrich, "Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 23, pp. 12 366–12 371, 1997.

[45] S. P. Behera, A. Dubey, W. N. Chen, V. S. De Paula, M. Zhang, N. G. Sgourakis, W. Bermel, G. Wagner, P. W. Coote, and H. Arthanari, "Nearest-neighbor NMR spectroscopy: categorizing spectral peaks by their adjacent nuclei," *Nature Communications*, vol. 11, no. 1, pp. 1–7, 2020.

[46] I. C. Felli and B. Brutscher, "Recent advances in solution nmr: Fast methods and heteronuclear direct detection," *ChemPhysChem*, vol. 10, no. 9-10, pp. 1356–1368, 2009.

[47] J. J. Ziarek, D. Baptista, and G. Wagner, "Recent developments in solution nuclear magnetic resonance (nmr)-based molecular biology," *Journal of Molecular Medicine*, vol. 96, pp. 1–8, 2018.

[48] S. Grzesiek and A. Bax, "Spin-locked multiple quantum coherence for signal enhancement in heteronuclear multidimensional NMR experiments," *Journal of Biomolecular NMR*, vol. 6, no. 3, pp. 335–339, 1995.

[49] G. W. Vuister and A. Bax, "Resolution enhancement and spectral editing of uniformly 13C-enriched proteins by homonuclear broadband 13C decoupling," *Journal of Magnetic Resonance (1969)*, vol. 98, no. 2, pp. 428–435, 1992.

[50] R. Lichtenecker, M. L. Ludwiczek, W. Schmid, and R. Konrat, "Simplification of protein NOESY spectra using bioorganic precursor synthesis and NMR spectral editing.," eng, *Journal of the American Chemical Society*, vol. 126, no. 17, pp. 5348–5349, 2004.

[51] R. J. Lichtenecker, N. Coudevylle, R. Konrat, and W. Schmid, "Selective Isotope Labelling of Leucine Residues by Using α-Ketoacid Precursor Compounds," *ChemBioChem*, vol. 14, no. 7, pp. 818–821, 2013.

[52] R. J. Lichtenecker, K. Weinhäupl, L. Reuther, J. Schörghuber, W. Schmid, and R. Konrat, "Independent valine and leucine isotope labeling in Escherichia coli protein overexpression systems," *Journal of Biomolecular NMR*, vol. 57, no. 3, pp. 205–209, 2013.

[53] R. J. Lichtenecker, J. Schoerghuber, and M. Bisaccia, "Synthesis of metabolic amino acid precursors: Tools for selective isotope labeling in cell-based protein overexpression," *Synlett*, vol. 26, no. 19, pp. 2611–2616, 2015.

[54] J. Schörghuber, T. Sára, M. Bisaccia, W. Schmid, R. Konrat, and R. J. Lichtenecker, "Novel approaches in selective tryptophan isotope labeling by using escherichia coli overexpression media," *ChemBioChem*, vol. 16, no. 5, pp. 746–751, 2015.

[55] J. Schörghuber, L. Geist, M. Bisaccia, F. Weber, R. Konrat, and R. J. Lichtenecker, "Anthranilic acid, the new player in the ensemble of aromatic residue labeling precursor compounds," *Journal of Biomolecular NMR*, vol. 69, pp. 13–22, 2017.

[56] J. Schörghuber, L. Geist, G. Platzer, R. Konrat, and R. J. Lichtenecker, "Highly selective stable isotope labeling of histidine residues by using a novel precursor in e. coli-based overexpression systems," *ChemBioChem*, vol. 18, no. 15, pp. 1487–1491, 2017.

[57] J. Schörghuber, L. Geist, G. Platzer, M. Feichtinger, M. Bisaccia, L. Scheibelberger, F. Weber, R. Konrat, and R. J. Lichtenecker, "Late metabolic precursors for selective aromatic residue labeling," *Journal of Biomolecular NMR*, vol. 71, pp. 129–140, 2018.

[58] M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. Mei Ono, and P. Güntert, "Optimal isotope labelling for NMR protein structure determinations.," eng, *Nature*, vol. 440, no. 7080, pp. 52–57, 2006.

[59] P. Lundström, K. Teilum, T. Carstensen, I. Bezsonova, S. Wiesner, D. F. Hansen, T. L. Religa, M. Akke, and L. E. Kay, "Fractional 13C enrichment of isolated carbons using [1-13C]- or [2-13C]-glucose facilitates the accurate measurement of dynamics at backbone Calpha and side-chain methyl positions in proteins.," eng, *Journal of biomolecular NMR*, vol. 38, no. 3, pp. 199–212, 2007.

[60] A. J. Wand, R. J. Bieber, J. L. Urbauer, R. P. McEvoy, and Z Gan, "Carbon relaxation in randomly fractionally 13C-enriched proteins.," eng, *Journal of magnetic resonance. Series B*, vol. 108, no. 2, pp. 173–175, 1995.

[61] N. K. Goto, K. H. Gardner, G. A. Mueller, R. C. Willis, and L. E. Kay, "A robust and cost-effective method for the production of Val, Leu, Ile ($\delta 1$) methyl-protonated 15N-, 13C-, 2H-labeled proteins," *Journal of Biomolecular NMR*, vol. 13, no. 4, pp. 369–374, 1999.

[62] M. K. Rosen, K. H. Gardner, R. C. Willis, W. E. Parris, T. Pawson, and L. E. Kay, "Selective methyl group protonation of perdeuterated proteins," *Journal of Molecular Biology*, vol. 263, no. 5, pp. 627–636, 1996.

[63] A. M. Ruschak, A. Velyvis, and L. E. Kay, "A simple strategy for 13C,1H labeling at the Ile-$\gamma 2$ methyl position in highly deuterated proteins," *Journal of Biomolecular NMR*, vol. 48, no. 3, pp. 129–135, 2010.

[64] K. H. Gardner and L. E. Kay, "Production and Incorporation of 15 N, 13 C, 2 H (1 H-$\delta 1$ Methyl) Isoleucine into Proteins for Multidimensional NMR Studies," *Journal of the American Chemical Society*, vol. 119, no. 32, pp. 7599–7600, 1997.

[65] M. H. AL Mughram, C. Catalano, N. B. Herrington, M. K. Safo, and G. E. Kellogg, "3D interaction homology: The hydrophobic residues alanine, isoleucine, leucine, proline and valine play different structural roles in soluble and membrane proteins," *Frontiers in Molecular Biosciences*, vol. 10, no. March, pp. 1–24, 2023.

[66] F. Desantis, M. Miotto, L. Di Rienzo, E. Milanetti, and G. Ruocco, "Spatial organization of hydrophobic and charged residues affects protein thermal stability and binding affinity," *Scientific Reports*, vol. 12, no. 1, pp. 1–13, 2022.

[67] C. W. Haigh and R. B. Mallion, "Ring current theories in nuclear magnetic resonance," *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 13, no. 4, pp. 303–344, 1979.

[68] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, and K. Tanabe, "The magnitude of the CH/π interaction between benzene and some model hydrocarbons," *Journal of the American Chemical Society*, vol. 122, no. 15, pp. 3746–3753, 2000.

[69] L. M. Salonen, M. Ellermann, and F. Diederich, "Aromatic rings in chemical and biological recognition: Energetics and structures," *Angewandte Chemie - International Edition*, vol. 50, no. 21, pp. 4808–4842, 2011.

[70] A. B. Sahakyan and M. Vendruscolo, "Analysis of the contributions of ring current and electric field effects to the chemical shifts of RNA bases," *Journal of Physical Chemistry B*, vol. 117, no. 7, pp. 1989–1998, 2013.

[71] G. Platzer, M. Mayer, A. Beier, S. Brüschweiler, J. E. Fuchs, H. Engelhardt, L. Geist, G. Bader, J. Schörghuber, R. Lichtenecker, B. Wolkerstorfer, D. Kessler, D. B. McConnell, and R. Konrat, "PI by NMR: Probing CH–π Interactions in Protein–Ligand Complexes by NMR Spectroscopy," *Angewandte Chemie International Edition*, vol. 59, no. 35, pp. 14 861–14 868, 2020.

[72] S. B. Shuker, P. J. Hajduk, R. P. Meadows, and S. W. Fesik, "Discovering High-Affinity Ligands for Proteins: SAR by NMR," *Science*, vol. 274, no. 5292, pp. 1531–1534, 1996.

[73] J. A. Pople, "Proton Magnetic Resonance of Hydrocarbons," *J. Chem. Phys*, vol. 24, pp. 1111–1111, 1956.

[74] T. Höfurthner, G. Toscano, G. Kontaxis, A. Beier, M. Mayer, L. Geist, D. B. McConnell, H. Weinstabl, R. Lichtenecker, and R. Konrat, "Synthesis of a 13C-methylene-labeled isoleucine precursor as a useful tool for studying protein side-chain interactions and dynamics," *Journal of Biomolecular NMR*, 2023.

[75] G. Toscano, T. Höfurthner, B. Nagl, A. Beier, M. Mayer, L. Geist, D. B. McConnell, H. Weinstabl, R. J. Lichtenecker, and R. Konrat, "13Cβ-Valine and 13Cγ-Leucine Methine Labeling To Probe Protein-Ligand Interaction," *ChemBioChem*, 2024.

[76] G. Platzer, A. L. Ptaszek, J. Böttcher, J. E. Fuchs, P. A. Sánchez-murcia, and M. Mayer, "Ligand 1H NMR Chemical Shifts as Accurate Reporters for Protein-Ligand Binding Interfaces in Solution," *ChemPhysChem*, vol. 202300636, no. 25, 2024.