



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

Theresa Höfurtherner^{1,3} · Giorgia Toscano^{2,3} · Georg Kontaxis⁴ · Andreas Beier¹ · Moriz Mayer⁵ · Leonhard Geist⁵ · Darryl B. McConnell⁵ · Harald Weinstabl⁵ · Roman Lichtenecker² · Robert Konrat^{1,4}

Received: 5 July 2023 / Accepted: 28 September 2023
© The Author(s) 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 - ^1H -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 ^1H -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 ^1H - ^{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 (^1H - ^{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

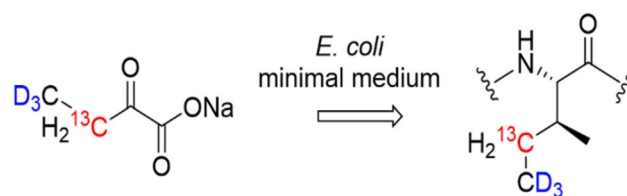
⁵ Boehringer Ingelheim RCV GmbH & Co. KG, Dr. Boehringer Gasse 5-11, 1121 Vienna, Austria

ligand binding modes and $^1\text{H}^{\text{N}}\text{-}^{15}\text{N}$ chemical shift perturbations (Williamson 2013). These problems can be largely overcome by resorting to $^1\text{H}\text{-}^{13}\text{C}$ 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 ^2H -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 $\gamma\text{-}^{13}\text{C}$ Isoleucine patterns are effectively induced by the corresponding precursor [$3\text{-}^{13}\text{C}$, $4,4,4\text{-}^2\text{H}_3$] 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 [$3\text{-}^{13}\text{C}$, $4,4,4\text{-}^2\text{H}_3$] 2-ketobutyrate in D_2O *E.coli* overexpression medium containing $U\text{-}^2\text{H}$ glucose to increase the deuteration grade in the β and γ_2 positions of the Ile-side chains. This additional deuteration attenuates line-broadening due to $^1\text{H}\text{-}^1\text{H}$ dipolar interaction and reduces the number of signals resulting from natural abundance methyl correlations.



Scheme 1 Conversion of [$3\text{-}^{13}\text{C}$; $4,4,4\text{-}^2\text{H}_3$] 2-ketobutyrate in *E. coli* minimal media to Isoleucine

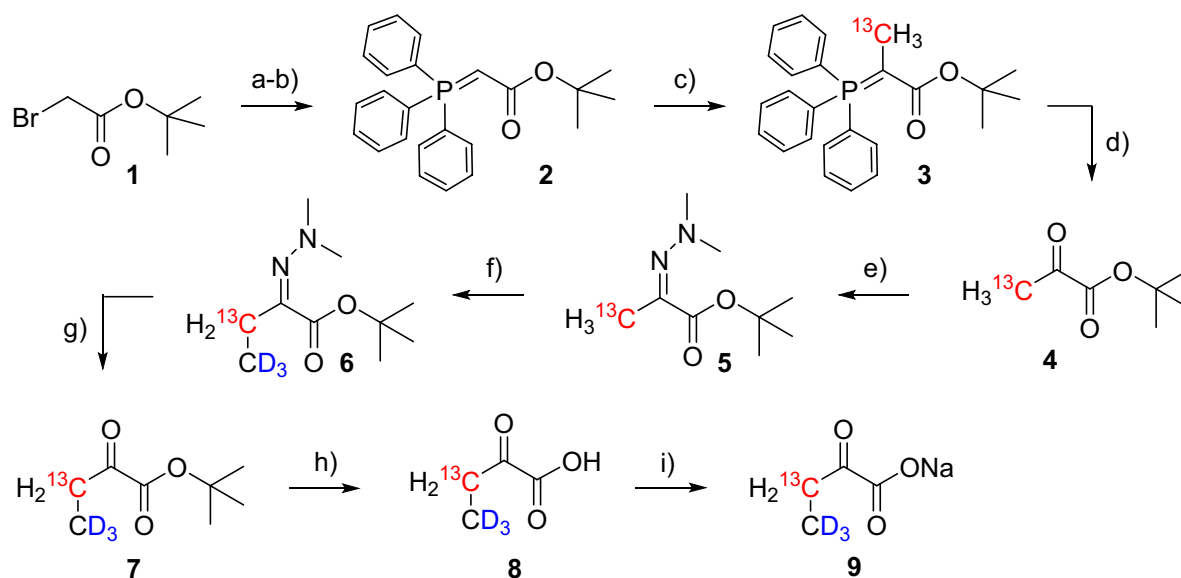
Materials and methods

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 $\gamma_1\text{-}^{13}\text{CH}_2/\delta\text{-CD}_3$ pattern (Scheme 2). Our approach combines a modified literature-known protocol to synthesize $3\text{-}^{13}\text{C}$ 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 [$3\text{-}^{13}\text{C}$] *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 [$^2\text{H}_3$] 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 [$3\text{-}^{13}\text{C}$; $4,4,4\text{-}^2\text{H}_3$] 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 [3-¹³C; 4,4,4-²H₃] sodium α-ketobutyrate **9**. Reagents and conditions: **a** PPh₃, toluene, rt, overnight; quant.; **b** NaOH_{aq} (1 M), rt, 15 min, quant.; **c** ¹³CH₃I, DCM, rt, overnight; **d** O₃, -78 °C, DCM, 5 min, then PPh₃, 3 h, rt, 70% over two steps; **e** H₂NNMe₂, Et₂O, 18 h, rt, 76%; **f** NaHMDS, CD₃I, THF, -78 °C,

5 min, 80%; **g** HCl_{aq} (1 M), Et₂O, 95%; **h** HCl_g, DCM/Et₂O 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 ¹⁵N and ¹³C-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 ¹⁵NH₄Cl, 3 g/L ¹³C₆-D-glucose. Uniformly ¹⁵N and selective γ₁-¹³C Isoleucine labeled Brd4-BD1 was expressed following the same protocol, M9 minimal medium was supplemented with 1 g/L ¹⁵NH₄Cl, 3 g/L ¹²C₆-D-glucose and 130 mg/L [3-¹³C, 4,4,4-²H₃] 2-ketobutyrate (Marley et al. 2001). Perdeuterated, uniformly ¹⁵N, D-Glucose-1,2,3,4,5,6,6-d₇ and selective γ₁-¹³C 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-H₂O minimal medium for 8 h at 37 °C shaking. From that culture, 250 μL were taken and inoculated in 10 mL fresh M9-D₂O 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-D₂O 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 Ni²⁺ affinity chromatography. The purified protein was treated with TEV protease and again loaded onto a Ni²⁺ 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 D₂O 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% D₂O, and 1 mM DTT. NMR samples of perdeuterated Brd4-BD1 were prepared in 10 mM sodium phosphate D₂O buffer containing 0.2 mM protein, 100 mM sodium chloride, and 1 mM deuterated Tris(2-Carboxyethyl)phosphine:DCI-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 γ₁-¹³C Isoleucine Brd4-BD1 sample concentration of 200 μM. For ¹³C relaxation studies, pulse sequence “hsqcctetgsp” for the constant time ¹H-¹³C 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 - $^1\text{H}_2$ heteronuclear triple/single quantum coherence, the pulse sequence “hmqcctetgp.2” (slightly modified) for the constant time ^1H - ^{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) \times 1024 (f2) real points (CT-HSQC) and 144 (f1) \times 1024 (f2) real points (CT-HTQC) with acquisition times of 0.06 \times 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 ^1H - ^{13}C HSQC spectra were acquired using the pulse sequence “hscqctfpgpsi2” 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) \times 1024 (f2) real points and acquisition times of 0.05 \times 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 $\text{C}\gamma_1$ 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 $[3\text{-}^{13}\text{C}; 4,4,4\text{-}^2\text{H}_3]$ 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 ^1H - ^{13}C -HSQC spectrum (Fig. 1, in red). As expected, two cross peaks are obtained for the diastereotopic methylene protons of each Ile $\text{C}\gamma_1$ differing in the hydrogen dimension, but with the same carbon chemical shift. Note that the peaks of two of the Isoleucine $\text{C}\gamma_1$ groups are overlapping (presumably I100 and I101). However, spurious cross peaks are found in the methyl group region of the ^1H - ^{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 $[3\text{-}^{13}\text{C}; 4,4,4\text{-}^2\text{H}_3]$ 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 $^{13}\text{CH}_2$ 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 $^{13}\text{CH}_2$ coherences relax more slowly than the corresponding single quantum coherences of the individual ^{13}C and ^1H 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 ^1H - ^{13}C -CT-HSQC and ^1H - ^{13}C -CT-HTQC spectra (Fig. 2) with different relaxation delays (see experimental section). Figure 3 shows a comparison of the rates extracted from the ^1H - ^{13}C -CT-HSQC and ^1H - ^{13}C -CT-HTQC, respectively. It can clearly be seen that the ^{13}C relaxation times for heteronuclear T/SQC extracted from the ^1H - ^{13}C -CT-HTQC were

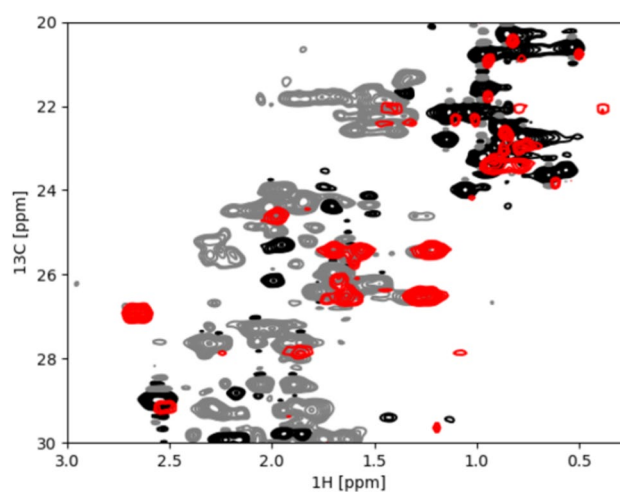


Fig. 1 Overlay of ^1H - ^{13}C -HSQC spectra of selectively Ile $^{13}\text{CH}_2$ labeled Brd4-BD1 (red) onto the CT- ^1H - ^{13}C -HSQC of ^{13}C -uniformly labeled Brd4-BD1 (^{13}CH and $^{13}\text{CH}_3$ signals in black and $^{13}\text{CH}_2$ signals in grey), zoomed into the CH_2 region

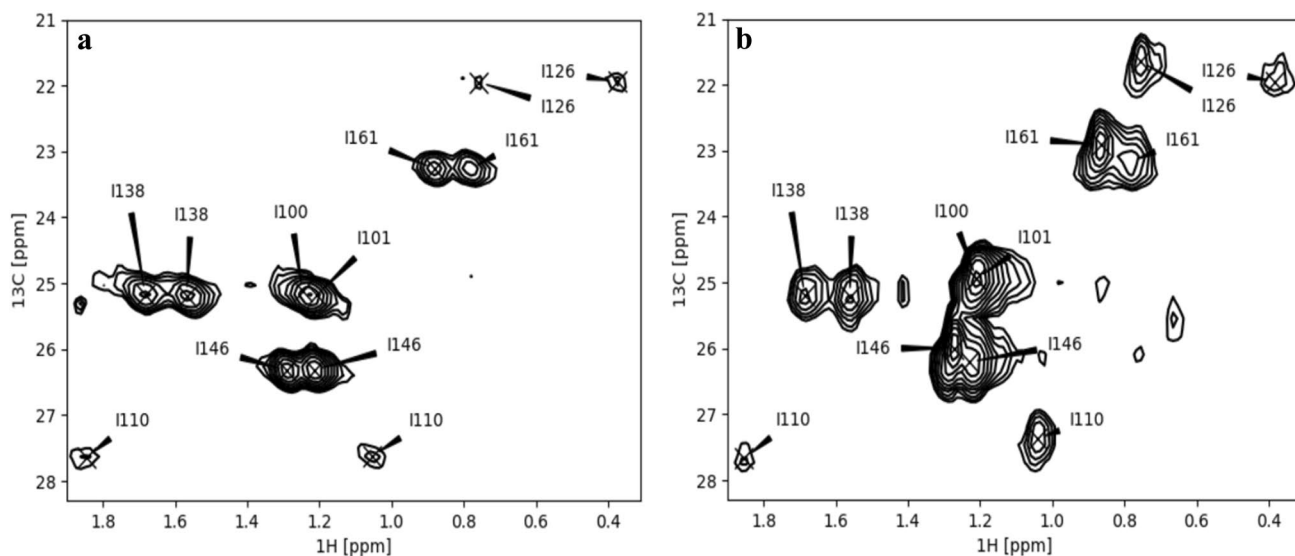


Fig. 2 Comparison between CT-HTQC (ct delay 0.028 s) (a) and CT-HSQC (ct delay 0.013 s) (b) spectra of perdeuterated ^{15}N -Brd4-BD1-Ile- $^{13}\text{C}\gamma 1$

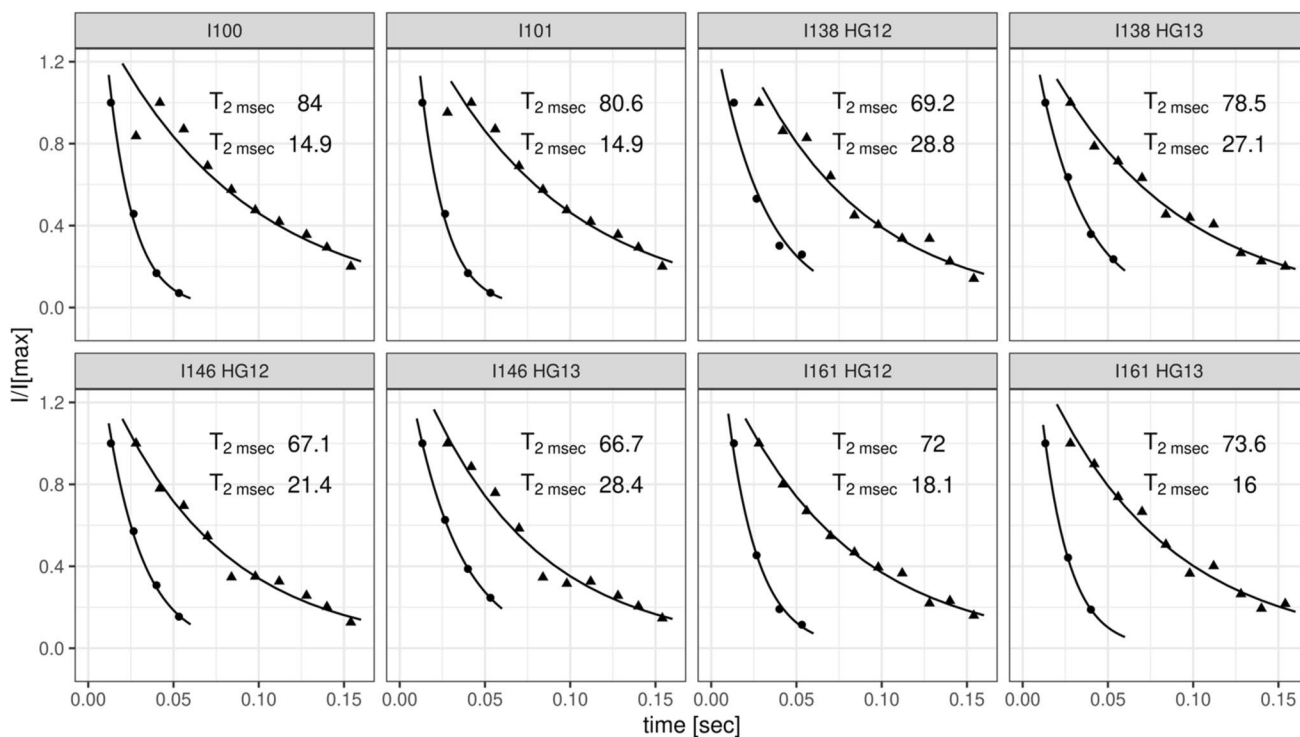


Fig. 3 Extracted ^{13}C relaxation times for specific Ile residues in perdeuterated Brd4-BD1. Dots represent relaxation times extracted from CT-HSQC, whereas pyramids represent rates extracted from CT-HTQCs

increased on average by a factor of three compared to the ^{13}C - T_2 from the ^1H - ^{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 CH₂-TROSY spectrum of the Ile-¹³CH₂ Brd-BD1 sample expressed in D₂O/deuterated glucose, which shows additional gain in spectral resolution – even relative to the HTQC spectrum (especially in the ¹H 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 CH₂-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 ¹H-¹³C-HSQC with ligands **A** and **B**, which are nanomolar Brd4-BD1 binders (Platzer et al. 2020). Figure 4a shows the chemical shift changes of ¹⁵N-Brd4-BD1-Ile-¹³C γ 1 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 γ 1H₂ 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 C γ 1-H₂ pairs in an otherwise

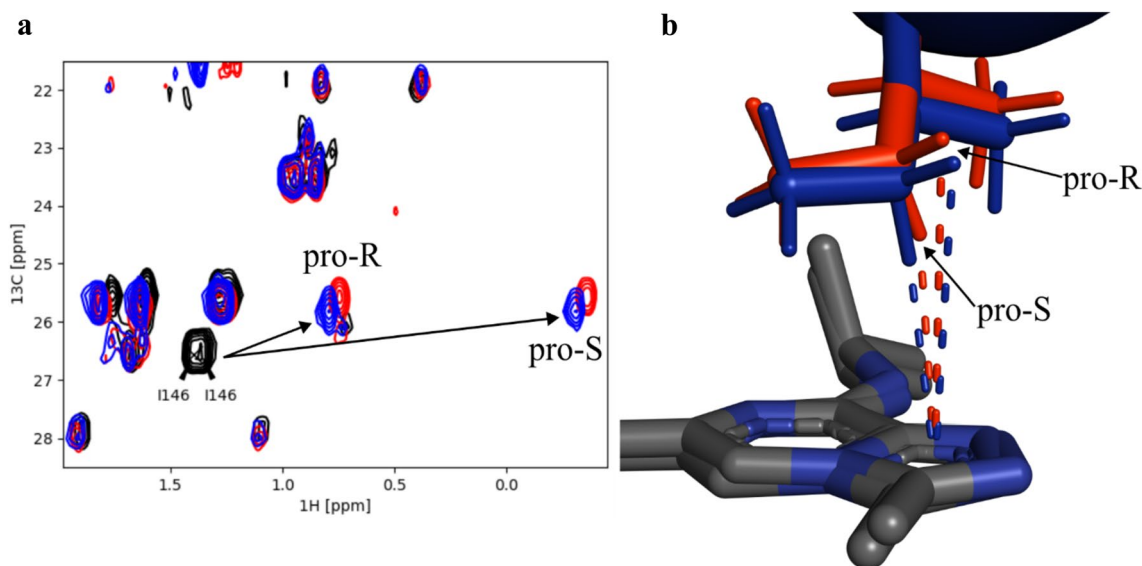


Fig. 4 **a** ¹H-¹³C-HSQC overlay of Brd4-BD1-Ile-¹³C γ 1 (black) with Ligand **A** (red) and Ligand **B** (blue). **b** Zoom into the Brd4-BD1 binding pocket with an overlay of Ligand **A** (red, 6XV3) and **B** (blue,

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 ^1H - ^{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 $\text{C}\gamma 1$ labeled Brd4-BD1 by studying its carbon transverse relaxation properties by measuring ^{13}C single and $^{13}\text{C}/^1\text{H}$ 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. Höfurtherner, 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. Brüscheweiler, 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.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source,

provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

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 , ^2H (^1H - $\delta 1$ 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 $\text{H}\alpha$ — $\text{H}\beta$ 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/ja0003501>
- 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 ¹³C-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](https://doi.org/10.1016/0022-2364(92)90097-Q)
- Platzer G, Mayer M, Beier A, Brüscheweiler 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 ¹³C,¹H labeling at the Ile- γ 2 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, Schedletsky 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;¹³CH₃]-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, Ollershaw JE, Kay LE (2003) Cross-correlated relaxation enhanced ¹H-¹³C 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-¹³C 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 ¹³C-enriched proteins by homonuclear broadband ¹³C decoupling. *J Magn Reson* 98:428–435. [https://doi.org/10.1016/0022-2364\(92\)90144-V](https://doi.org/10.1016/0022-2364(92)90144-V)
- Werkhoven TM, van Nispen R, Lugtenburg J (1999) Specific 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>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.