J Biomol NMR (2013) 57:327–331
DOI 10.1007/s10858-013-9796-9

COMMUNICATION

α-Ketoacids as precursors for phenylalanine and tyrosine labelling in cell-based protein overexpression

Roman J. Lichtenecker · Katharina Weinhäupl · Walther Schmid · Robert Konrat

Received: 4 October 2013 / Accepted: 14 November 2013 / Published online: 22 November 2013

Abstract ^{13}C-α-ketoacid metabolic precursors of phenylalanine and tyrosine effectively enter the metabolism of a protein overexpressing *E. coli* strain to label Phe- and Tyr-residues devoid of any cross-labelling. The methodology gives access to highly selective labelling patterns as valuable tools in protein NMR spectroscopy without the need of ^{15}N-chiral amino acid synthesis using organic chemistry.

Keywords Isotope labeling · Protein NMR · Protein overexpression · Phenylalanine · Tyrosine

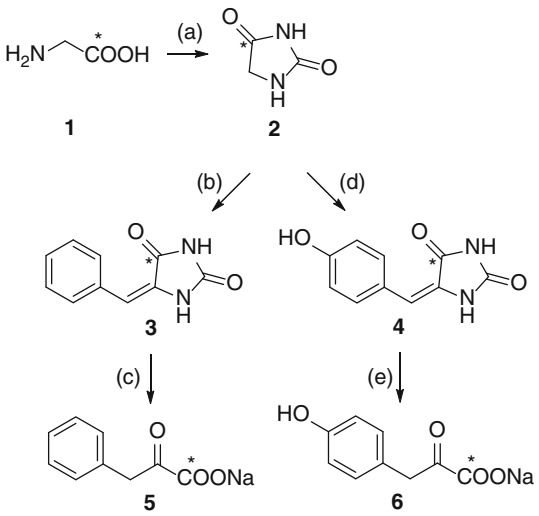
Main text

As aromatic residues are highly overrepresented in the hydrophobic core regions of proteins, the unambiguous identification of NOEs between aromatic and aliphatic side-chains is of high importance for accurate structure determination (Lin et al. 2006; Slupsky et al. 1998; Löhr et al. 2007). Tyrosine and phenylalanine residues play a decisive role in protein binding by contributing to hydrophobic-, cationic-π-, and polar-π-interactions (Bogan and Thorn 1998). Their exact orientation in space and dynamic properties are of special interest in fields such as structural biology, cell physiology and drug development. However, the marginal shift dispersion of aromatic ring signals often hinders an accurate NMR signal assignment or reduces the number of attainable structure parameters. The elucidation of aromatic side chain dynamic motion by detecting transverse relaxation rates is a difficult task as well, since this process follows a non-exponential evolution due to ^{13}C–^{13}C scalar coupling effects within the aromatic side chain spin system. Site specific ^{13}C- and ^{2}H-labelled aromatic residues have been applied to tackle these sensitivity and resolution problems. Aromatic isotope patterns render heteronuclear NMR experiments possible and close unwanted relaxation pathways to significantly increase NMR spectra resolution. The value of selectively labelled aromatic residues for accurate structure calculation based on NMR data has been shown impressively in the scope of the stereo-array isotope labelling technique (SAIL), where the application of labelled phenylalanine and tyrosine in a cell-free expression system resulted in a very exact NMR derived structure of the EPPIb protein (Takeda et al. 2010). Cell-free (CF) expression systems allow for the assembly of highly specific protein isotope patterns and have been applied to straightforward phenylalanine and tyrosine labelling devoid of cross-labelling to other residues (Kainosho et al. 2006; Torizawa et al. 2005). However, CF-systems require the laborious and expensive synthesis of ^{15}N- amino acids and often struggle with low protein yields (Staunton et al. 2006). The application of chemically less complex precursor compounds in combination with an overexpression system of a host organism provides a more general approach to produce labelled proteins at lower synthetic costs (Hoogstraten and Johnson 2008). Chiral labelled phenylalanine (Wang et al. 1999; Lee et al. 1995), glycerol (LeMaster and Kushlan 1996), glucose (Teilum et al. 2006), erythrose

Electronic supplementary material The online version of this article (doi:10.1007/s10858-013-9796-9) contains supplementary material, which is available to authorized users.

R. J. Lichtenecker (✉) · W. Schmid
Institute of Organic Chemistry, University of Vienna, Währingerstr. 38, 1090 Vienna, Austria
e-mail: roman.lichtenecker@univie.ac.at

K. Weinhäupl · R. Konrat
Department of Structural and Computational Biology, Max F. Perutz Laboratories, University of Vienna, Dr. Bohr-Gasse 9, 1030 Vienna, Austria



Scheme 1 Reagents and conditions: **a** KOCN, H_2O, 100 °C, 2 h; then HCl, 120 °C, 3 h, 95 %; **b** benzaldehyde, AcOH, NH_4OAc, 4 h, 120 °C, 85 %; **c** NaOH (20 %), 5 h, 100 °C, 90 %; **d** 4-hydroxybenzaldehyde, piperidine, 30 min., 130 °C, 65 %; **e** NaOH (20 %), 5 h, 100 °C, 83 %. *Asterisks* denote ^{13}C labelling

(Kasinath et al. 2013) and shikimic acid (Rajesh et al. 2003) have been used as nutrients in *Escherichia coli* expression hosts, resulting in aromatic residues enriched with ^{13}C or/and ^{2}H at distinct positions. However, these methods often suffer from low isotope incorporation levels, low selectivity concerning the isotope's spatial distribution in the protein sequence or require the use of auxotrophic host organisms.

E. coli-based overexpression systems containing α-ketoacid metabolic precursor compounds have become routinely used tools for protein labelling and have been realized by now for leucine- (Lichtenecker et al. 2013), valine- (Goto et al. 1999; Lichtenecker et al. 2004), isoleucine- (Goto et al. 1999), methionine- (Fischer et al. 2007), and alanine- (Ayala et al. 2009) labelling. Surprisingly, reports on the application of phenylalanine- or tyrosine α-ketoacid precursors are rare with the exception of a recently published article, describing reversed labelling by adding unlabelled sodium phenylpyruvate and 4-hydroxyphenylpyruvic acid to a bacterial growth medium (Rasia et al. 2012). In this study, relative high amounts

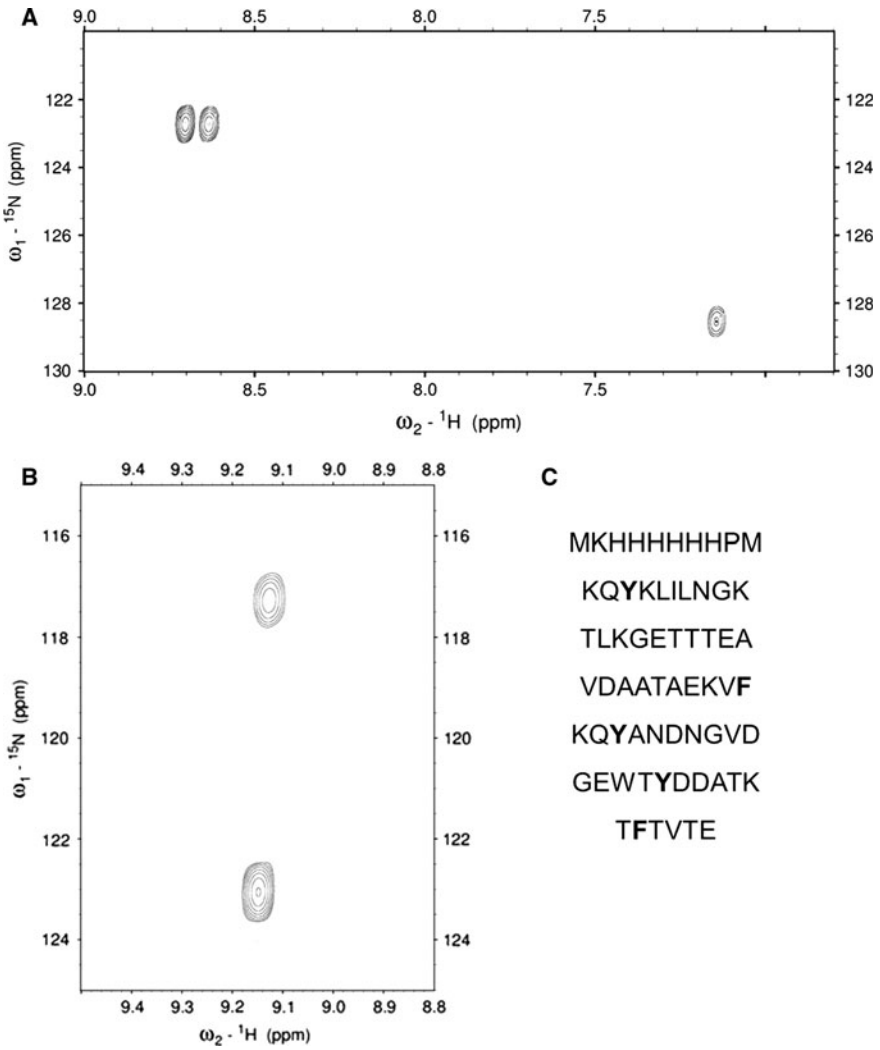


Fig. 1 HNCO spectra of His-tag GB1 overexpressed in presence of sodium [1-^{13}C]-4-hydroxyphenylpyruvate **6** (200 mg/L, **a**) and sodium [1-^{13}C] phenylpyruvate **5** (100 mg/L, **b**). Primary sequence of His-tag GB1 featuring three Tyr- and two Phe- residues, respectively (**c**)

of unlabelled precursors have been used (400 and 800 mg/L). We want to demonstrate, that a lower concentration of labelled phenylpyruvate and 4-hydroxyphenylpyruvate is sufficient to introduce specific stable isotope patterns into Phe- and Tyr- residues, respectively. To validate our hypothesis we chose an *E. coli* overexpressing organism for the synthesis of a immunoglobulin binding protein G containing a polyhistidine tag (His-tag GB1, 66 residues). The protein was expressed in presence of different concentrations of the labelled aromatic precursors sodium [1-^{13}C] phenylpyruvate **5** and sodium [1-^{13}C] 4-hydroxyphenylpyruvate **6**, which were both prepared from the same low prized source of label [1-^{13}C] glycine **1** (scheme 1) (Billek 1961a, b, c, modified).

Condensation of hydantoin **2** with benzaldehyde or 4-hydroxybenzaldehyde, respectively, and subsequent hydrolysis resulted in the formation of the corresponding α-ketoacids, which were obtained as their sodium-salts by lyophilisation (details are provided in the supporting information). His-tag GB1 was overexpressed in an *E. coli* medium supplemented with $^{15}NH_4Cl$ and glucose in presence of increasing concentrations of precursors **5** or **6**. The ^{13}C content was quantified by comparing HNCO/HSQC signal intensity ratios.

Notably, no interference between Phe- and Tyr-metabolism or cross labelling to other residues was observed under these conditions, since the resulting HNCO spectra revealed signals of the targeted residues only (Fig. 1). Our results clearly show that independent labelling of Phe- and Tyr- residues is feasible using α-ketoacid precursor compounds at concentrations, which are comparable to the ones used for aliphatic residue labelling. The HNCO signals displayed high isotope incorporation levels at already 100 mg/L of the corresponding precursor in the medium (Fig. 2) and reached the ^{13}C-content at the C_1 position of phenylalanines and tyrosines in an *all*-^{13}C protein (overexpressed in presence of $^{13}C_6$-glucose) at a precursor concentration of 120–150 mg/L (Fig. 3).

Various stable isotope patterns of phenylpyruvate **5** and 4-hydroxyphenylpyruvate **6** have been already described in literature, as these compounds represent important intermediates in synthetic routes to access the corresponding labelled amino acids (Raap et al. 1999). These aromatic-ring $^{13}C/^{2}H$ containing metabolic precursors of phenylalanine and tyrosine can now be directly applied to in-cell protein overexpression systems without the need of laborious ^{15}N-introduction and additional assembly of a chiral C_α using synthetic organic chemistry. As a first example of ring-labelled aromatic precursors, we prepared sodium [3,3-$^{2}H_2$] ([3,5-$^{13}C_2$; 2,4,6-$^{2}H_3$] phenyl) pyruvate **7** using the cheap isotope sources [1,3-$^{13}C_2$] acetone and $^{2}H_2O$ (for NMR characterization see supporting information). Figure 4a demonstrates that different isotope patterns are

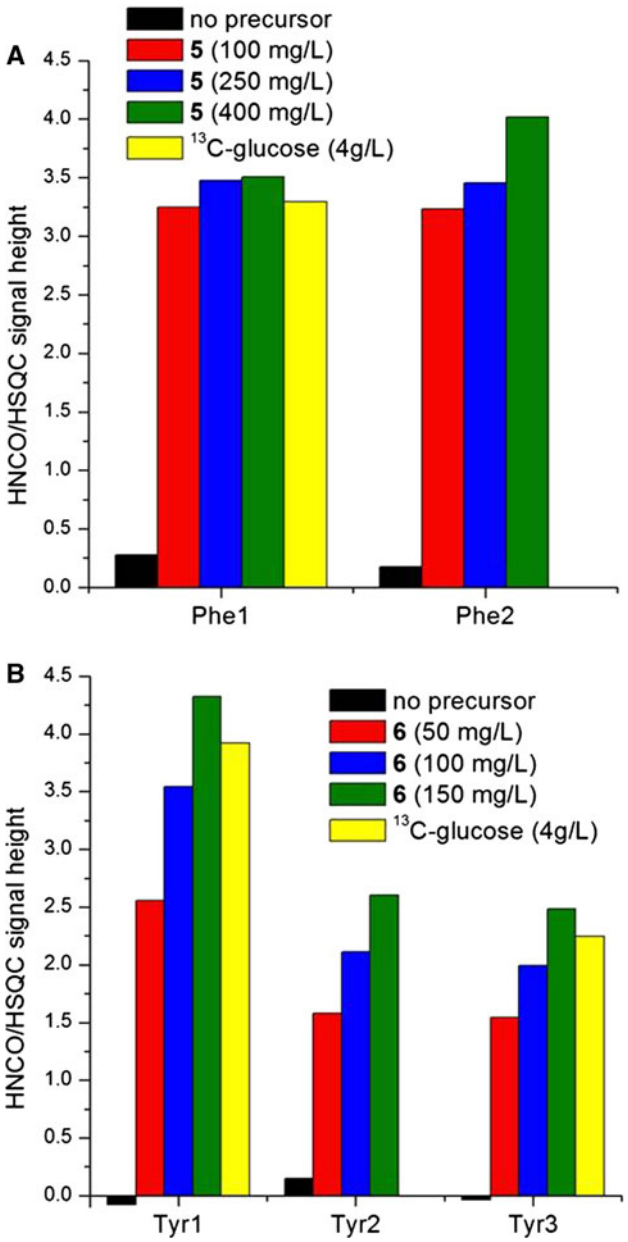


Fig. 2 HNCO/HSQC signal intensities of the His-tag GB1 protein, overexpressed in presence of different concentrations of sodium [1-^{13}C] phenylpyruvate **5** (**a**) and sodium [1-^{13}C] 4-hydroxyphenylpyruvate **6** (**b**). The signal intensities for a His-tag GB1 sample expressed in presence of $^{13}C_6$ glucose (4 g/L) are shown in yellow. In the case of Phe2 and Tyr2 no fully resolved signal resulting from this *all*-^{13}C model protein was available for signal intensity determination

accessible by this approach. We built the aromatic system by reacting acetone (commercially available in the [1, 3-$^{13}C_2$], [2-^{13}C] and [$^{13}C_3$]-form) with sodium nitromalonaldehyde (Akira et al. 1995). Deuteration at the aromatic positions is possible independently from backbone deuteration. Moreover, ^{13}C can be introduced at position 1 and 2 of compound **7** using labelled glycine (commercially available in all possible ^{13}C-patterns). His-tag GB1 was

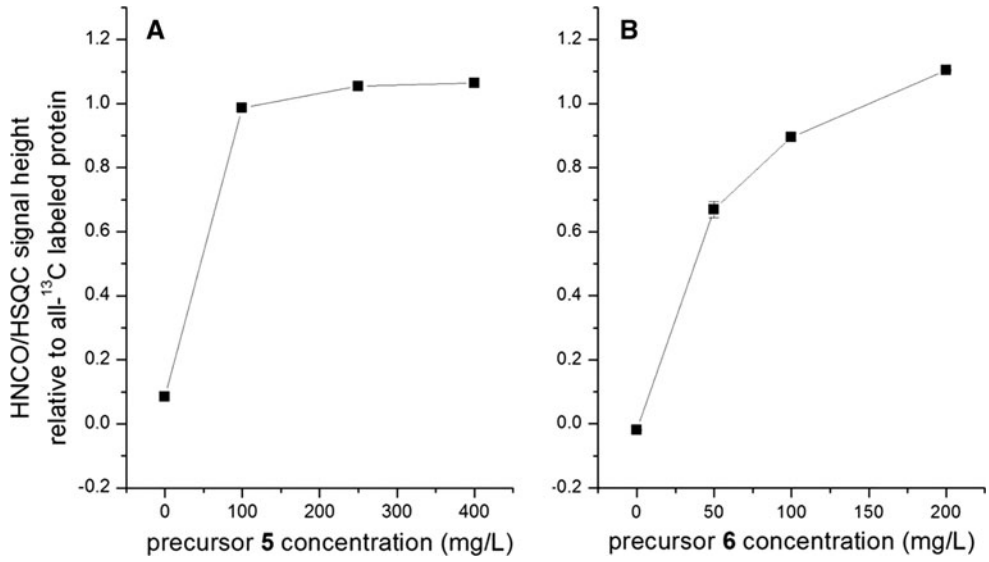


Fig. 3 HNCO/HSQC signal intensity of Phe1 (**a**) and Tyr1/Tyr3 (**b**) relative to *all*-^{13}C His-tag GB1 signals as a function of the precursor concentration in the growth medium

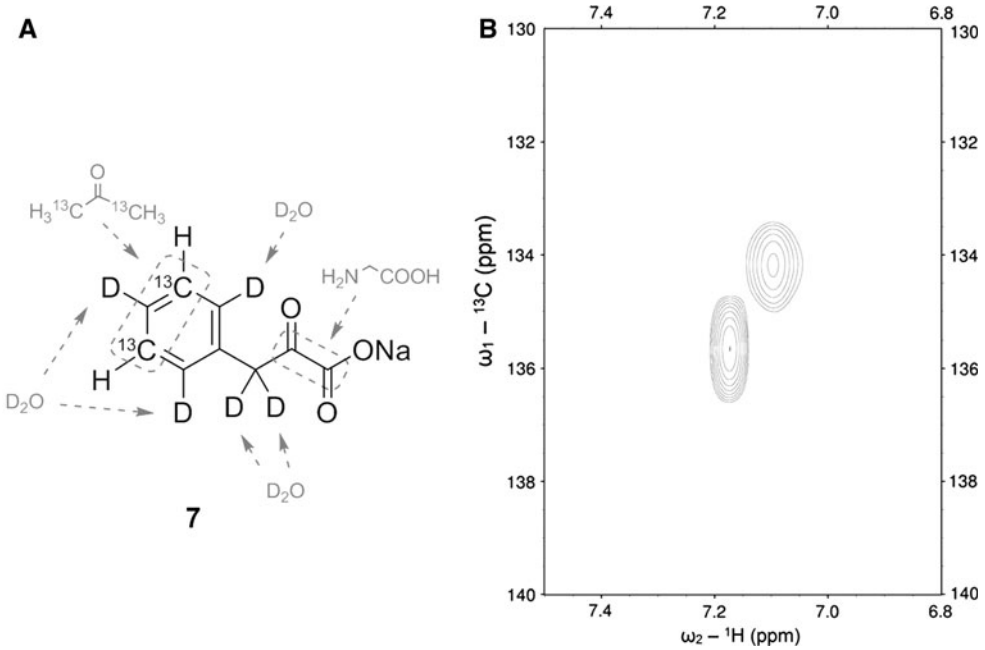


Fig. 4 **a** Structure of ring-labelled sodium phenylpyruvate **7**; the isotope sources [1,3-$^{13}C_2$] acetone and D_2O, as well as the potential ^{13}C-source glycine are shown in *grey*; **b** ^{13}C-HSQC of His-tag GB1 overexpressed in presence of compound **7**

overexpressed in 250 mL of minimal medium similar to the protocol described above for selective backbone labelling but containing 25 mg of precursor **7**. The resulting ^{13}C–HSQC spectra display selective incorporation of the ring labelled aromatic α-ketoacid (Fig. 4b). Compound **7** is highly expected to find extensive applicability in protein NMR, since this precursor enables the installation of isolated ^{13}C–^{1}H spin systems in a highly deuterated environment for detecting spin-relaxation rates (Kasinath et al. 2013) and features alternating ^{13}C–C–^{13}C ring patterns to reduce ^{13}C–^{13}C scalar coupling (Takeda et al. 2010). The synthesis of ring-labelled aromatic precursor compounds is currently ongoing in our laboratories. Therefore, details concerning the synthesis of precursor **7** will be published in due course, together with the preparation of alternative isotope patterns of phenylpyruvate and 4-hydroxyphenylpyruvate.

Acknowledgments This work was supported in part by the FWF (P20549-N19 and W-1221-B03 to R.K.).

References

Akira K, Hasegawa H, Baba S (1995) Synthesis of selectively ^{13}C labelled benzoic acid for nuclear magnetic resonance. J Label Compd Radiopharm 36:845–853

Ayala I, Sounier R, Usé N, Gans P, Boisbouvier J (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. J Biomol NMR 43:111–119

Billek G (1961a) Eine neue Synthese der 4-Hydroxyphenylbrenztraubensäure. Monatsh Chem 92:335–342

Billek G (1961b) Zur Synthese der Phenylbrenztraubensäuren, 3. Mitt Monatsh Chem 92:343–351

Billek G (1961c) Über die Kondensation aromatischer Aldehyde mit Hydantoin. Monatsh Chem 92:352–360

Bogan AA, Thorn KS (1998) Anatomy of hot spots in protein interfaces. J Mol Biol 280:1–9

Fischer M, Kloiber K, Hauesler J, Ledolter K, Konrat R, Schmid W (2007) Synthesis of a 13C-methyl-group-labeled methionine precursor as a useful tool for simplifying protein structural analysis by NMR spectroscopy. ChemBioChem 8:610–612

Goto NK, Gardner KH, Mueller GA, Willis KC, Kay LE (1999) A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated 15 N-, 13C-, 2H-labeled proteins. J Biomol NMR 13:369–374

Hoogstraten CG, Johnson JE Jr (2008) Metabolic labeling: taking advantage of bacterial pathways to prepare spectroscopically useful isotope patterns in proteins and nucleic acids. Concepts Magn Reson A 32A:34–55

Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Ono AM, Güntert P (2006) Optimal isotope labelling for NMR protein structure determinations. Nature 440:52–57

Kasinath V, Valentine KG, Wand AJ (2013) A 13C labeling strategy reveals a range of aromatic side chain motion in calmodulin. J Am Chem Soc 135:9560–9563

Lee KM, Androphy EJ, Baleja JD (1995) A novel method for selective isotope labeling of bacterially expressed proteins. J Biomol NMR 5:93–96

LeMaster DM, Kushlan DM (1996) Dynamical mapping of *E. coli* thioredoxin via ^{13}C NMR relaxation analysis. J Am Chem Soc 118:9255–9264

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

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

Lin Z, Xu Y, Yang S, Yang D (2006) Sequence-specific assignment of aromatic resonances of uniformly ^{13}C,^{15}N-labeled proteins by using ^{13}C- and ^{15}N-edited NOESY spectra. Angew Chem 118:1994–1997

Löhr F, Hänsel R, Rogov VV, Dötsch V (2007) Improved pulse sequences for sequence specific assignment of aromatic proton resonances in proteins. J Biomol NMR 37:205–224

Raap J, Niewenhuis S, Creemers A, Hexspoor S, Kragl U, Lugtenburg J (1999) Synthesis of Isotopically Labelled L-Phenylalanine and L-Tyrosine. Eur J Org Chem 2609–2621

Rajesh S, Nietlispach D, Nakayama H, Takio K, Laue ED, Shibata T, Ito Y (2003) A novel method for the biosynthesis of deuterated proteins with selective protonation at the aromatic rings of Phe, Tyr and Trp. J Biomol NMR 27:81–86

Rasia RM, Brutscher B, Plevin MJ (2012) Selective isotopic unlabeling of proteins using metabolic precursors: application to NMR assignment of intrinsically disordered proteins. ChemBioChem 13:732–739

Slupsky CM, Gentile LN, McIntosh LP (1998) Assigning the NMR spectra of aromatic amino acids in proteins: analysis of two Ets pointed domains. Biochem Cell Biol 76:379–390

Staunton D, Schlinkert R, Zanetti G, Colebrook SA, Campbell ID (2006) Cell-free expression and selective isotope labelling in protein NMR. Magn Reson Chem 44:S2–S9

Takeda M, Ono AM, Terauchi T, Kainosho M (2010) Application of SAIL phenylalanine and tyrosine with alternative isotope-labeling patterns for protein structure determination. J Biomol NMR 46:45–49

Teilum K, Brath U, Lundström P, Akke M (2006) Biosynthetic ^{13}C labeling of aromatic side chains in proteins for NMR relaxation measurements. J Am Chem Soc 128:2506–2507

Torizawa T, Ono AM, Terauchi T, Kainosho M (2005) NMR assignment methods for the aromatic ring resonances of phenylalanine and tyrosine residues in proteins. J Am Chem Soc 127:12620–12626

Wang H, Janowick DA, Schkeryantz JM, Liu X, Fesik SWJ (1999) A method of assigning phenylalanines in proteins. J Am Chem Soc 121:1611–1612