J Biomol NMR (2013) 57:205–209
DOI 10.1007/s10858-013-9786-y

COMMUNICATION

Independent valine and leucine isotope labeling in *Escherichia coli* protein overexpression systems

Roman J. Lichtenecker · Katharina Weinhäupl · Lukas Reuther · Julia Schörghuber · Walther Schmid · Robert Konrat

Received: 6 September 2013 / Accepted: 20 September 2013 / Published online: 28 September 2013

Abstract The addition of labeled α-ketoisovalerate to the growth medium of a protein-expressing host organism has evolved into a versatile tool to achieve concomitant incorporation of specific isotopes into valine- and leucine-residues. The resulting target proteins represent excellent probes for protein NMR analysis. However, as the side-chain resonances of these residues emerge in a narrow spectral range, signal overlap represents a severe limitation in the case of high-molecular-weight NMR probes. We present a protocol to eliminate leucine labeling by supplying the medium with unlabeled α-ketoisocaproate. The resulting spectra of a model protein exclusively feature valine signals of increased intensity, confirming the method to be a first example of independent valine and leucine labeling employing α-ketoacid precursor compounds.

Keywords Isotope labeling · Protein NMR · Protein overexpression · Leucine · Valine

NMR correlations of side-chain methyl groups represent valuable reporters for protein structure determination, due to favorable relaxation properties and high abundance in hydrophobic cores and on interaction surfaces of proteins (Tugarinov and Kay 2005a). Progress in NMR methodology led to a considerable shift of the upper size limit of NMR-attainable protein structures, especially with the advent of heteronuclear-edited NOESY techniques (Morshauser and Zuiderweg 1999; Diercks et al. 1999) or the methyl transverse relaxation optimized spectroscopy (methyl-TROSY; Ollerenshaw et al. 2003). Concomitant to novel NMR-techniques, robust protocols for selective labeling of isoleucines, valines, leucines, alanines and methionines using cell-based expression systems have been designed and were summarized previously (Ohki and Kainosho 2008; Ruschak and Kay 2010). These methods have to fulfil the criteria of providing defined isotope patterns at distinct atomic positions of the macromolecular assembly to suit the NMR experiment applied. At the same time, these stable isotopes should be implemented with reasonable synthetic and economic efforts. A straightforward approach to generate methyl labeled proteins is given by the use of an overexpressing, plasmid transformed *Escherichia coli* strain. The bacteria can be grown in minimal medium, supplemented with labeled amino acids or metabolic precursors thereof (Hoogstraten and Johnson 2008). The application of α-ketoacids in this context avoids the costly challenge of isotope labeled amino acid synthesis and thus the introduction of the stereocenter at the C_α position using synthetic chemistry. This approach additionally enables facile introduction of nitrogen-15 by using ^{15}N-ammonium chloride as a source of label. Combinations of 2-ketobutyrate and 2-ketoisovalerate as central metabolites in the biosynthesis of isoleucine, leucine and valine have been shown to yield appropriate isotope methyl patterns (often referred to as ILV-labeling; Tugarinov et al. 2006; Sprangers et al. 2007). Synthetic routes to access various labeled 2-ketobutyrates and 2-ketoisovalerates have been published (Gardner and

Electronic supplementary material The online version of this article (doi:10.1007/s10858-013-9786-y) contains supplementary material, which is available to authorized users.

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

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

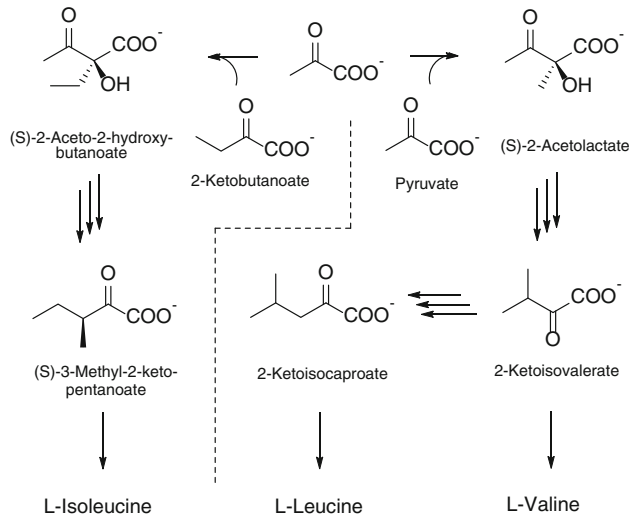

Fig. 1 Precursor compounds in ILV-labeling using *E. coli* expression systems

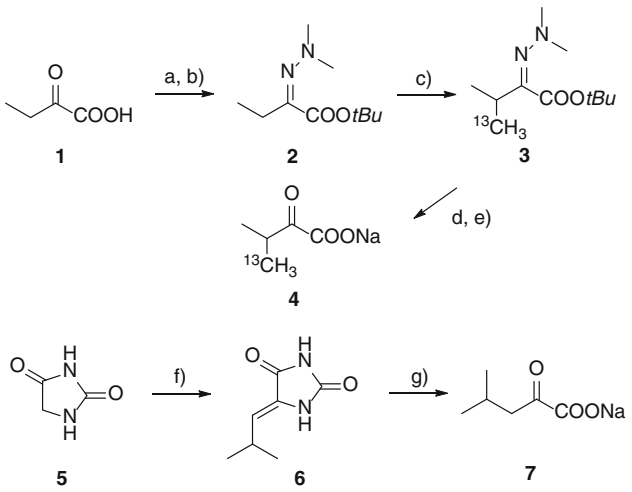

Fig. 2 Synthesis of sodium-($^{13}CH_3$)-2-ketoisovalerate **4** and sodium-2-ketoisocaproate **7**. Conditions *a* *t*BuOH, methanesulfonyl chloride, pyridine, THF, −5 °C − RT, 12 h, 64 %, *b* *N*,*N*′-dimethylhydrazine, diethyl ether, 12 h, 73 %, *c* lithium diisopropylamide, −78 °C, 1 h, then ^{13}C-methyliodide, THF, 2 h, 83 %, *d* 1N HCl, THF, 3 h, 77 %, *e* HCl_g, Et_2O/CH_2Cl_2 1:1, 0 °C − RT, then 1 M NaOH until pH = 8, lyophilize, 95 %, *f* ethanolamine, isobutyraldehyde, H_2O, 110 °C, 9 h, 74 %, *g* NaOH (20 %), 100 °C, 5 h, 81 %

Kay 1997; Hajduk et al. 2000; Gross et al. 2003; Lichtenecker et al. 2004; Schedlbauer et al. 2008) and some of them are commercially available. ILV-labeling has been successfully applied to study very large proteins or protein complexes, including the malate synthase G (82 kDa) (Tugarinov and Kay 2005b), the aspartate transcarbamoylase (306 kDa) (Velyvis et al. 2007), the 20S core particle proteasome (670 kDa) (Sprangers and Kay 2007) or the translocase motor SecA (204 kDa) (Gelis et al. 2007). Despite these impressive applications, extensive signal overlap and restricted sensitivity remain the most profound challenges in studies of high molecular weight NMR probes. The limited resolution of complex spectra often results in incomplete signal assignment and a significant loss of structural restraints.

Figure 1 gives an overview of the main metabolic intermediates in ILV-labeling. Selective isoleucine labeling can easily be realized using either 2-ketobutyrate (Goto et al. 1999) or (*S*)-2-aceto-2-hydroxybutanoate (Ruschak et al. 2010; Ayala et al. 2012) as unique precursor compounds. On the other hand, leucines and valines are labeled simultaneously in the presence of 2-ketoisovalerate (Goto et al. 1999) or (*S*)-2-acetolactate (Gans et al. 2010). Separation of Val- and Leu-residue methyl resonances has been realized by suitable NMR experiments which allow for selective residue type editing (Hu et al. 2012). We implied a more general approach based on the introduction of 2-ketoisocaproate as a new precursor for selective leucine labeling. This methodology was successfully applied and resulted in selective labeling without interfering with the valine metabolic pathway (Lichtenecker et al. 2013). In the present study, we have further developed our approach of independent ILV-labeling to exclusive incorporation of isotopes into valine residues by applying the well-established 2-ketoisovalerate as a metabolic precursor in the presence of unlabeled 2-ketoisocaproate to saturate the metabolic branch of leucine biosynthesis. To validate our hypothesis of independent isotope incorporation into aliphatic residues we used a His-Tag GB1 66 amino acid protein as a model system and tested the appropriate quantities of additives in the *E. coli* growth media with the intention to maximize valine-, while at the same time repressing leucine-labeling.

Sodium-($^{13}CH_3$)-2-ketoisovalerate **4** was prepared similar to the synthetic routes described previously (Fig. 2; Hajduk et al. 2000; Lichtenecker et al. 2004). As a first step, α-ketobutanoic acid **1** was converted to the corresponding dimethylhydrazone-*tert*-butyl ester **2**. Subsequent methylation using iodo(^{13}C)methane gave compound **3**, which was treated with 1N HCl to cleave the hydrazone to the α-ketoester. Final hydrolysis was performed with gaseous HCl, followed by lyophilization from an aqueous solution (pH = 8) to a white powder of product **4** in a total yield of 28 % from **1** and 40 % from $^{13}CH_3I$. Sodium 2-ketoisocaproate **7** was synthesized as reported (Lichtenecker et al. 2013) via hydrolysis of isobutylidenehydantoin **6** in a total yield of 60 %.

His-tagged-GB1 protein was expressed and purified according to a protocol described previously (Lichtenecker et al. 2013). In brief, BL21-pLysS cells, containing the corresponding expression plasmid were grown in LB medium at 37 °C. Dilution in $^{15}NH_4Cl$ (1 g/L) and glucose (4 g/L) supplemented M9 medium (250 mL) was followed by addition of sodium-($^{13}CH_3$)-2-ketoisovalerate **4** (100 mg/L)

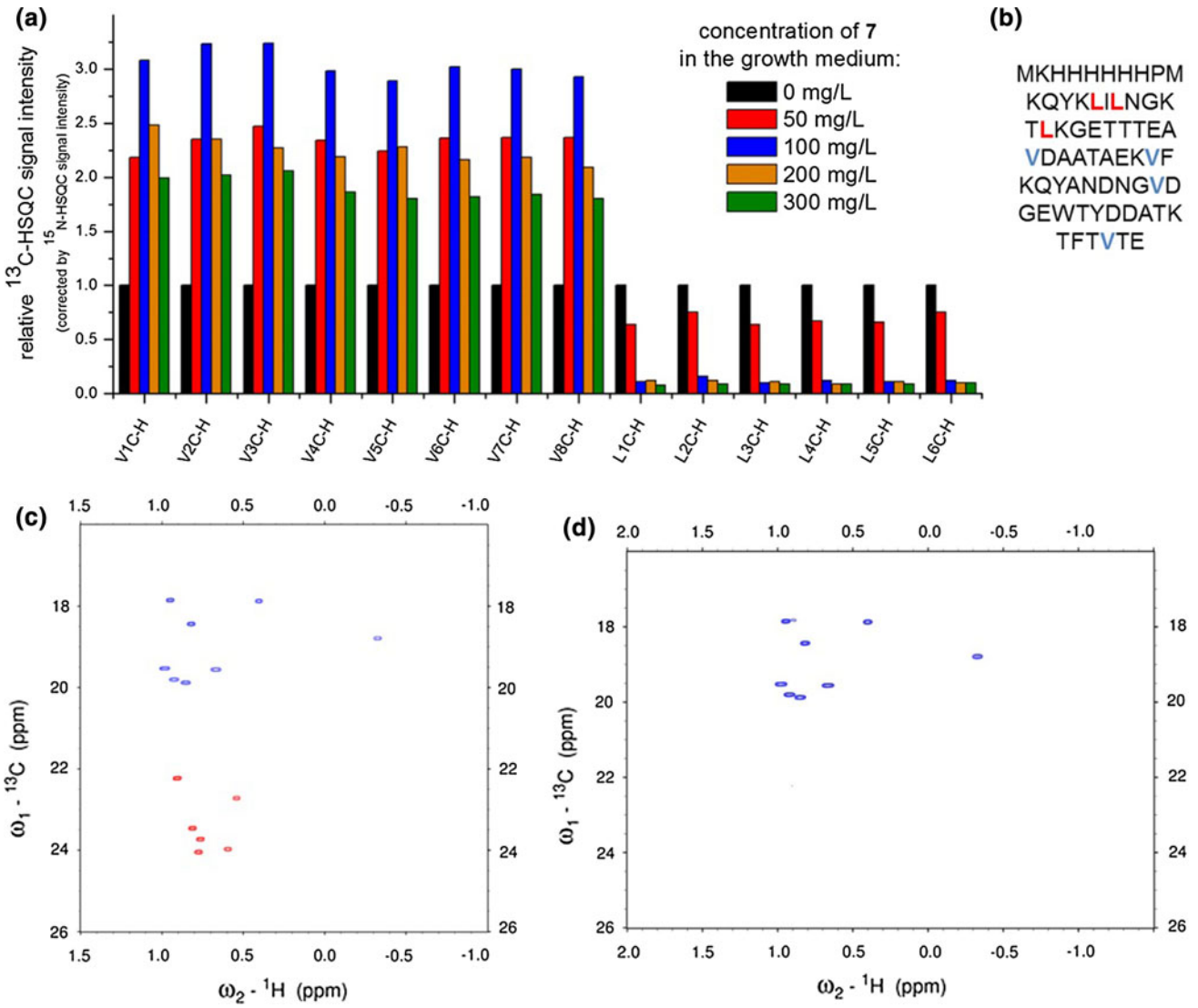

Fig. 3 **a** Relative ^{13}C-HSQC signal intensities of valine—(V1–V8) and leucine methyl groups (L1–L6) of His-tagged GB1 expressed in presence of sodium-($^{13}CH_3$)-2-ketoisovalerate **4** (100 mg/L) and different concentrations of sodium-2-ketoisocaproate **7** (0–300 mg/L). **b** Sequence of His-tag-GB1. **c** ^{13}C-HSQC of His-tag-GB1 expressed in presence of sodium-($^{13}CH_3$)-2-ketoisovalerate **4** (100 mg/L) showing valine (*blue*) and leucine (*red*) residues. **d** ^{13}C-HSQC of His-tag-GB1 expressed in presence of sodium-($^{13}CH_3$)-2-ketoisovalerate **4** (100 mg/L) and sodium-2-ketoisocaproate **7** (100 mg/L) exclusively showing valine residues

and sodium-2-ketoisocaproate **7** (0, 50, 100, 200 and 300 mg/L) 1 h prior to induction. Addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.8 mM at an OD_{600} of 0.8 induced protein overexpression. After 16 h of expression at 30 °C, the cells were collected by centrifugation at 6,000 rpm for 15 min and suspended in ice-cold PBS. Subsequently, the cells were passed through a French press and the cell lysate was cleared by centrifugation at 18,000 rpm for 30 min. Purification was performed by loading the supernatant onto a Ni^{2+} HiTrap 5 mL affinity column, washed with PBS and eluted with high imidazole buffer (0.5 M). The protein containing fractions were collected, concentrated and further purified by gel filtration over a Superdex75 column. ^{15}N-HSQC and ^{13}C-HSQC experiments were measured at a protein concentration of 0.5 mM in 20 mM Tris-buffer (pH 7.4, 50 mM NaCl, supplemented with 10 % ^2H_2O). In order to compare ^{13}C incorporation levels, the ^{13}C-HSQC signal intensities were divided by the sum of 10 signal intensities of the corresponding ^{15}N-HSQC spectrum.

Figure 3 shows the ^{13}C-HSQC intensities of the 66 amino acid model peptide His-tagged GB1, expressed in presence of sodium-($^{13}CH_3$)-2-ketoisovalerate **4** (100 mg/L) and increasing amounts of sodium-2-ketoisocaproate **7** (50-300 mg/L) relative to the control protein, which was grown in a medium devoid of unlabeled compound **7** (Fig. 3a black). The results depicted in Fig. 3a evidence decreasing Leu-signal intensities as a function of the precursor **7**

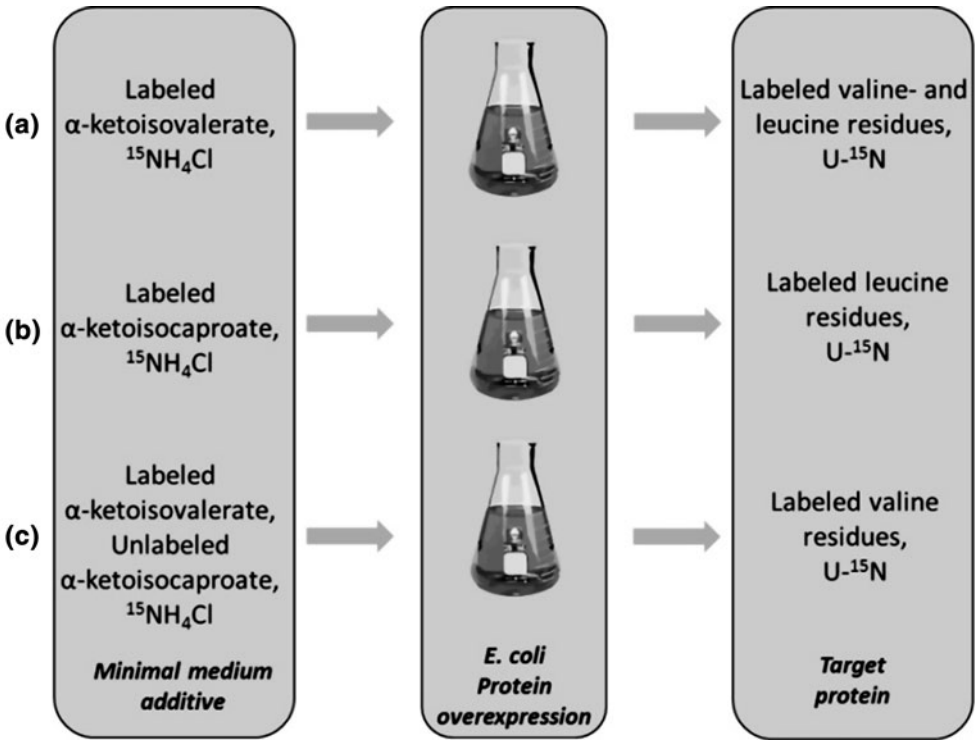

Fig. 4 Approaches for Leu/Val-labeling. *a* Concomitant Leu/Val labeling, *b* selective Leu labeling, *c* selective Val labeling

concentration in the overexpression medium. Complete extinction of the methyl signals resulting from the three leucines present in the protein sequence was reached at 100 mg/L of sodium-2-ketoisocaproate **7** (Fig. 3c, d). Under these conditions, a significant increase of the valine-methyl-resonance intensities was observed, proving the transformation of sodium-($^{13}CH_3$)-2-ketoisovalerate **4** to Val-residues to be more effective in presence of unlabeled sodium 2-ketoisocaproate **7** (Fig. 3a). Higher concentrations of unlabeled 2-ketoisocaproate (>100 mg/L) led to a slight decay of valine signal intensity. Therefore, the optimal ratio of precursor compounds in the minimal medium was identified as a 1:1 mixture (100 mg/L each) since these conditions resulted in complete deletion of leucine signals and a maximum increase of valine signal intensities.

When it comes to large proteins, a high deuteration grade may be required in order to decrease transverse relaxation rates and improve sensitivity. In this case, fully ^{2}H-labeled 2-ketoisocaproate may be applied. This deuterated precursor is synthetically accessible via a route similar to the one we described recently (Lichtenecker et al. 2013). In summary, we could show that addition of unlabeled sodium 2-ketoisocaproate **7** to an *E. coli* overexpression medium efficiently saturates the leucine metabolic pathway. The metabolic conversion of labeled sodium 2-ketoisovalerate **4** to leucine was thus blocked, leading to exclusive valine labeling from precursor **4**. At the same time, signal intensities derived from valine-methyl groups were significantly increased. These results complete our efforts of separating valine and leucine labeling, which are summarized in Fig. 4. Protein overexpression in presence of labeled 2-ketoisovalerates leads to concomitant labeling of Val- and Leu-residues (Fig. 4a; Goto et al. 1999; Lichtenecker et al. 2004). Labeled 2-ketoisocaproates can be used as selective leucine precursors without any cross-labeling to valine residues (Fig. 4b; Lichtenecker et al. 2013). Finally, a combination of labeled 2-ketoisovalerates and unlabeled 2-ketoisocaproates results in exclusive valine labeling (Fig. 4c). We want to point out that in using our methodology different isotope patterns for Val- and Leu-residues in the same protein sample become available for the first time. We are currently working on patterns that will allow for the discrimination of valine-derived through-space connectivities from those derived from leucines. The resulting spectra simplification will increase the number of available distance restraints in the case of large proteins significantly.

Acknowledgments This work was supported by grants from the Austrian Science foundation FWF P20549, P22125 and W1221B03.

References

Ayala I, Hamelin O, Amero C, Pessey O, Plevin MJ, Gans P, Boisbouvier J (2012) An optimized isotope labelling strategy of isoleucine-γ2 methyl groups for solution NMR studies of high molecular weight proteins. Chem Commun 48:1434–1436

Diercks T, Coles M, Kessler H (1999) An efficient strategy for assignment of cross-peaks in 3D heteronuclear NOESY experiments. J Biomol NMR 15:177–180

Gans P, Hamelin O, Sounier R, Ayala I, Asucion Dura M, Amero CD, Noirclerc-Savoye M, Franzetti B, Plevin MJ, Boisbouvier J (2010) Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. Angew Int Ed 49:1958–1962

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 119:7599–7600

Gelis I, Bonvin AM, Keramisanou D, Koukaki M, Gouridis G, Karamanou S, Economou A, Kalodimos CG (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. Cell 131:756–769

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

Gross JD, Gelev VM, Wagner G (2003) A sensitive and robust method for obtaining intermolecular NOEs between side chains in large protein complexes. J Biomol NMR 25:235–242

Hajduk PJ, Augeri DJ, Mack J, Mendoza R, Yang JG, Betz SF, Fesik SW (2000) NMR-based screening of proteins containing ^{13}C-labelled methyl groups. J Am Chem Soc 122:7898–7904

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

Hu W, Namanja AT, Wong S, Chen Y (2012) Selective editing of Val and Leu methyl groups in high molecular weight protein NMR. J Biomol NMR 53:113–124

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

Morshauser RC, Zuiderweg ERP (1999) High-resolution four-dimensional HMQC-NOESY-HSQC spectroscopy. J Magn Reson 139:232–239

Ohki S, Kainosho M (2008) Stable isotope labeling methods for protein NMR spectroscopy. Prog Nucl Magn Reson Spectrosc 53:208–226

Ollerenshaw JE, Tugarinov V, Kay LE (2003) Methyl TROSY: explanation and experimental verification. Magn Reson Chem 41:843–852

Ruschak A, Kay LE (2010) Methyl groups as probes of supramolecular structure, dynamics and function. J Biomol NMR 46:75–87

Ruschak AM, Velyvis A, Kay LE (2010) A simple strategy for ^{13}C, ^{1}H labeling at the Ile-γ2 methyl position in highly deuterated proteins. J Biomol NMR 48:129–135

Schedlbauer A, Auer R, Ledolter R, Tollinger M, Kloiber K, Lichtenecker R, Ruedisser S, Hommel U, Schmid W, Konrat R (2008) Direct methods and residue type specific isotope labeling in NMR structure determination and model-driven sequential assignment. J Biomol NMR 42:111–127

Sprangers R, Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature 445:618–622

Sprangers R, Velyvis A, Kay LE (2007) Solution NMR of supramolecular complexes: providing new insights into function. Nat Methods 4:697–703

Tugarinov V, Kay LE (2005a) Methyl groups as probes of structure and dynamics in NMR studies of high-molecular-weight proteins. ChemBioChem 6:1567–1577

Tugarinov V, Kay LE (2005b) Quantitative ^{13}C and ^{2}H NMR relaxation studies of the 723-residue enzyme malate synthase G reveal a dynamic binding interface. Biochemistry 44:15970–15977

Tugarinov V, Kanelis V, Kay LE (2006) Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. Nat Protoc 1:749–754

Velyvis A, Yang YR, Schachmann HK, Kay LE (2007) A solution NMR study showing that active site ligands and nucleotides directly perturb the allosteric equilibrium in aspartate transcarbamoylase. Proc Natl Acad Sci USA 104:8815–8820