

A NMR experiment for simultaneous correlations of valine and leucine/isoleucine methyls with carbonyl chemical shifts in proteins

Vitali Tugarinov · Vincenzo Venditti ·
G. Marius Clore

Received: 8 November 2013 / Accepted: 6 December 2013 / Published online: 18 December 2013
© Springer Science+Business Media Dordrecht (outside the USA) 2013

Abstract A methyl-detected ‘out-and-back’ NMR experiment for obtaining simultaneous correlations of methyl resonances of valine and isoleucine/leucine residues with backbone carbonyl chemical shifts, SIM-HMCM(CGBCA)CO, is described. The developed pulse-scheme serves the purpose of convenience in recording a single data set for all Ile^{δ1}, Leu^δ and Val^γ (ILV) methyl positions instead of acquiring two separate spectra selective for valine or leucine/isoleucine residues. The SIM-HMCM(CGBCA)CO experiment can be used for ILV methyl assignments in moderately sized protein systems (up to ~100 kDa) where the backbone chemical shifts of ¹³C^α, ¹³C_β and ¹³CO are known from prior NMR studies and where some losses in sensitivity can be tolerated for the sake of an overall reduction in NMR acquisition time.

Keywords Chemical shift assignments · Methyl labeling · Enzyme I · Methyl TROSY · Isotope shifts

Selective ¹³CH₃ labeling of Ile^{δ1}, Leu^δ and Val^γ (ILV) methyl positions on a deuterated background (Goto et al. 1999; Ruschak and Kay 2010; Sheppard et al. 2010; Tugarinov et al. 2006; Tugarinov and Kay 2004) in combination

with Methyl-TROSY (Amero et al. 2009; Ollerenshaw et al. 2003; Tugarinov et al. 2003) have had a profound impact on NMR studies of structure and dynamics in high-molecular-weight proteins (Gelis et al. 2007; Hamel and Dahlquist 2005; Religa et al. 2010; Rosenzweig et al. 2013; Ruschak et al. 2010; Sprangers and Kay 2007; Sprangers et al. 2007; Tugarinov et al. 2004; Velyvis et al. 2007). The success of NMR studies that make use of the ILV labeling methodology is predicated upon the availability of methyl resonance assignments. While a variety of approaches have been proposed for assignment of methyl resonances in very large (>>100-kDa) proteins and protein complexes over the past decade (John et al. 2007; Sprangers and Kay 2007; Sprangers et al. 2007; Velyvis et al. 2009; Venditti et al. 2011; Xu et al. 2009), methyl-detected ‘out-and-back’ NMR experiments that correlate the (¹H, ¹³C) chemical shifts of ILV methyl groups with the backbone and ¹³C_β carbon positions (Sheppard et al. 2009a, b; Tugarinov et al. 2004; Tugarinov and Kay 2003), provide the most sensitive and robust methodology for unambiguous methyl assignments in medium-sized (<~100 kDa) protein systems without recourse to site-directed mutagenesis and/or analysis of NOE data. These experiments benefit from isotope labeling strategies that ‘linearize’ the ¹³C spin-systems of Leu and Val side-chains via the use of α-keto-acid (Tugarinov et al. 2006) or aceto-lactate (Gans et al. 2010) biosynthetic precursors that have only one of their methyl positions labeled with ¹³CH₃, while the other remains at natural abundance and deuterated (¹²CD₃). The ‘linearization’ of carbon spin-systems of (branched) LV side-chains has been shown to simplify NMR spectra and provide significant sensitivity gains in methyl resonance assignment experiments (Tugarinov and Kay 2003).

Commonly, the HMCM[CG]BCA experiment that correlates methyl chemical shifts with the chemical shifts of

Electronic supplementary material The online version of this article (doi:10.1007/s10858-013-9803-1) contains supplementary material, which is available to authorized users.

V. Tugarinov (✉) · V. Venditti · G. Marius Clore (✉)
Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, USA
e-mail: vitali.tugarinov@nih.gov

G. Marius Clore
e-mail: mariusc@intra.niddk.nih.gov; mariusc@mail.nih.gov

$^{13}\text{C}_\alpha/^{13}\text{C}_\beta$ nuclei (Tugarinov and Kay 2003) is sufficient for reliable assignments of ILV methyls in proteins with a moderate content of ILV methyl groups. The degeneracy of $^{13}\text{C}_\alpha/^{13}\text{C}_\beta$ pairs of chemical shifts that inevitably occurs for some residues, even in medium-sized proteins, can be resolved by recording spectra that correlate ILV methyl resonances with carbonyl chemical shifts: one experiment selective for Ile and Leu, ILE/LEU-HMCM(CGCBCA)CO, and a separate data set for Val, VAL-HMCM(CBCA)CO (Tugarinov and Kay 2003). Here, we describe a three-dimensional methyl-detected ‘out-and-back’ NMR experiment developed for simultaneously correlating the methyl resonances of Val and Ile/Leu residues with the backbone carbonyl chemical shifts, SIM-HMCM(CGCBCA)CO. The pulse scheme provides the convenience of acquiring only a single data set for all ILV methyl positions instead of two separate spectra for Val and Ile/Leu residues. The SIM-HMCM(CGCBCA)CO experiment can be used for ILV methyl assignments in medium-sized proteins (up to ~ 100 kDa) where the backbone chemical shifts of $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$ and ^{13}CO are known from prior NMR studies and where some losses in sensitivity can be tolerated for the sake of overall reduction in NMR acquisition time. We primarily address the problem of how carbon magnetization in homonuclear ^{13}C spin-systems of different topologies (as in Val vs. Ile/Leu) can be manipulated to obtain the desired correlations for all ILV positions with minimal losses in sensitivity. As a corollary result, we report the assignments of ILV methyl resonances for the 70-kDa homo-dimeric C-terminal domain of *E. coli* Enzyme I (EIC) (Venditti et al. 2012) obtained via combined analysis of methyl-detected ‘out-and-back’ HMCM[CG]CBCA and SIM-HMCM(CGCBCA)CO data sets.

Despite the fact that Val and Ile/Leu side-chains have different topologies, careful manipulation of magnetization via application of selective pulses allows one to ‘store’ the magnetization terms of (shorter) Val side-chains to ensure the subsequent ‘read-out’ of carbonyl frequencies of Val in synchrony with that of Ile/Leu residues. In the following, we briefly describe how this can be achieved in practice.

Figure 1 shows the SIM-HMCM(CGCBCA)CO pulse-scheme that allows methyl-carbonyl correlations in Val and Ile/Leu residues to be obtained simultaneously. The pulse-scheme in Fig. 1 utilizes the Methyl-TROSY principle (Tugarinov et al. 2003) where possible by keeping the methyl magnetization in a ^1H – ^{13}C multiple-quantum state as long as it resides on methyl groups (i.e. before the $^{13}\text{C}_{\phi 2}/\text{H}_y$ pair of pulses and after the $^{13}\text{C}_{\phi 7}/\text{H}_y$ pair of pulses). At time-point *a* of the scheme, the generated magnetization terms are of the form $4H_z^m C_x^\beta C_z^\alpha$ for Val and $4H_z^m C_x^\gamma C_z^\beta$ for Ile/Leu residues, where C_r^i and $H_r^m = H_r^1 + H_r^2 + H_r^3$ are product operators for carbon nucleus *i* and the three methyl

proton nuclei, respectively, $r = x, y, z$, and the signs of the terms are arbitrary. Fortunately, the region of Val $^{13}\text{C}_\alpha$ chemical shifts is well separated from the rest of the carbon chemical shifts in the terms above. Therefore, a selective 90° pulse can be chosen that excites all carbon nuclei in these terms except for $^{13}\text{C}_\alpha$ of Val. Application of a universal Gaussian Cascade (Q5) pulse (Emsley and Bodenhausen 1987) with phase ‘y’ centered at 33 ppm with an excitation bandwidth of ± 12 ppm (600 MHz) converts the terms above to $4H_z^m C_z^\beta C_z^\alpha$ for Val and $4H_z^m C_z^\gamma C_x^\beta$ for Ile/Leu (time point *b*; Fig. 1). Thus, the magnetization of Val residues is effectively ‘stored’ in a triple-spin-order state that does not evolve due to ^{13}C – ^{13}C couplings during the following $4T_C$ period. Subsequently, the evolution of magnetization proceeds via the following pathways:

$$4H_z^m C_z^\beta C_z^\alpha(b) \xrightarrow{4T_C, 90^\circ_{C, \phi 2}} 4H_z^m C_x^\beta C_x^\alpha(c) \xrightarrow{4\tau_d, 90^\circ_{C, y}, 90^\circ_{CO, \phi 4}} 8H_z^m C_z^\beta C_y^\alpha C_y'(d) \quad (1)$$

for Val and

$$4H_z^m C_z^\gamma C_x^\beta(b) \xrightarrow{4T_C, 90^\circ_{C, \phi 2}} 4H_z^m C_z^\beta C_x^\alpha(c) \xrightarrow{4\tau_d, 90^\circ_{C, y}, 90^\circ_{CO, \phi 4}} 4H_z^m C_z^\alpha C_y'(d) \quad (2)$$

for Ile/Leu, where the letters in parentheses correspond to the time points in the pulse scheme (Fig. 1) and the signs of the terms are neglected. The $^{13}\text{C}_\beta$ – $^{13}\text{C}_\gamma$ couplings in Ile side-chains are refocused by application of a pair of ‘I γ 2’ pulses selective for the Ile $^{13}\text{C}_\gamma$ methyl region between time-points *b* and *c*, while evolution due to $^{13}\text{C}_\beta$ – $^{13}\text{C}_\gamma$ couplings in Val between time-points *c* and *d* is eliminated by application of a pair of ‘V γ ’ pulses selective for the Val $^{13}\text{C}_\gamma$ methyl region. The terms on the right-hand side of pathways (1) and (2) evolve during the t_1 period, with the chemical shifts of $^{13}\text{C}_\alpha$ nuclei of Val refocused by the ^{13}C 180° pulse ‘ α ’ applied in the middle of t_1 and selective for the $^{13}\text{C}_\alpha$ chemical shifts of all ILV residues (Fig. 1). The selectivity of this pulse also ensures that the $^{13}\text{C}_\alpha$ – $^{13}\text{C}_\beta$ *J* couplings in Val are refocused during the t_1 evolution period. After the terms generated at time point *d* of pathways (1) and (2) are labeled with ^{13}CO chemical shifts, the magnetization is transferred back to methyl carbons that evolve during the time t_2 , and subsequently to methyl protons for detection by reversing the pathways.

Figure 2 shows plots of the magnetization transfer function for Val methyl (green curve) and Ile/Leu methyl (red curve) (^1H , ^{13}C)– ^{13}CO correlations as a function of the delay $4\tau_d$ in the SIM-HMCM(CGCBCA)CO pulse scheme. These plots take into account the fact that Val signals relax faster during the $4\tau_d$ period than Ile/Leu signals as Val magnetization is present in a ^{13}C – ^{13}C multiple-quantum state, with both $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ nuclei relaxing in the

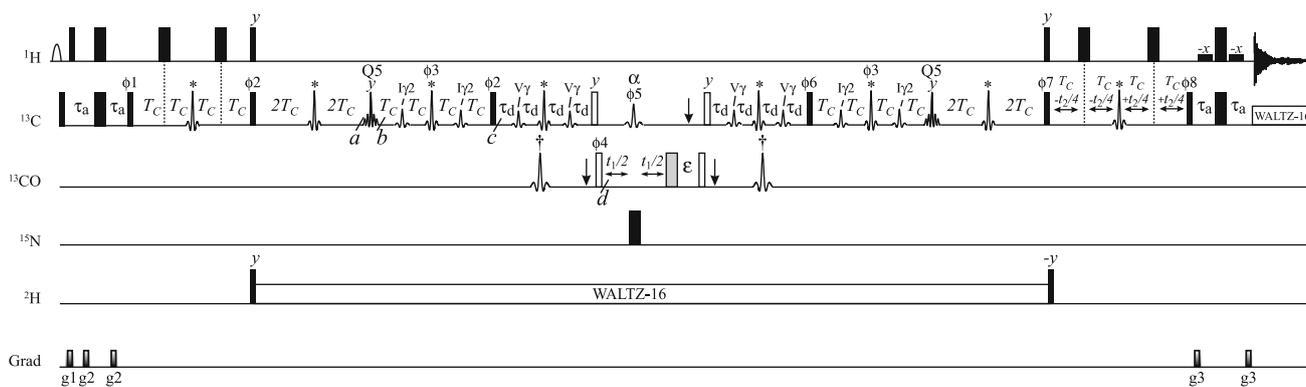


Fig. 1 The SIM-HMCM(CGBCA)CO pulse-scheme for simultaneous methyl–carbonyl correlations in Val and Ile/Leu residues. All narrow and wide rectangular pulses are applied with flip angles of 90° and 180°, respectively. All the pulses (including shaped pulses) are applied along the x-axis unless indicated otherwise. The ¹H, ²H and ¹⁵N carrier positions were set to 4.7, 2.5 and 119 ppm, respectively. The ¹³C carrier is positioned at 18 ppm at the beginning of the scheme, switched to 33 ppm before the ¹³C pulse with phase ϕ_1 , and returned to 18 ppm after the ¹³C pulse with phase ϕ_8 . All the pulses shown with black rectangles are applied with the highest possible power except for the ~1.4-ms long water-selective rectangular pulses flanking the last 180 ¹H pulse and the water-selective ~7.0-ms ¹H E-BURP1 pulse (Geen and Freeman 1991) shown with the open arc. ¹³C WALTZ-16 decoupling (Shaka et al. 1983) during acquisition is achieved using a 2.5 kHz field, while ²H WALTZ-16 decoupling uses a 0.8 kHz field. The ¹³C shaped pulses marked with asterisks are high-power 350- μ s (600 MHz) RE-BURP pulses (Geen and Freeman 1991) applied on-resonance. The ¹³C shaped pulses marked with daggers are high-power 350- μ s RE-BURP pulses applied at 176 ppm by phase modulation of the carrier (Boyd and Soffe 1989; Patt 1992). The ¹³C pulses marked with ‘I γ 2’ are 4.0 ms RE-BURP pulses selective for Ile ¹³C γ^2 positions and are applied at 16 ppm by phase-modulation of the carrier. The ¹³C pulses marked with ‘V γ ’ are 3.0 ms RE-BURP pulses selective for ¹³C γ methyls of valine residues and are applied at 21 ppm by phase modulation of the carrier. The ¹³C pulse labeled ‘ α ’ is a 1.8 ms RE-BURP pulse

(600 MHz) selective for ¹³C α positions of ILV residues and applied at 60 ppm by phase modulation of the carrier. The ¹³C pulses labeled ‘Q5’ are 1.4 ms Q5 Gaussian Cascade pulses (Emsley and Bodenhausen 1987) applied on-resonance (see text for details). ¹³C and ¹³CO 90° pulses shown with open rectangles are applied at 45 and 176 ppm, respectively, by phase modulation of the carrier with a field strength of $\Delta\sqrt{15}$ where Δ is the difference (in Hz) between ¹³C α/β (45 ppm) and ¹³CO (176 ppm) chemical shifts (Kay et al. 1990). The 180° ¹³CO pulse following the t_1 period shown with the shaded rectangle is applied at 176 ppm by phase modulation of the carrier with the same field strength as the 90° ‘open’ pulses. This ensures complete refocusing of ‘Bloch–Siegert’ shifts of both ¹³C α and ¹³CO nuclei during t_1 . Vertical arrows at the start and end of the $4\tau_d$ periods and after the t_1 period indicate the position of Bloch–Siegert compensation pulses (Kay et al. 1990). Delays are: $\tau_a = 2.0$ ms; $T_C = 3.5$ ms; $\tau_d = 2.6$ ms; the delay ϵ is adjusted to accommodate the shaped Bloch–Siegert compensation pulse ‘ α ’. The phase-cycle is: $\phi_1 = x, -x$; $\phi_2 = 2(y), 2(-y)$; $\phi_3 = x$; $\phi_4 = 2(x), 2(-x)$; $\phi_5 = 4(x), 4(-x)$; $\phi_6 = 4(y), 4(-y)$; $\phi_7 = 8(y), 8(-y)$; $\phi_8 = x$; receiver = $x, -x, -x, x$. Note that if the phase of one of the ¹³C ϕ_3 pulses is shifted by 90° ($\pm y$ instead of x), the signs of Ile/Leu correlations will be inverted while retaining the signs of Val correlations. Quadrature detection in F_1 and F_2 is achieved by States-TPPI (Marion et al. 1989) of ϕ_4 and ϕ_8 , respectively. Durations and strengths of the pulsed-field gradients in units of (ms; G/cm) are: $g_1 = (1.0; 15)$, $g_2 = (0.3; 8)$, $g_3 = (0.45; 10)$

transverse plane. This is also true for Val residues with respect to the t_1 evolution period where the multiple-quantum ¹³C α –¹³CO state relaxes. The duration of the τ_d delays should be chosen to ensure the desired compromise between the intensities of Val and Ile/Leu correlations ($\tau_d = 2.6$ ms in this work; Fig. 2), and the time $t_{1,max}$ should be adjusted with faster relaxation of Val correlations in mind ($t_{1,max} \approx 15$ ms at 600 MHz allows the observation of all Val correlations for EIC). The terms $8H_z^m C_x^\beta C_y^\alpha C'_y$ in Ile/Leu residues are omitted from the end of pathway (2) above but are not filtered out before the t_1 evolution period. These terms (blue curve in Fig. 2) are, however, much lower in intensity for the duration of τ_d chosen here, providing correlations are observed for only a few very intense peaks in the 3D spectra of EIC at the frequencies $\Omega_{CO} \pm \Omega_{C\beta}$, where Ω_i is the offset of nucleus i from the carbon carrier. We note that the same transfer function as shown for the $8H_z^m C_x^\beta C_y^\alpha C'_y$ terms of Ile/Leu in

Fig. 2 also applies for the correlations of Val residues in the ILE/LEU-HMCM(CGBCA)CO data sets optimized for observation of Ile, Leu methyl-(¹H, ¹³C)–¹³CO correlations. That is why the correlations of Val residues are either not observed or observed with low intensities in the ILE/LEU-HMCM(CGBCA)CO data sets.

Analysis of magnetization flow for the pulse-scheme in Fig. 1 shows that shifting the phase of one of the ¹³C 180° pulses with phase ϕ_3 (¹³C ϕ_3 ; in the time interval between points b and c in Fig. 1) by 90° inverts the sign of the correlations belonging to Ile/Leu residues while preserving the sign of Val correlations. Although this may seem as an especially attractive feature of the SIM-HMCM(CGBCA)CO experiment, in practice, Val residues are easily distinguished from Ile/Leu by the relative signs of the peaks obtained in the 3D HMCM[CG]CBCA data sets (Tugarinov and Kay 2003). Nevertheless, in the regions of the 3D HMCM[CG]CBCA spectra where Val correlations may overlap with Leu peaks

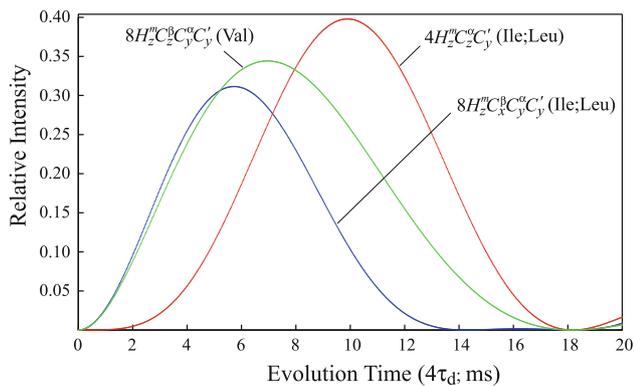


Fig. 2 Plots comparing the magnetization transfer functions for Valine methyl (^1H , ^{13}C)- ^{13}C O correlations (green curve; $[\sin(4\pi\tau_d^1 J_{C\alpha-CO})\exp(-4R_{C\alpha}\tau_d)\exp(-4R_{C\beta}\tau_d)]^2$, where $^1 J_{C\alpha-CO}$ is the one bond $^{13}\text{C}_\alpha$ - ^{13}C O J coupling (55 Hz), and $R_{C\alpha}$ and $R_{C\beta}$ are the relaxation rates of $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ nuclei, respectively, both assumed to be equal to 30 s^{-1} for deuterated EIC), and Ile/Leu methyl (^1H , ^{13}C)- ^{13}C O correlations (red and blue curves corresponding to $[\sin(4\pi\tau_d^1 J_{C\alpha-CO})\sin(4\pi\tau_d^1 J_{C\alpha-C\beta})\exp(-4R_{C\alpha}\tau_d)]^2$ and $[\sin(4\pi\tau_d^1 J_{C\alpha-CO})\cos(4\pi\tau_d^1 J_{C\alpha-C\beta})\exp(-4R_{C\alpha}\tau_d)]^2$, respectively, where $^1 J_{C\alpha-C\beta}$ is the one bond $^{13}\text{C}_\alpha$ - $^{13}\text{C}_\beta$ J coupling (35 Hz) and the rest of the notation defined above) as a function of the delay $4\tau_d$ (ms) in the SIM-HMCM(CGCBBCA)CO experiment of Fig. 1. Methyl (^1H , ^{13}C)- ^{13}C O correlations arise from the magnetization terms indicated on the plot created at time point d and evolving during the t_1 evolution period of the scheme in Fig. 1

(an unlikely event due to significantly different $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ chemical shifts for Val and Leu), it may prove useful to be able to differentiate between the two types of residues from the SIM-HMCM(CGCBBCA)CO data.

The SIM-HMCM(CGCBBCA)CO experiment was tested on the C-terminal domain of Enzyme I (EIC) from *E. coli*, a symmetric 70-kDa homodimer. Enzyme I [PDB access code 2hwg (Teplyakov et al. 2006)] is the first enzyme in the phosphotransferase system (PTS), a signal transduction pathway that couples phosphoryl transfer to active sugar transport across the cell membrane (Meadow et al. 1990). Auto-phosphorylation of Enzyme I by phosphoenolpyruvate (PEP) in a Mg^{2+} -dependent manner is followed by the transfer of the phosphoryl group to the histidine phosphocarrier protein HPr (Weigel et al. 1982a, b). Recently, robust protocols for expression and purification of EIC, as well as the backbone assignments and conformational transitions occurring in EIC upon PEP binding have been reported by our laboratory (Venditti and Clore 2012; Venditti et al. 2012). NMR analysis and earlier crystallographic data show that the isolated C-terminal domain of enzyme I adopts a stable tertiary fold and exists in a monomer/dimer equilibrium that is modulated by ligand binding (Chauvin et al. 1994, 1996; Patel et al. 2006; Seok et al. 1996, 1998; Venditti and Clore 2012). The symmetric homo-dimer of EIC is shown in Fig. 3a. Under the sample conditions used in this study

(0.5 mM protein and 4 mM Mg^{2+} concentration), $\sim 95\%$ of the protein is present in solution in the dimeric state (Venditti and Clore 2012). The Ile and Leu/Val regions of the high-resolution ^1H - ^{13}C methyl constant time (CT)-HMQC spectrum (Bax et al. 1983; Mueller 1979; Tugarinov et al. 2003) of $[\text{U-}^2\text{H}, ^{15}\text{N}, ^{13}\text{C}; \text{Ile}^{\delta 1}\text{-}\{^{13}\text{CH}_3\}; \text{Leu, Val-}\{^{13}\text{CH}_3/^{12}\text{CD}_3\}]$ -labeled EIC are shown in Fig. 3b, c. Among 27 Ile $\delta 1$, 56 Leu δ and 36 Val γ methyl groups present in the EIC monomer, all Ile, 52 Leu and 26 Val methyls have been unambiguously assigned via combined analysis of 3D HMCM[CG]CBCA and 3D SIM-HMCM(CGCBBCA)CO data sets. The backbone chemical shifts of the remaining 5 Val methyls (Val²⁷⁸, Val²⁸¹, Val⁴²⁸, Val⁴³⁰ and Val⁴⁴⁶) were not assigned in the previous NMR study of EIC (Venditti and Clore 2012). A pair of resonances with outstanding ^{13}C and/or ^1H chemical shifts from this subset of methyls (Val⁴³⁰ and Val⁴⁴⁶) could be assigned tentatively based on the magnitude of the ring-current effects (Haigh and Mallion 1979; Koradi et al. 1996; Kuszewski et al. 1995) calculated from the x-ray coordinates (Teplyakov et al. 2006).

Figure 3d–f show the superposition of a selected region of the 2D ^1H - ^{13}C O planes from the 3D VAL-HMCM(CBCA)CO (Fig. 3d), ILE/LEU-HMCM(CGCBBCA)CO (Fig. 3e) and SIM-HMCM(CGCBBCA)CO (Fig. 3f) spectra corresponding to the area of ^1H - ^{13}C chemical shifts highlighted in the Leu-Val region of the methyl ^1H - ^{13}C correlation map in Fig. 3c. Clearly, the cross-peaks in the Val-selective VAL-HMCM(CBCA)CO (Fig. 3d) and Ile/Leu-selective ILE/LEU-HMCM(CGCBBCA)CO (Fig. 3e) experiments represent only subsets of the correlations obtained in the SIM-HMCM(CGCBBCA)CO data set. The convenience of recording a single data set comes at a relatively low price: we estimate that the sensitivity of the SIM-HMCM(CGCBBCA)CO data set is 25 % lower on average for Ile/Leu correlations and 50 % lower for Val correlations compared to the separate (Ile/Leu- and Val-selective) experiments recorded for the same measurement time. These losses in signal-to-noise per unit time in the simultaneous data set are associated with (1) faster relaxation of the multiple-quantum magnetization terms $8H_z^m C_z^\beta C_{\alpha'}^\gamma C_{\alpha'}^\gamma$ present during the t_1 evolution period for Val residues, (2) the non-idealities as well as relaxation and phase evolution during the long Q5 Gaussian Cascade pulses (Emsley and Bodenhausen 1987) for all residues, and (3) the need to adjust the length of the delay $4\tau_d$ to achieve the desired compromise in sensitivity of Val versus Ile/Leu correlations. We note that even the loss of about half of the signal for Val residues in the SIM-HMCM(CGCBBCA)CO data set is tolerable as the Val-selective VAL-HMCM(CBCA)CO spectra are usually much more sensitive than their Ile, Leu-selective counterparts (4.6-fold more sensitive in the 82-kDa Malate Synthase G at the same

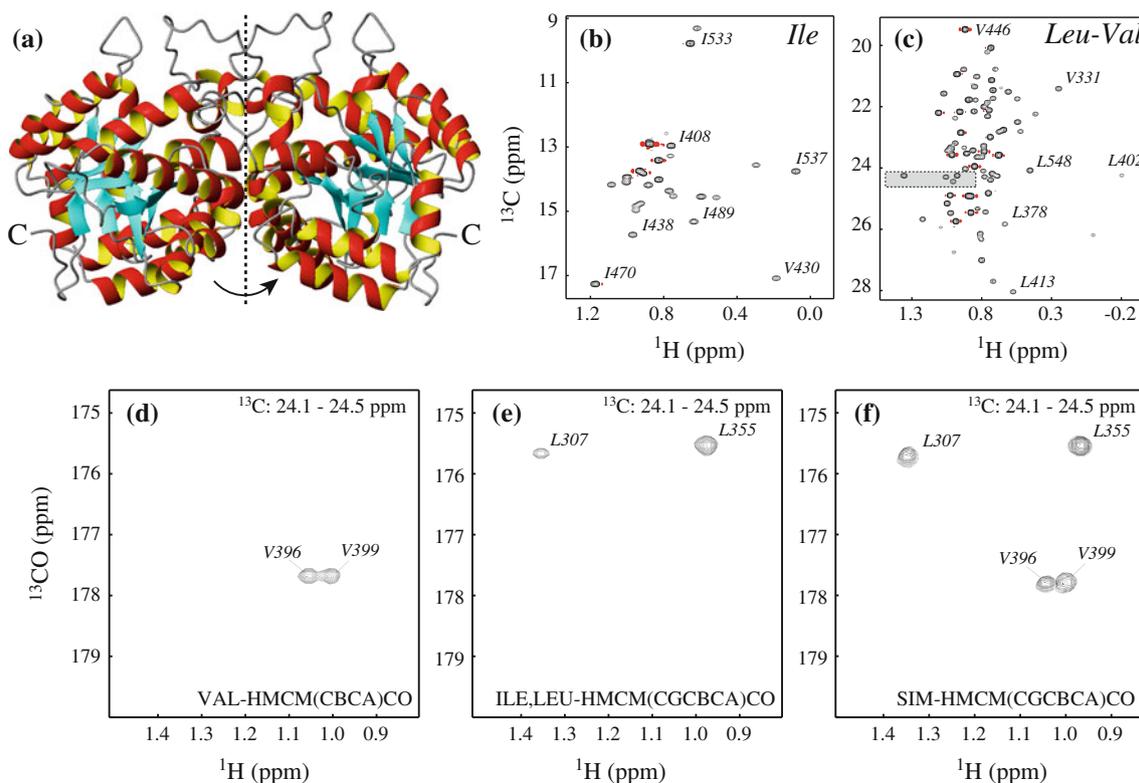


Fig. 3 **a** *Ribbon diagram* representation of the structure of the biological assembly of the C-terminal domain of Enzyme I (EIC dimer; pdb code 2hwg (Teplyakov et al. 2006)). The molecular dyad axis of the homodimer is shown with a *dashed vertical line*. Regions of the ^1H - ^{13}C methyl CT-HMQC correlation map of $[\text{U-}^2\text{H}, ^{15}\text{N}, ^{13}\text{C}; \text{Ile}^{\delta 1}\text{-}\{^{13}\text{CH}_3\}; \text{Leu}, \text{Val}\text{-}\{^{13}\text{CH}_3/^{12}\text{CD}_3\}]$ -labeled EIC (90 % $\text{H}_2\text{O}/10$ % D_2O ; 600 MHz; 37 °C) featuring the correlations of **b** Ile $^{\delta 1}$ and **c** Val $^{\gamma}$ /Leu $^{\delta}$ methyl groups. The high-resolution CT-HMQC correlation spectrum was acquired with the constant time period adjusted to $2/{}^1J_{\text{CC}} = 56$ ms. Selected residues are labeled with methyl peak residue types and numbers. The methyl-containing residues are numbered in the context of full-length EI. The assignments of Val 430 and Val 446 are tentative. Superposition of a selected region of the 2D ^1H - ^{13}CO planes plotted from the ^1H - ^{13}C chemical shift area highlighted in the Leu-Val region of the methyl ^1H - ^{13}C correlation

temperature (Tugarinov and Kay 2003) and can therefore be recorded with shorter total acquisition times.

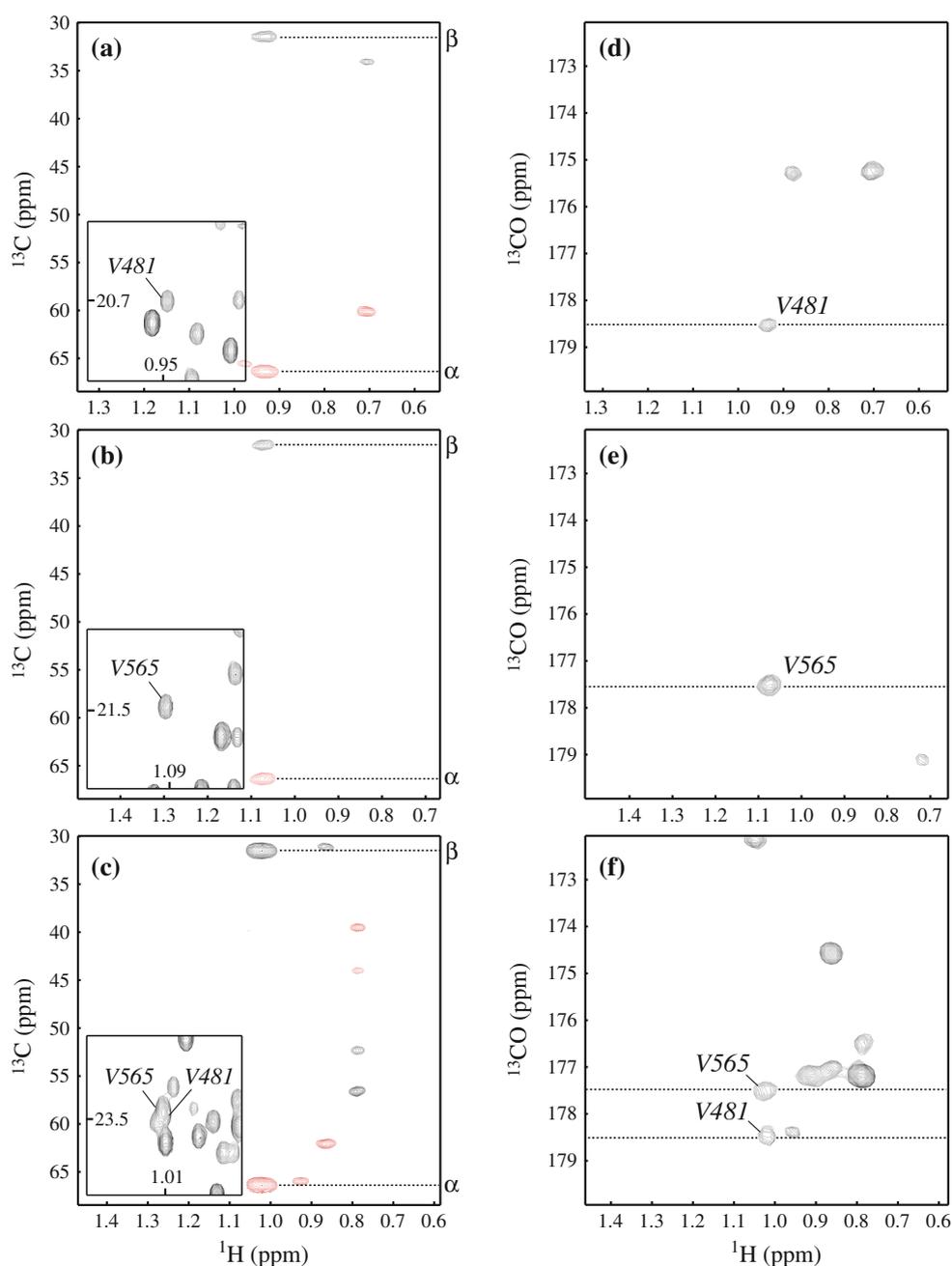
Although a single HMCM[CG]CBCA dataset is usually sufficient for assignments of the vast majority of methyl sites, and correlations of methyl resonances with backbone carbonyls are obtained primarily for verification purposes, even in such medium-sized systems as the EIC dimer, some cases can be identified when the correlations with carbonyl chemical shifts are absolutely indispensable for unambiguous assignments. For example, Val 481 and Val 565 of the EIC dimer have practically identical $^{13}\text{C}^{\alpha}$ and $^{13}\text{C}^{\beta}$ chemical shifts (31.2 and 66.4 ppm, respectively). Figure 4a–c show selected regions of the 2D ^1H - ^{13}C planes plotted from the 3D HMCM[CG]CBCA spectrum of the $[\text{U-}^2\text{H}, ^{15}\text{N}, ^{13}\text{C}; \text{Ile}^{\delta 1}\text{-}\{^{13}\text{CH}_3\}; \text{Leu}, \text{Val}\text{-}\{^{13}\text{CH}_3/^{12}\text{CD}_3\}]$ -EIC at the methyl ^{13}C chemical shifts of these two residues. The

map in panel **c** from **d** 3D VAL-HMCM(CBCA)CO, **e** 3D ILE/LEU-HMCM(CGBCA)CO and **f** 3D SIM-HMCM(CGBCA)CO data sets. The sample comprised 0.5 mM EIC dissolved in 20 mM Tris buffer, pH = 7.4, 100 mM NaCl, 2 mM DTT, 4 mM MgCl_2 , and 1 mM EDTA (Venditti and Clore 2012). All experiments were performed on a Bruker Avance III 600 MHz spectrometer equipped with a cryogenic probe operating at 37 °C. Typically, the 3D data sets were recorded with 18*, 40* and 512* complex points in the ^{13}CO , ^{13}C and ^1H dimensions, respectively, corresponding to respective acquisition times of 15, 14, and 64 ms. The recovery delay of 1.4 s and 32 scans per fid led to total acquisition times of ~40 h per 3D experiment. NMR spectra were processed with NMRPipe/NMRDraw software (Delaglio et al. 1995) and analyzed with NMRView (Johnson and Blevins 1994) using a tcl/tk interface written in-house

peaks of one of the two methyl sites of these two valines overlap in the ^1H - ^{13}C CT-HMQC map (Fig. 4c). The correlations provided by the SIM-HMCM(CGBCA)CO data set is the sole means of distinguishing between the two methyl positions, as the chemical shifts of backbone ^{13}CO nuclei of this pair of valines happen to be different by 1 ppm, as can be seen in the selected region of the 2D ^1H - ^{13}CO planes from the SIM-HMCM(CGBCA)CO data set plotted at the methyl ^{13}C chemical shifts of Val 481 and Val 565 shown in Fig. 4d–f.

Methyl assignments based on HMCM[CG]CBCA data sets are complicated by the fact that $^{13}\text{C}^{\beta}$ (and to a smaller extent $^{13}\text{C}^{\alpha}$) chemical shifts of ILV residues in $^{13}\text{CH}_3$ -labeled protein samples are different from those of fully deuterated proteins by three times the value of the multiple-bond deuterium isotope shifts (Gardner and Kay 1998;

Fig. 4 Selected regions of the 2D ^1H - ^{13}C planes drawn from the 3D HMCM(CG)CBCA spectrum of $[\text{U-}^2\text{H}, ^{15}\text{N}, ^{13}\text{C}; \text{Ile}^{\delta 1}\text{-}\{^{13}\text{CH}_3\}; \text{Leu, Val}\{^{13}\text{CH}_3/^{12}\text{CD}_3\}]$ -EIC (90 % $\text{H}_2\text{O}/10$ % D_2O ; 600 MHz; 37 °C) at the methyl ^{13}C chemical shifts of **a** one of the two Val^{481} methyls of EIC, **b** one of the two Val^{565} methyls of EIC, and **c** the position where the methyl peaks of Val^{481} and Val^{565} overlap. The insets in **a**–**c** show the regions of the 2D ^1H - ^{13}C CT-HMQC correlation map of EIC centered at the methyl positions in question. Negative contours are shown in red. ^{13}C chemical shift positions of ^1H - $^{13}\text{C}_\alpha$ and ^1H - $^{13}\text{C}_\beta$ correlations are indicated on the right side of each plot. Selected regions of the 2D ^1H - ^{13}CO planes from the SIM-HMCM(CG)CBCA data set plotted at the methyl ^{13}C chemical shifts of **d** one of Val^{481} methyls, **e** one of Val^{565} methyls, and **f** the region where the methyl peak of Val^{481} overlaps with that of Val^{565}



Rosen et al. 1996; Venters et al. 1996). Usually, however, $^{13}\text{C}^\beta$ and $^{13}\text{C}^\alpha$ chemical shifts are available from the assignment experiments acquired on fully deuterated protein samples as was the case in the previous study of EIC (Venditti and Clore 2012). For example, deuterium isotope shifts of $^{13}\text{C}^\beta$ nuclei arising from the substitution of three methyl deuterons for protons, $^n\Delta\text{C}^\beta(D_m)$ where n is the number of bonds between the carbon β and the site of the substitution ($n = 2$ for Val and $n = 3$ for Ile/Leu), are measured to be 250 ± 50 ppb for Val and 80 ± 30 ppb for Ile/Leu residues in EIC from a comparison of shifts in the 3D TROSY-HNCACB spectra acquired on the $^{13}\text{CH}_3$ -

labeled and fully deuterated protein samples. The variability of these isotope effects may compromise the exact ‘matching’ of $^{13}\text{C}^\beta$ and $^{13}\text{C}^\alpha$ frequencies in the HMCM[CG]CBCA and HNCACB data sets if recorded on different samples. Moreover, these deuterium isotope effects on $^{13}\text{C}^\beta$ nuclei may be somewhat offset by three-bond isotope shifts resulting from backbone amide $^1\text{H}_\text{N}$ -to- D_N substitutions, $^3\Delta\text{C}^\beta(\text{ND})$ (~ 37 ppb on average ranging from 2 to 65 ppb in model proteins (Zhang and Tugarinov 2013), when the HMCM[CG]CBCA spectra are recorded in D_2O . The assignments based on correlations with carbonyl shifts, however, do not suffer from these isotope

effects as (1) carbonyl carbons are removed by 4 (in Val) and 5 (in Ile/Leu) bonds from the sites of the methyl ^1H -to-D replacement, and (2) generally, deuterium isotope effects of carbonyl nuclei tend to be much smaller (Garrett et al. 1997; Sun and Tugarinov 2012; Zhang and Tugarinov 2013). In this context, we note that ^{13}C O chemical shifts in the SIM-HMCM(CGCBBCA)CO experiments recorded in H_2O and D_2O solvents are different by the sum of two- and three-bond deuterium isotope shifts arising from the ^1H -to-D substitutions at backbone amide positions ($^2\Delta\text{CO}(\text{N}_{i+1}\text{D}) + ^3\Delta\text{CO}(\text{N}_i\text{D})$); equal on average to 102 ± 25 ppb in EIC in good agreement with the measurements reported earlier in model protein systems (Feeney et al. 1974; Kainosho et al. 1987; LiWang and Bax 1996; Tüchsen and Hansen 1991; Zhang and Tugarinov 2013).

In summary, we describe a methyl-detected ‘out-and-back’ NMR experiment for obtaining simultaneous correlations of methyl resonances of Val and Ile/Leu residues with backbone carbonyl chemical shifts. The developed pulse-scheme serves the purpose of convenience in recording a single data set for all ILV methyl positions instead of acquiring two separate spectra for Val and Ile/Leu correlations. The SIM-HMCM(CGCBBCA)CO experiment can be used for ILV methyl assignments in moderately sized protein systems (up to ~ 100 kDa) where the backbone chemical shifts of $^{13}\text{C}^\alpha$, $^{13}\text{C}_\beta$ and ^{13}C O are known from prior NMR studies and where some losses in sensitivity can be tolerated for the sake of reduction in net NMR acquisition time. We note that though Val and Leu side-chains have been non-stereospecifically labeled in EIC via the use of [$^{13}\text{CH}_3/^{12}\text{CD}_3$]-labeled α -keto-isovalerate (Tugarinov et al. 2006; Tugarinov and Kay 2004), the developed methodology is equally applicable to stereospecifically [$^{13}\text{CH}_3/^{12}\text{CD}_3$]-labeled Val and Leu isopropyl moieties, i.e. with $^{13}\text{CH}_3$ labels restricted to either pro-*R* or pro-*S* methyl positions as obtained via the use of appropriately labeled aceto-lactate precursors (Gans et al. 2010).

Acknowledgments This work was supported by funds from the Intramural Program of the NIH, NIDDK, and the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the NIH (to G.M.C.).

References

Amero C, Schanda P, Durá MA, Ayala I, Marion D, Franzetti B, Brutscher B, Boisbouvier J (2009) Fast two-dimensional NMR spectroscopy of high molecular weight protein assemblies. *J Am Chem Soc* 131:3448–3449

Bax A, Griffey RH, Hawkins BL (1983) Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR. *J Magn Reson* 55:301–315

Boyd J, Soffe N (1989) Selective excitation by pulse shaping combined with phase modulation. *J Magn Reson* 85:406–413

Chauvin F, Brand L, Roseman S (1994) Sugar transport by the bacterial phosphotransferase system. Characterization of the *Escherichia coli* enzyme I monomer/dimer transition kinetics by fluorescence anisotropy. *J Biol Chem* 269:20270–20274

Chauvin F, Brand L, Roseman S (1996) Enzyme I: the first protein and potential regulator of the bacterial phosphoenolpyruvate: glyucose phosphotransferase system. *Res Microbiol* 147:471–479

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

Emsley L, Bodenhausen G (1987) Gaussian pulse cascades: new analytical functions for rectangular selective inversion and in-phase excitation in NMR. *Chem Phys Lett* 165:469–476

Feeney J, Partington P, Roberts GCK (1974) The assignment of carbon-13 resonances from carbonyl groups in peptides. *J Magn Reson* 13:268–274

Gans P, Hamelin O, Sounier R, Ayala I, Durá MA, 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 Chem Int Ed Engl* 49:1958–1962

Gardner KH, Kay LE (1998) The use of ^2H , ^{13}C , ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct* 27:357–406

Garrett DS, Seok YJ, Liao DI, Peterkofsky A, Gronenborn AM, Clore GM (1997) Solution structure of the 30 kDa N-terminal domain of enzyme I of the *Escherichia coli* phosphoenolpyruvate: sugar phosphotransferase system by multidimensional NMR. *Biochemistry* 36:2517–2530

Geen H, Freeman R (1991) Band-selective radiofrequency pulses. *J Magn Reson* 93:93–141

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 (δ 1) methyl-protonated ^{15}N -, ^{13}C -, ^2H -labeled proteins. *J Biomol NMR* 13:369–374

Haigh CW, Mallion RB (1979) Ring current theories in nuclear magnetic resonance. *Prog Nucl Magn Reson Spectrosc* 13:303–344

Hamel DJ, Dahlquist FW (2005) The contact interface of a 120 kD CheA–CheW complex by methyl TROSY interaction spectroscopy. *J Am Chem Soc* 127:9676–9677

John M, Schmitz C, Park AY, Dixon NE, Huber T, Otting G (2007) Sequence-specific and stereospecific assignment of methyl groups using paramagnetic lanthanides. *J Am Chem Soc* 129:13749–13757

Johnson BA, Blevins RA (1994) NMRView: a computer program for the visualization and analysis of NMR data. *J Biomol NMR* 4:603–614

Kainosho M, Nagao H, Tsuji T (1987) Local structural features around the C-terminal segment of Streptomyces subtilisin inhibitor studied by carbonyl carbon nuclear magnetic resonances of three phenylalanyl residues. *Biochemistry* 26:1068–1075

Kay LE, Ikura M, Tschudin R, Bax A (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J Magn Reson* 89:496–514

Koradi R, Billeter M, Wüthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 14:51–55

Kuszewski J, Gronenborn AM, Clore GM (1995) The impact of direct refinement against proton chemical shifts in protein structure determination by NMR. *J Magn Reson Ser B* 107:293–297

- LiWang AC, Bax A (1996) Equilibrium protium/deuterium fractionation of backbone amides in U- $^{13}\text{C}/^{15}\text{N}$ labeled human ubiquitin by triple resonance NMR. *J Am Chem Soc* 118:12864–12865
- Marion D, Ikura M, Tschudin R, Bax A (1989) Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange. *J Magn Reson* 85:393–399
- Meadow ND, Fox DK, Roseman S (1990) The bacterial phosphoenolpyruvate: glyucose phosphotransferase system. *Annu Rev Biochem* 59:497–542
- Mueller L (1979) Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence. *J Am Chem Soc* 101:4481–4484
- Ollershaw JE, Tugarinov V, Kay LE (2003) Methyl TROSY: explanation and experimental verification. *Magn Reson Chem* 41:843–852
- Patel HV, Vyas KA, Savtchenko R, Roseman S (2006) The monomer/dimer transition of enzyme I of the *Escherichia coli* phosphotransferase system. *J Biol Chem* 281:17570–17578
- Patt SL (1992) Single- and multiple-frequency-shifted laminar pulses. *J Magn Reson* 96:94–102
- Religa TL, Sprangers R, Kay LE (2010) Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. *Science* 328:98–102
- Rosen MK, Gardner KH, Willis RC, Parris WE, Pawson T, Kay LE (1996) Selective methyl group protonation of perdeuterated proteins. *J Mol Biol* 263:627–636
- Rosenzweig R, Moradi S, Zarrine-Afsar A, Glover JR, Kay LE (2013) Unraveling the mechanism of protein disaggregation through a ClpB–DnaK interaction. *Science* 339:1080–1083
- Ruschak AM, Kay LE (2010) Methyl groups as probes of supramolecular structure, dynamics and function. *J Biomol NMR* 46:75–87
- Ruschak AM, Religa TL, Breuer S, Witt S, Kay LE (2010) The proteasome antechamber maintains substrates in an unfolded state. *Nature* 467:868–871
- Seok YJ, Lee BR, Zhu PP, Peterkofsky A (1996) Importance of the carboxyl-terminal domain of enzyme I of the *Escherichia coli* phosphoenolpyruvate: sugar phosphotransferase system for phosphoryl donor specificity. *Proc Natl Acad Sci USA* 93:347–351
- Seok YJ, Zhu PP, Koo BM, Peterkofsky A (1998) Autophosphorylation of enzyme I of the *Escherichia coli* phosphoenolpyruvate: sugar phosphotransferase system requires dimerization. *Biochem Biophys Res Commun* 250:381–384
- Shaka AJ, Keeler J, Frenkiel T, Freeman R (1983) An improved sequence for broadband decoupling: WALTZ-16. *J Magn Reson* 52:335–338
- Sheppard D, Guo C, Tugarinov V (2009a) 4D ^1H - ^{13}C NMR spectroscopy for assignments of alanine methyls in large and complex protein structures. *J Am Chem Soc* 131:1364–1365
- Sheppard D, Guo C, Tugarinov V (2009b) Methyl-detected ‘out-and-back’ NMR experiments for simultaneous assignments of Ala β and Ile γ 2 methyl groups in large proteins. *J Biomol NMR* 43:229–238
- Sheppard D, Sprangers R, Tugarinov V (2010) Experimental approaches for NMR studies of sidechain dynamics in high-molecular-weight proteins. *Prog Nucl Magn Reson Spectrosc* 56:1–45
- 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
- Sun H, Tugarinov V (2012) Precision measurements of deuterium isotope effects on the chemical shifts of backbone nuclei in proteins: correlations with secondary structure. *J Phys Chem B* 116:7436–7448
- Teplyakov A, Lim K, Zhu PP, Kapadia G, Chen CC, Schwartz J, Howard A, Reddy PT, Peterkofsky A, Herzberg O (2006) Structure of phosphorylated enzyme I, the phosphoenolpyruvate: sugar phosphotransferase system sugar translocation signal protein. *Proc Natl Acad Sci USA* 103:16218–16223
- Tüchsen E, Hansen PE (1991) Hydrogen bonding monitored by deuterium isotope effects on carbonyl ^{13}C chemical shift in BPTI: intra-residue hydrogen bonds in antiparallel β -sheet. *Int J Biol Macromol* 13:2–8
- Tugarinov V, Kay LE (2003) Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J Am Chem Soc* 125:13868–13878
- Tugarinov V, Kay LE (2004) An isotope labeling strategy for methyl TROSY spectroscopy. *J Biomol NMR* 28:165–172
- Tugarinov V, Hwang PM, Ollershaw JE, Kay LE (2003) Cross-correlated relaxation enhanced ^1H - ^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc* 125:10420–10428
- Tugarinov V, Hwang PM, Kay LE (2004) Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu Rev Biochem* 73:107–146
- 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, Schachman 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
- Velyvis A, Schachman HK, Kay LE (2009) Assignment of Ile, Leu, and Val methyl correlations in supra-molecular systems: an application to aspartate transcarbamoylase. *J Am Chem Soc* 131:16534–16543
- Venditti V, Clore GM (2012) Conformational selection and substrate binding regulate the monomer/dimer equilibrium of the C-terminal domain of *Escherichia coli* enzyme I. *J Biol Chem* 287:26989–26998
- Venditti V, Fawzi NL, Clore GM (2011) Automated sequence- and stereo-specific assignment of methyl-labeled proteins by paramagnetic relaxation and methyl–methyl nuclear Overhauser enhancement spectroscopy. *J Biomol NMR* 51:319–328
- Venditti V, Fawzi NL, Clore GM (2012) An efficient protocol for incorporation of an unnatural amino acid in perdeuterated recombinant proteins using glucose-based media. *J Biomol NMR* 52:191–195
- Venters RA, Farmer BT, Fierke CA, Spicer LD (1996) Characterizing the use of perdeuteration in NMR studies of large proteins: ^{13}C , ^{15}N and ^1H assignments of human carbonic anhydrase II. *J Mol Biol* 264:1101–1116
- Weigel N, Kukuruzinska MA, Nakazawa A, Waygood EB, Roseman S (1982a) Sugar transport by the bacterial phosphotransferase system. Phosphoryl transfer reactions catalyzed by enzyme I of *Salmonella typhimurium*. *J Biol Chem* 257:14477–14491
- Weigel N, Waygood EB, Kukuruzinska MA, Nakazawa A, Roseman S (1982b) Sugar transport by the bacterial phosphotransferase system. Isolation and characterization of enzyme I from *Salmonella typhimurium*. *J Biol Chem* 257:14461–14469
- Xu Y, Liu M, Simpson PJ, Isaacson R, Cota E, Marchant J, Yang D, Zhang X, Freemont P, Matthews S (2009) Automated assignment in selectively methyl-labeled proteins. *J Am Chem Soc* 131:9480–9481
- Zhang D, Tugarinov V (2013) Accurate measurements of the effects of deuteration at backbone amide positions on the chemical shifts of ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$, $^{13}\text{C}\text{O}$ and $^1\text{H}\alpha$ nuclei in proteins. *J Biomol NMR* 56:169–182