A straightforward method for stereospecific assignment of val and leu prochiral methyl groups by solid-state NMR: Scrambling in the [2-13C]Glucose labeling scheme

Guohua Lv, Hannes Klaus Faßhuber, Antoine Loquet, Jean-Philippe Demers, Vinesh Vijayan, Karin Giller, Stefan Becker, Adam Lange

Max Planck Institute for Biophysical Chemistry, Department of NMR-based Structural Biology, Am Fassberg 11, 37077 Göttingen, Germany

Abstract

The unambiguous stereospecific assignment of the prochiral methyl groups in Val and Leu plays an important role in the structural investigation of proteins by NMR. Here, we present a straightforward method for their stereospecific solid-state NMR assignment based on [2-13C]Glucose ([2-13C]Glc) as the sole carbon source during protein expression. The approach is fundamentally based on the stereoselective biosynthetic pathway of Val and Leu, and the co-presence of [2-13C]Pyruvate produced mainly by glycolysis and [3-13C]/[1,3-13C]Pyruvate most probably formed through scrambling in the pentose phosphate pathway. As a consequence, the isotope spin pairs $^{13}C_b-^{13}C_c$ and $^{13}C_d-^{13}C_e$ in Val, and $^{13}C_c-^{13}C_d$ in Leu are obtained. The approach is successfully demonstrated with the stereospecific assignment of the methyl groups of Val and Leu of type 3 secretion system (T3SS) PrgI needles and microcrystalline ubiquitin.

1. Introduction

The stereospecific assignment of prochiral methyl groups of Val and Leu can greatly improve the precision and accuracy in protein structure determination by liquid-state NMR [1–4]. The methyl groups of Leu and Val are useful sources of structural information, as they are often found abundantly in the structurally important hydrophobic protein core and involved in numerous inter-residue and long-range contacts [5]. Additionally, it was recently shown that the methyl carbon chemical shifts can be used for obtaining $\chi_1$ rotamer distributions in Val and Leu residues, and that the methyl groups are a rich source of information about side-chain conformation and dynamics [6,7]. ssNMR spectroscopy is an emerging powerful tool for the structure determination of non-crystalline and insoluble proteins [8–12], which are not amenable to X-ray crystallography or solution NMR spectroscopy. The stereospecific assignment of the prochiral methyl groups of Val and Leu by ssNMR, should play an equally important role for structural and dynamical studies as in liquid-state NMR.

In order to address the stereospecific assignment of the prochiral methyl groups of Val and Leu, several isotope-labeling approaches have been developed [2,3,13–17]. The most widely used method to date relies on the use of a minimal culture medium containing 10% [1-13C]Glucose and 90% [12C]Glucose as the carbon source during biosynthetic protein production [2,18,19]. This elegant method has been successfully applied for obtaining the stereospecific assignment of the isopropyl groups of Val and Leu by liquid-state NMR and has also been extended to ssNMR [20]. However, a major practical drawback of this method is that the fractional labeling requires an additional sample to be prepared only for the purpose of stereospecific assignments.

Here, we propose an alternative approach for the stereospecific assignment of the methyl groups of Val and Leu based on the [2-13C]Glucose ([2-13C]Glc) labeling scheme. Bacterial growth in media containing [2-13C]Glc results in high 13C spin dilution in the protein produced with only one out of six carbons labeled [21,22]. We have reported earlier that the [2-13C]Glc labeling scheme leads to a significant resolution enhancement in ssNMR spectra and also to improvements in polarization transfer efficiencies due to the reduction of dipolar truncation effects [21,23]. In combination with uniformly [13C]Glc labeled ([U-13C]Glc-labeled) and [1-13C]Glc-labeled samples, sequential resonance assignments can be readily obtained [23]. Herein, we present that the [2-13C]Glc labeling scheme additionally allows for obtaining stereospecific assignments of the isopropyl groups of Val and Leu in a straightforward manner using ssNMR spectroscopy. As an illustration, we present stereospecific assignments of Val and Leu in type 3 secretion system (T3SS) PrgI needles and microcrystalline ubiquitin.
2. Results and discussion

The $^{13}$C enrichment pattern of [2-$^{13}$C]Glc has been well established by Lundström et al. [22]. To further investigate the labeling pattern and possible scrambling, we recorded a $^1$H-$^1$H HSQC spectrum of T3SS PrgI protomers that were produced with [2-$^{13}$C]Glc as the sole carbon source. As shown in Fig. S1, the observation of $^{13}$C$_{1}$/2 of Val, $^{13}$C$_{6}$/2 of Leu, and $^{13}$C$_{8}$/ of Ala with strong intensity indicates that metabolic scrambling indeed occurred. Starting from glucose, glycolysis leads to pyruvate as its final product (Fig. S2a) [24]. As a consequence, about 50% [2-$^{13}$C]pyruvate is expected when starting from [2-$^{13}$C]Glc. As an alternative pathway to glycolysis, the pentose phosphate pathway (PPP) [24] yields [1-$^{13}$C]Fructose-6-phosphate (F6P) and Glyceraldehyde-3-phosphate (GAP) starting from [2-$^{13}$C]Glucose-6-phosphate (G6P; Fig. 1). The stereoselective biosynthesis of Val and Leu. As an example, the formation of the isotope pairs (a) $^{13}$C$_{a}$-$^{13}$C$_{b}$ of Val and (b) $^{13}$C$_{a}$-$^{13}$C$_{b}$ of Leu originating from [2-$^{13}$C]pyruvate and [3-$^{13}$C]pyruvate is shown. The carbons that are $^{13}$C labeled are colored in red. Abbreviations: TPP, thiamine pyrophosphate; CoA: Coenzyme A.

As shown in Fig. 1, the labeled spin pairs $^{13}$C$_{a}$-$^{13}$C$_{b}$ of Val and $^{13}$C$_{a}$-$^{13}$C$_{b}$ in Val (Fig. 1a), and $^{13}$C$_{a}$-$^{13}$C$_{b}$2 and $^{13}$C$_{a}$-$^{13}$C$_{b}$1 in Leu (Fig. 1b) are obtained. Thus, by means of $^{13}$C-$^{13}$C correlation ssNMR spectroscopy, the two prochiral methyl groups of Val and Leu can be readily distinguished and the stereospecific assignment obtained. The resulting $^{13}$C-$^{13}$C correlation spectrum will exhibit strong intensity correlations for $^{13}$C$_{b}$-$^{13}$C$_{a}$ of Val and $^{13}$C$_{b}$-$^{13}$C$_{a}$1 of Leu, respectively. In contrast, the correlations $^{13}$C$_{a}$-$^{13}$C$_{b}$2 and $^{13}$C$_{a}$-$^{13}$C$_{b}$1 of Val and $^{13}$C$_{a}$-$^{13}$C$_{b}$2 and $^{13}$C$_{a}$-$^{13}$C$_{b}$1 of Leu are not expected. In case of a crowded methyl region in $^{13}$C-$^{13}$C correlation ssNMR spectra, this method thus bears another advantage, i.e., the simplification by the decreased number of detected correlations, in addition to the spectroscopic differentiation of the prochiral methyl groups.

The stereospecific $^{13}$C labeling pattern of Val and Leu resulting from [2-$^{13}$C]Glc is successfully demonstrated with applications to T3SS PrgI needles (Fig. 2a) and ubiquitin (Fig. S4), where stereospecific assignments of Val and Leu are obtained unambiguously. As shown in Fig. 2a, generally only one methyl carbon ($^{13}$C$_{a}$1 of Val and $^{13}$C$_{a}$2) correlates to $^{13}$C$_{b}$ of Val in the 2D $^{13}$C–$^{13}$C PDSD spectrum of [2-$^{13}$C]Glc-labeled T3SS PrgI needles (in magenta), while four correlations ($^{13}$C$_{a}$1/2–$^{13}$C$_{b}$ and $^{13}$C$_{a}$1/2–$^{13}$C$_{b}$1) are present in the 2D PDSD spectrum of [U–$^{13}$C]Glc-labeled T3SS PrgI needles (in black). For Val12, 27, and 65, as expected, the correlations of $^{13}$C$_{a}$–$^{13}$C$_{b}$1 and $^{13}$C$_{a}$–$^{13}$C$_{b}$2 were observed for [2-$^{13}$C]Glc-labeled T3SS PrgI needles (magenta spectrum in Fig. 2a). For Val20, only the correlation of $^{13}$C$_{a}$–$^{13}$C$_{b}$1 was observed, but with weak intensity, while the correlation of $^{13}$C$_{a}$–$^{13}$C$_{b}$2 was absent. This observation is consistent with the observation of weak-intensity correlations in the 2D PDSD spectrum of [U–$^{13}$C]Glc-labeled T3SS PrgI needles (in black) and can be attributed to structural plasticity [10]. For Val67, in addition to the $^{13}$C$_{a}$–$^{13}$C$_{b}$1 correlation, a $^{13}$C$_{a}$–$^{13}$C$_{b}$2 correlation was also observed, but with much weaker intensity compared to the expected correlations. Peaks that are observed in the [2-$^{13}$C]Glc-labeled spectrum in addition to the ones observed in the [U–$^{13}$C]Glc-labeled sample corre-
spond to sequential, medium- and long-range correlations, as the PDSD mixing time used was longer (400 ms).

Similarly for Leu, the stereoselective $^{13}$C enrichment pattern, i.e., $^{13}$C$_b$–$^{13}$C$_d$$_1$ and $^{13}$C$_c$–$^{13}$C$_d$$_2$, is illustrated using [2-13C]Glc-labeled ubiquitin (magenta spectrum in Fig. S4b). Taking Leu56 as an example, the correlations $^{13}$C$_b$–$^{13}$C$_d$$_1$ and $^{13}$C$_c$–$^{13}$C$_d$$_2$ were observed without the presence of the correlations $^{13}$C$_b$–$^{13}$C$_d$$_2$ and $^{13}$C$_c$–$^{13}$C$_d$$_1$. However, for Leu50, only the correlation of $^{13}$C$_c$–$^{13}$C$_d$$_2$ was clearly present, while the correlation of $^{13}$C$_b$–$^{13}$C$_d$$_1$ was very weak and only observable on one side of the diagonal. This might be due to the longer distance involved in this correlation (two bonds) compared to the $^{13}$C$_c$–$^{13}$C$_d$$_2$ (one bond) correlation. Additionally, strong $^{13}$C$_b$–$^{13}$C$_c$ cross-peaks are observed for Leu. These peaks are expected from the labeling pattern and are very intense compared to the $^{13}$C$_b$–$^{13}$C$_d$$_2$ and $^{13}$C$_c$–$^{13}$C$_d$$_1$ peaks that are only observed due to scrambling.

In principle, one can directly identify Val and Leu methyl groups by their expected correlations in PDSD spectra of [2-13C]Glc-labeled samples, i.e., $^{13}$C$_x$–$^{13}$C$_y$$_1$ and $^{13}$C$_p$–$^{13}$C$_q$$_2$ for Val, and $^{13}$C$_p$–$^{13}$C$_q$$_1$ and $^{13}$C$_y$–$^{13}$C$_z$$_2$ for Leu. However, in case of absence of expected peaks, e.g., as seen in this study for Prgl Val20 $^{13}$C$_p$–$^{13}$C$_q$$_2$, one also needs [U–13C]Glc spectra to assist the assignment of $^{13}$C$_y$1/2.

In addition to the scrambling due to the PPP detailed above, it is noteworthy that (a) amino acid degradation together with the C1 metabolism [24], (b) the breakdown of oxaloacetate via gluconeogenesis [24] as well as (c) the conversion of malate [24] might contribute to scrambling resulting in (a) [3-13C]pyruvate, [1-13C]pyruvate, [1,3-13C]pyruvate, and [2,3-13C]pyruvate, (b) [1,3-13C]pyruvate and [1-13C]pyruvate, and (c) [1,3-13C]pyruvate and [1-13C]pyruvate ([2-13C]pyruvate is not considered here, as it is always present in the [2-13C]Glc-labeling scheme). As mentioned above, the spin pairs $^{13}$C$_p$–$^{13}$C$_q$$_1$ and $^{13}$C$_x$–$^{13}$C$_y$$_2$ of Val originate from carbons 2 and 3 of a single pyruvate (Fig. 1). Therefore, the presence of [2,3-13C]pyruvate could explain the presence of the $^{13}$C$_x$–$^{13}$C$_y$$_2$ correlation of Val67 in T3SS Prgl needles (Fig. 2a).

In Fig. 2, (a) 2D PDSD spectra of [U–13C]Glc- (in black) and [2-13C]Glc-labeled (in magenta) T3SS Prgl needles. The spectrum of the [U–13C]Glc-labeled sample was recorded with a mixing time of 50 ms, using maximum acquisition times of 20 ms (direct dimension) and 15 ms (indirect dimension). The total experimental time was 3 days. The spectrum of the [2-13C]Glc-labeled sample was recorded with a mixing time of 400 ms, using maximum acquisition times of 15 ms (direct dimension) and 10 ms (indirect dimension). The total experimental time was 1.5 days. Spin systems for V12, V20, V27, V65, and V67 are highlighted in blue. The stereospecific assignment of $^{13}$C$_y$1/2 in Val (a) was obtained based on the correlations $^{13}$C$_x$–$^{13}$C$_y$$_1$ and $^{13}$C$_p$–$^{13}$C$_q$$_2$ as illustrated in (b). 1D slices for the residues Val65 and Val67 are shown in Fig. S3.
generally of minor importance when *Escherichia coli* cells have access to glucose [24,28]. This implies that the enrichment level of [2,3-13C]pyruvate is so low that it does not disturb the stereospecific assignment of Val and Leu. This is also corroborated by the fact that no Ala Cα–Cβ peaks are observed in ssNMR spectra of [2-13C]Glc-labeled proteins [21,23,29], which would be the case if [2,3-13C]pyruvate was present.

3. Conclusion

In conclusion, we have presented an alternative stereospecific labeling pattern of Val and Leu formed by the co-presence of [2-13C]pyruvate and [3-13C] ([1.3-13C]pyruvate starting from [2-13C]Glc, enabling us to obtain stereospecific assignments of the isopropyl groups of Val and Leu in a simple way. For this purpose, only a unique [2-13C]Glc-labeled sample is required, the isotope expense is thus reduced compared to fractional labeling approaches [2,18–20]. Given the resolution enhancement and spectral simplification that we have recently exploited for the *de novo* sequential assignment of the T3SS PrgI needles [23] and of mouse α-synuclein fibrils [29], the detection of long-range distance restraints [10], and the stereospecific assignment of the isopropyl groups of Val and Leu presented here, the [2-13C]Glc labeling scheme has proven to be a remarkably versatile scheme, and should therefore become an attractive tool for structural investigations of proteins by ssNMR.

4. Materials and methods

4.1. Sample preparation

For [U-13C]Glc- and [2-13C]Glc-labeled T3SS PrgI needles were prepared as described in Loquet et al. [10]. [U-13C]Glc- and [2-13C]Glc-labeled ubiquitin was prepared as described in Seidel et al. [30]. [2-13C]Glc-labeled PrgI protomers were prepared as in Loquet et al. [10] with the exception that the monomers were not polymerized. The buffer used for the solution NMR studies contained 20 mM MES, pH 5.5, 10% D2O. All samples were uniformly 15N labeled.

4.2. Solid-state NMR experiments and data processing

Samples were packed in 4 mm MAS rotors, using protein quantities of ~10 mg for T3SS PrgI needles and ~20 mg for ubiquitin. All spectra were recorded at a spinning frequency of 11 kHz and the 13C chemical shifts were calibrated with DSS as an internal Ref. [31]. The temperature-dependent position of the water proton resonance was used to measure the temperature inside the MAS rotor [32]. High-power 1H-13C decoupling (SPINAL-64 [33]) with a radio-frequency amplitude of 83 kHz was applied during evolution and detection periods.

For [U-13C]Glc- and [2-13C]Glc-labeled T3SS PrgI needles, two-dimensional 13C-13C ssNMR experiments were conducted on a 20.0 Tesla (1H resonance frequency: 850 MHz) wide-bore spectrometer (Bruker Biospin, Germany), equipped with a 4 mm triple-resonance (1H, 13C, 15N) MAS probe. The experiments were conducted on a sample temperature of 278 K (+5 °C). 13C-13C transfer was achieved via proton-driven spin-diffusion (PDSD) with mixing times of 50 ms and 400 ms for [U-13C]Glc- and [2-13C]Glc-labeled PrgI needles, respectively. These two spectra were processed with NMRpipe [34] and analyzed using CcpNmr [35].

For [U-13C]Glc- and [2-13C]Glc-labeled ubiquitin, two-dimensional 13C-13C ssNMR experiments were conducted on a 18.8 Tesla (1H resonance frequency: 800 MHz) standard-bore spectrometer (Bruker Biospin, Germany), equipped with a 4 mm triple-resonance (1H, 13C, 15N) MAS probe. 13C-13C transfer was achieved via proton-driven spin-diffusion (PDSD) with mixing times of 50 ms and 100 ms for [U-13C]Glc- and [2-13C]Glc-labeled ubiquitin, respectively. The spectra were processed with Topspin (Bruker Biospin, Germany) and analyzed in SPARKY version 3.1 (T.D. Goddard & D.G. Kneller, University of California).

4.3. Solution NMR experiment and data processing

A 1H-13C HSQC spectrum was recorded on a 21.1 Tesla (1H resonance frequency: 900 MHz) standard-bore spectrometer (Bruker Biospin, Germany), equipped with a cryogenically-cooled triple resonance probe head, with 1024 (τ2) × 1152 (τ1) complex points and maximum acquisition times of 33.6 ms (τ2) and 33.9 ms (τ1). 32 scans were collected per indirect increment. The 1H carrier frequency was set to the water resonance (4.67 ppm), and the 13C carrier frequency was set to 45 ppm. The spectrum was zero-filled automatically, apodized with a squared sine bell window function along τ2 and τ1 axes, and then Fourier transformed. The spectrum was processed with NMRpipe [34] and analyzed using SPARKY version 3.1 (T.D. Goddard & D.G. Kneller, University of California).

Acknowledgments

We are grateful to Christian Griesinger for continuous support of this project, Zrinka Gattin for discussions, and Gitta Angerstein for expert technical assistance. We thank the Max Planck Society, the DFG (Emmy Noether fellowship to A. Lange), the China Scholarship Council (Ph.D. scholarship to G. Lv), and EMBO (long-term fellowship to A. Loquet) for financial support.

Appendix A. Supplementary data

HSQC spectrum of T3SS PrgI protomers; glycolysis and PPP pathways involved in the formation of pyruvate; 1D traces for residues Val65 and Val67 in PrgI; the stereospecific assignment of Val and Leu in ubiquitin. Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmr.2012.12.017.

References
