Journal of Magnetic Resonance 248 (2014) 81–95

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

Time-shared experiments for efficient assignment of triple-selectively labeled proteins

Frank Löhr [a], Aisha Laguerre [a], Christoph Bock [b], Sina Reckel [a,1], Peter J. Connolly [c], Norzehan Abdul-Manan [c], Franz Tumulka [b], Rupert Abele [b], Jonathan M. Moore [c], Volker Dötsch [a,*]

[a] Institute of Biophysical Chemistry & Center for Biomolecular Magnetic Resonance, Goethe University, Max-von-Laue-Str. 9, D-60438 Frankfurt, Germany
[b] Institute of Biochemistry, Goethe University, Max-von-Laue-Str. 9, D-60438 Frankfurt, Germany
[c] Vertex Pharmaceuticals Inc., Cambridge, MA 02139, USA

ARTICLE INFO

Article history:
Received 23 July 2014
Revised 10 September 2014
Available online 30 September 2014

Keywords:
BEST-TROSY
Cell-free expression
Deuteration
Hadamard decoding
Isotope labeling
Membrane proteins
Triple-resonance NMR

ABSTRACT

Combinatorial triple-selective labeling facilitates the NMR assignment process for proteins that are subject to signal overlap and insufficient signal-to-noise in standard triple-resonance experiments. Aiming at maximum amino-acid type and sequence-specific information, the method represents a trade-off between the number of selectively labeled samples that have to be prepared and the number of spectra to be recorded per sample. In order to address the demand of long measurement times, we here propose pulse sequences in which individual phase-shifted transients are stored separately and recombined later to produce several 2D HN(CX) type spectra that are usually acquired sequentially. Sign encoding by the phases of ^{13}C 90° pulses allows to either select or discriminate against $^{13}C'$ or $^{13}C^{\alpha}$ spins coupled to ^{15}N. As a result, ^{1}H–^{15}N correlation maps of the various isotopomeric species present in triple-selectively labeled proteins are deconvoluted which in turn reduces problems due to spectral overlap. The new methods are demonstrated with four different membrane proteins with rotational correlation times ranging from 18 to 52 ns.

1. Introduction

Obtaining complete backbone resonance assignment is desirable for detailed NMR studies of the structure of proteins, their dynamics, and interactions with other molecules. However, in cases where severe degeneracy of peak positions meets large linewidths due to slow tumbling and/or conformational exchange, well-established assignment strategies in solution NMR exclusively based on matching intraresidual and sequential correlations of amide groups with other nuclei may fail. Typical examples include α-helical membrane proteins [1–5].

For uniformly labeled proteins considerable simplification of ^{1}H–^{15}N correlation spectra as well as valuable residue-type information can be gained from amino-acid type selective experiments [6–10]. Unfortunately, an increased number of coherence transfer steps or editing elements in the corresponding pulse sequences may render this approach inefficient for larger systems. Alternatively, amino-acid specific isotope labeling methods are available for which molecular size restrictions apply to a much lesser extent and which may take advantage of cell-free expression systems. The latter offer reduced isotope scrambling and requires relatively low amounts of labeled amino acids [11–17]. While amino-acid type information is obtained via selective ^{15}N labeling [12,14,18–20], dual selective labeling with ^{15}N and ^{13}C isotopes additionally permits selective detection of specific amino acid pairs based on interresidual $^{1}J(^{15}N,^{13}C')$ scalar interactions [11,13,21–26]. The number of samples required to obtain sufficient anchor points for sequence-specific assignments can be minimized by designing labeling patterns in a combinatorial fashion [27–31].

Recently, we have introduced combinatorial triple-selective labeling [32] to further increase residue-type and sequence-specific information that can be collected with a given number of samples. The disadvantage to this strategy is that a greater number of different 2D ^{1}H–^{15}N detected triple-resonance experiments must be acquired including less sensitive ones such as HN(COCA), (CO)HN(CA) and DQ-HN(CA). Additionally, and common to all combinatorial labeling protocols, is the problem that some amide signals can remain undiscovered in the case of chemical shift degeneracy in the two dimensions. Here we present modified pulse sequences that allow recording several experiments in a

* Corresponding author.
E-mail address: vdoetsch@em.uni-frankfurt.de (V. Dötsch).
[1] Current address: ISREC Chair Foundation in Translational Oncology UPHAN, EPFL, CH-1015 Lausanne, Switzerland.

http://dx.doi.org/10.1016/j.jmr.2014.09.014
1090-7807/

time-shared (ts) manner to address both prolonged measurement times and spectral overlap.

Time-shared schemes, originally proposed by Sørensen [33], have been extensively used to correlate nuclei involved in independent coherence transfer pathways in a single experiment, resulting in appreciable time savings [34–49]. Although different experiment types required to distinguish individual sites in triple-selectively labeled proteins are recorded simultaneously, identical nuclei (^{1}H and ^{15}N) are correlated in the implementations described here. Measurement times are dictated by the least sensitive of the combined experiments, while the remaining experiments are obtained "for free". This is achieved by separate storage of transients acquired with different pulse phases followed by appropriate data recombination [50–57] in order to select amide signals involved in different scalar coupling networks. As a desired side-product of the new methods isotope filtering against either $^{13}C'$ or $^{13}C^{\alpha}$ simplifies [^{15}N, ^{1}H]-HSQC type spectra [58–60], enabling the identification of otherwise overlapping signals.

2. Materials and methods

2.1. Sample preparations

Production of isotopically labeled proteins was carried out using a continuous-exchange cell-free expression system based on an *Escherichia coli* S30 extract [61]. Samples of TMD0 [62] as well as KvAP VSD and proteorhodopsin [32] were prepared as described previously. Lipoprotein signal peptidase II (LspA) was cloned into pET28b with an N-terminal streptactin, 6× histag and thrombin cleavage site. Continuous-exchange cell-free expression in the presence of preformed MSP1D1ΔH5-DMPC nanodiscs was performed using the standard protocol as described elsewhere [63,64]. Empty nanodiscs were prepared before use in cell-free expression system as previously published [65]. LspA – MSP1D1ΔH5-DMPC nanodisc NMR samples were purified by metal affinity chromatography. Empty nanodiscs were removed from the samples by streptactin purification. On overview of the isotopic labeling patterns employed for the four proteins is provided in Table S1 of the Supplementary Material.

Final sample conditions were as follows: 0.55 mM TMD0 in 1% dihexanoyl-phosphatidylcholine (diC6PC), 25 mM sodium acetate (pH 5), 100 mM NaCl; 0.3 mM KvAP VSD in a detergent mixture of 5% *n*-dodecylphophocholine (DPC) and 1.7% *n*-dodecyl-N,N-dimethylamine-N-oxide (LDAO), 25 mM sodium acetate (pH 5), 2 mM dithiothreitol (DTT); 0.5 mM proteorhodopsin in 4% diheptanoyl-phosphatidylcholine (diC7PC), 25 mM sodium acetate (pH 5), 2 mM DTT; 0.25 mM LspA in MSP1D1ΔH5/DMPC nanodiscs, 20 mM sodium acetate (pH 4). The samples of KvAP VSD, proteorhodopsin and LspA had a volume of 320 µl including 5% D_2O and 0.15 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) as internal standard and were placed in 5 mm susceptibility-matched Shigemi tubes. The sample volume of TMD0 was 190 µl, contained in a 4 mm Shigemi tube that was inserted into a standard 5 mm tube filled with a 0.15 mM DSS solution in D_2O.

2.2. NMR spectroscopy

All NMR experiments were carried out at a Bruker Avance 600 MHz spectrometer equipped with a cryogenic ^{1}H[^{13}C/^{15}N] triple-resonance probe with sample temperatures set to either 318 K (proteorhodopsin and LspA) or 323 K (TMD0 and KvAP VSD). Pulse sequences were of the [^{15}N, ^{1}H]-TROSY [66,67] type including a sensitivity-enhanced gradient coherence selection scheme [68–70]. Shaped proton pulses centered at 8.5 ppm were employed in all time-shared and separately recorded reference experiments to accelerate longitudinal relaxation between scans according to the Band selective-Excitation Short-Transient (BEST) [71,72] protocol. Proton magnetization longitudinally relaxed during the pulse sequences is transferred to ^{15}N by the final single-transition-to-single-transition (ST2-PT) element. In order to constructively add the resulting ^{15}N polarization in the subsequent scan a ^{15}N 180° pulse was applied immediately following ^{1}H acquisition [73]. Pulse schemes to obtain ^{13}C-edited 2D ^{1}H–^{15}N correlations were based on magnetization transfer pathways followed in HNCO [74], HNCA [74], HN(CO)CA [75], COHNCA [76] and double-quantum (DQ) HNCA [77] triple-resonance experiments. Carbon-13 evolution periods were replaced by fixed 3-µs delays and, where applicable, ^{1}H 180° decoupling pulses required in the corresponding higher dimensional sequences were omitted [32]. Reference 2D HN(CO), HN(CA), (CO)HN(CA), HN(COCA) and DQ-HN(CA) spectra were recorded using INEPT [78] type ^{15}N–^{13}C polarization transfer steps and standard phase cycling. Modified pulse sequences employed in the current study (Fig. 1) are based on HMQC-like [79,80] ^{15}N–^{13}C transfer rather and involve individual storage of FIDs recorded with different phase settings of ^{13}C 90° pulses. Details are given in the figure legend.

The time between the end of acquisition and the beginning of the next scan of triple-resonance pulse sequences was set to 0.3 s in experiments on TMD0, KvAP VSD, and proteorhodopsin and to 0.7 s in experiments on LspA. Recycle delays in HSQC experiments were 0.3 s for TMD0 and 0.7 s for the remaining proteins. All spectra were acquired with spectral widths of 13 ppm in the ^{1}H dimension and 50 ppm in the ^{15}N dimension, with carrier frequencies placed on the water resonance and in the center of the amide region, respectively. The ^{1}H acquisition time was uniformly set to 41 ms (320 complex data points) whereas acquisition times in the ^{15}N acquisition was varied between 42.1 and 73.6 ms for different experiments and samples according to the number of expected signals. The exact number of t_1 data points and corresponding acquisition times employed in the individual experiments are summarized in the Supplementary Material (Table S1) along with delay durations for ^{15}N–^{13}C and ^{13}C–^{13}C magnetization transfer and measurement times.

Spectra processing and analysis was performed with TopSpin 3.2 (Bruker). Cosine-squared window functions were applied for apodization in both dimensions. Contour levels were drawn on an exponential scale using a factor of $2^{1/2}$ in all plots shown in the following.

3. Results and discussion

As shown in a previous publication [32], recording two-dimensional ^{13}C-edited proton-nitrogen correlation spectra on a set of combinatorial ^{15}N, 1-^{13}C and fully ^{13}C/^{15}N-labeled protein samples can significantly aid the backbone assignment in situations where high-quality three-dimensional spectra on uniformly labeled samples are difficult to obtain. The method termed combinatorial triple-selective labeling involves acquisition of a total of six experiments, i.e. [^{15}N,^{1}H]-TROSY-HSQC and TROSY-based HN(CO), HN(CA), (CO)HN(CA), HN(COCA), and DQ-HN(CA) to unambiguously distinguish ^{15}N–^{1}H cross peaks from the six possible combinations of ^{15}N- or ^{13}C/^{15}N-labeled amino acids in position i and either 1-^{13}C-labeled, ^{13}C/^{15}N-labeled or non-^{13}C-labeled amino acids in position $i-1$. Compared to combinatorial dual-selective labeling which employs only ^{15}N- and ^{13}C/^{15}N- [13,28] or ^{15}N- and 1-^{13}C-labeled [29,31] amino acids, the percentage of amino acid types as well as sequential pairs that can be identified from a limited number of samples is increased. As a drawback, considerably longer measurement times per sample are required to acquire the supplementary 2D HN(CA), (CO)HN(CA), HN(COCA) and DQ-HN(CA) data sets. On the other hand, dispensing with the ^{13}C