J Biomol NMR (2015) 62:239–245
DOI 10.1007/s10858-015-9939-2

COMMUNICATION

Methyl labeling and TROSY NMR spectroscopy of proteins expressed in the eukaryote *Pichia pastoris*

Lindsay Clark[1] · Jacob A. Zahm[1,2] · Rustam Ali[1,2] · Maciej Kukula[4] · Liangqiao Bian[4] · Steven M. Patrie[3] · Kevin H. Gardner[5] · Michael K. Rosen[1,2] · Daniel M. Rosenbaum[1]

Received: 11 February 2015 / Accepted: 27 April 2015 / Published online: 30 May 2015

Abstract ^{13}C Methyl TROSY NMR spectroscopy has emerged as a powerful method for studying the dynamics of large systems such as macromolecular assemblies and membrane proteins. Specific ^{13}C labeling of aliphatic methyl groups and perdeuteration has been limited primarily to proteins expressed in *E. coli*, preventing studies of many eukaryotic proteins of physiological and biomedical significance. We demonstrate the feasibility of efficient ^{13}C isoleucine δ1-methyl labeling in a deuterated background in an established eukaryotic expression host, *Pichia pastoris*, and show that this method can be used to label the eukaryotic protein actin, which cannot be expressed in bacteria. This approach will enable NMR studies of previously intractable targets.

Keywords Methyl labeling · TROSY · Eukaryotic expression system · *Pichia pastoris* · Deuteration · Actin

Electronic supplementary material The online version of this article (doi:10.1007/s10858-015-9939-2) contains supplementary material, which is available to authorized users.

✉ Daniel M. Rosenbaum
dan.rosenbaum@utsouthwestern.edu

[1] Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

[2] Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

[3] Department of Pathology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

[4] Shimadzu Center for Advanced Analytical Chemistry, University of Texas at Arlington, Arlington, TX 76019, USA

[5] Structural Biology Initiative, CUNY Advanced Science Research Center, New York, NY 10031, USA

Well-resolved multi-dimensional NMR spectra are essential for obtaining structural and dynamic information on backbone and sidechain moieties within proteins. However, obtaining such spectra of large macromolecules is complicated by poor peak dispersion and line broadening due to rapid transverse relaxation of nuclear magnetization and spectral crowding. To overcome this problem for aliphatic sidechains, proteins can be specifically labeled with ^{13}C in the methyl groups of isoleucine, leucine, and valine residues using ^{13}C α-ketoacid precursors in *E. coli* (Gardner and Kay 1997; Goto et al. 1999). When paired with selective protonation in an otherwise deuterated background (Rosen et al. 1996), this approach takes advantage of the favorable relaxation properties of ^{13}C-methyl groups with the application of transverse relaxation optimized spectroscopy (TROSY) (Pervushin et al. 1997; Ollerenshaw et al. 2003). However these methods have remained unavailable for many eukaryotic proteins due to poor expression and folding in *E. coli* resulting from lack of required chaperones, lack of proper post-translational modifications, or improper membrane composition.

Several different eukaryotic hosts, including fungi (Miyazawa-Onami et al. 2013), insect cells (Nygaard et al. 2013; Kofuku et al. 2014), and mammalian cells (Werner et al. 2008), have been used to overexpress proteins for NMR. While these systems have succeeded in producing amino acid-specific and uniformly ^{15}N or ^{13}C labeled material (Chen et al. 2005; Fan et al. 2011; Gossert et al. 2011; Strauss et al. 2005; Hansen et al. 1992), the high expense and difficulties in perdeuteration have limited their widespread use for larger proteins. The methylotrophic yeast *Pichia pastoris* is a well established expression host (Cereghino and Cregg 2000) for proteins that cannot be made in *E. coli*—eukaryotic membrane

proteins such as ATP transporters (Lee et al. 2002), ion pumps (Strugatsky et al. 2003) and G-protein coupled receptors (Shimamura et al. 2011; Hino et al. 2012) have all been successfully overexpressed in and purified from this organism. Genetic manipulation, transformation, and growth of *P. pastoris* are more rapid than for higher eukaryotes such as insect cells and mammalian cells.

Overexpression using the tightly regulated AOX1 promoter often yields milligram quantities of recombinant protein per liter of *P. pastoris* suspension culture (Cereghino and Cregg 2000). *P. pastoris* is also favorable for NMR studies given its ability to grow on defined minimal media, uptake isotope-containing precursors, and efficiently incorporate deuterium at non-exchangeable sites (Morgan et al. 2000). Despite conservation of branched-chain amino acid biosynthesis pathways from *E. coli* (Fig. 1), site-specific methyl labeling using α-ketoacid precursors has not been reported in. *P. pastoris.*

We initially explored the use of ^{13}C-methyl α-ketobutyrate in *P. pastoris* cultures to label maltose binding protein (MBP) with ^{13}C at the δ1-methyl groups of isoleucine (Ile) residues. MBP has well-characterized ^{1}H-^{13}C 2D NMR spectra (Gardner et al. 1998) and is highly expressed in *P. pastoris* (Li et al. 2010). We collected ^{1}H-^{13}C heteronuclear single quantum coherence (HSQC) spectra on MBP that was labeled by addition of ^{13}C-methyl α-ketobutyrate to the culture media (Fig. 2a).

Resonances for all 22 Ile δ1-methyl groups of MBP (Gardner et al. 1998) are observed in our spectrum (Fig. 2a, Fig. S1), while little signal is present in other regions (indicating lack of "bleed-through" of the isotope into other amino acids—see Fig. 3). Based on tryptic peptide mass spectra (Fig. 2b), we estimate the efficiency of incorporation for the α-ketobutyrate-derived ^{13}C methyl probe to be 51 ± 7 % in a protonated background (see Supporting Information). The power of TROSY to obtain spectra of high-molecular weight species can only be exploited in the context of partial or full deuteration (Gardner et al. 1997; Wider and Wüthrich 1999; Ruschak and Kay 2010), which eliminates dipolar relaxation effects of surrounding protons on a given ^{13}C-methyl spin system. To assess simultaneous ^{13}C methyl labeling and perdeuteration in our system, we made samples of MBP in both *P. pastoris* and *E. coli*. We quantified the level of Ile δ1-methyl labeling in *P. pastoris*-derived deuterated MBP by comparing intensities to a concentration-matched *E. coli* sample (with assumed full incorporation at Ile δ1-methyl sites), yielding a labeling efficiency of 45 ± 6 % (Figure S2). The total deuteration level of *P. pastoris*-expressed MBP was estimated at 90 % through ESI-LC–MS analysis (Figure S3; a comparison of labeling efficiency and

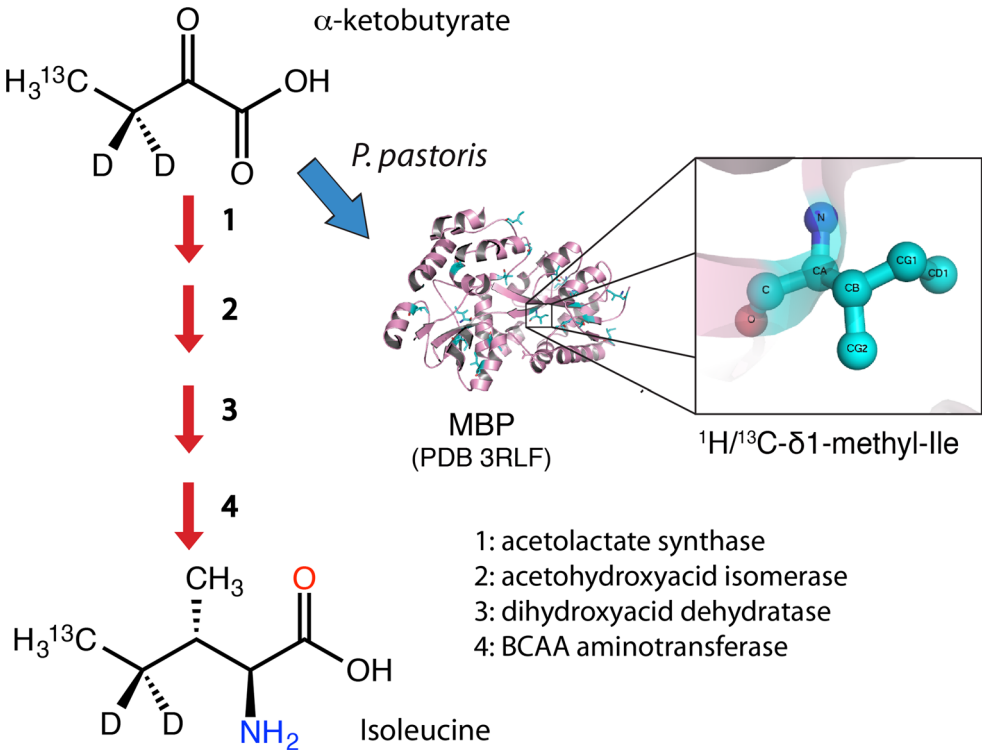

Fig. 1 Incorporation of ^{13}C-methyl groups at the δ1 position of isoleucine residues in proteins expressed in *P. pastoris*

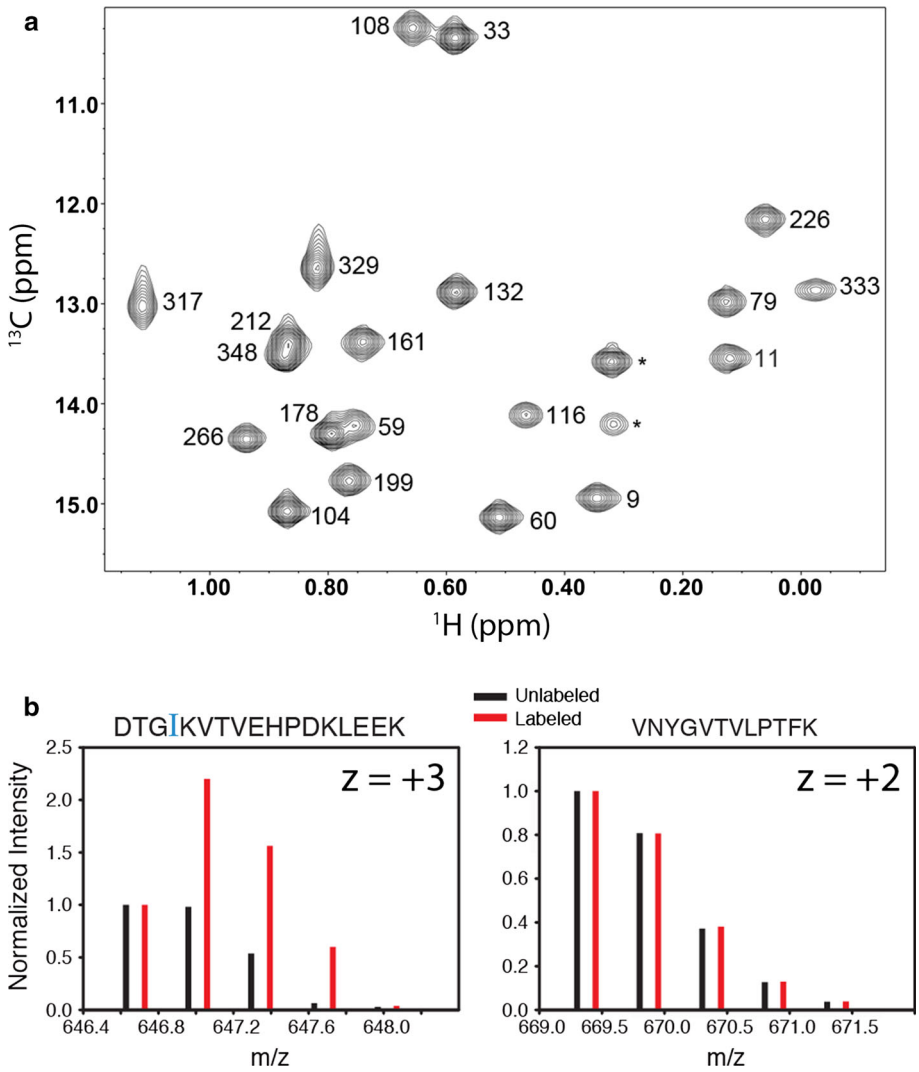

Fig. 2 Labeling of δ1-methyl groups of MBP expressed in *Pichia pastoris*. **a** ^{1}H-^{13}C HSQC spectrum of 225 μM MBP labeled with α-ketobutyrate. Spectrum was recorded at 25 °C on a Varian 800 MHz spectrometer. Peaks corresponding to Ile δ1-methyl groups are labeled in reference to assigned spectra (Gardner et al. 1998). Two unassigned peaks likely arising from differences in constructs are denoted with an asterisk. **b** Mass spectra of tryptic peptides containing Ile (*left*) and not containing Ile (*right*)

yields of recombinant MBP from *P. pastoris* vs. *E. coli* is shown in Figure S4). Addition of α-ketoisovalerate led to very modest labeling of leucine δ- and valine γ-methyl groups (<5 %, not shown), suggesting that labeling of these sites would require significant optimization, perhaps through cytoplasmic overexpression of branched-chain-amino-acid aminotransferase as reported for a *K. lactis* expression system (Miyazawa-Onami et al. 2013).

The impetus for using *P. pastoris* for ^{13}C methyl labeling is to access proteins that are not amenable to expression and purification from *E. coli*—for example, the eukaryotic cytoskeletal protein actin. Actin's capacity to change between monomeric and polymeric states arises from its conformational dynamics between distinct globular and filamentous forms (Oda et al. 2009; Pollard and Cooper 1986). NMR dynamics measurements would represent a significant new tool to study the biophysics of actin polymerization and interactions with regulatory molecules (Schmid et al. 2004; Kudryashov and Reisler 2013). While structures of actin monomers have been determined by X-ray crystallography (Otterbein et al. 2001; Rould et al. 2006; Nair et al. 2008) and actin filaments have been characterized by electron microscopy (Fujii et al. 2010; Ecken et al. 2015), expression of isotopically labeled actin for NMR has not been reported. Actin cannot be expressed at high levels in *E. coli* because of the lack of eukaryotic chaperone systems that are necessary for folding.

Biophysical characterization of actin is intrinsically difficult because actin polymerizes at concentrations above 100 nM. We therefore attempted to express a non-

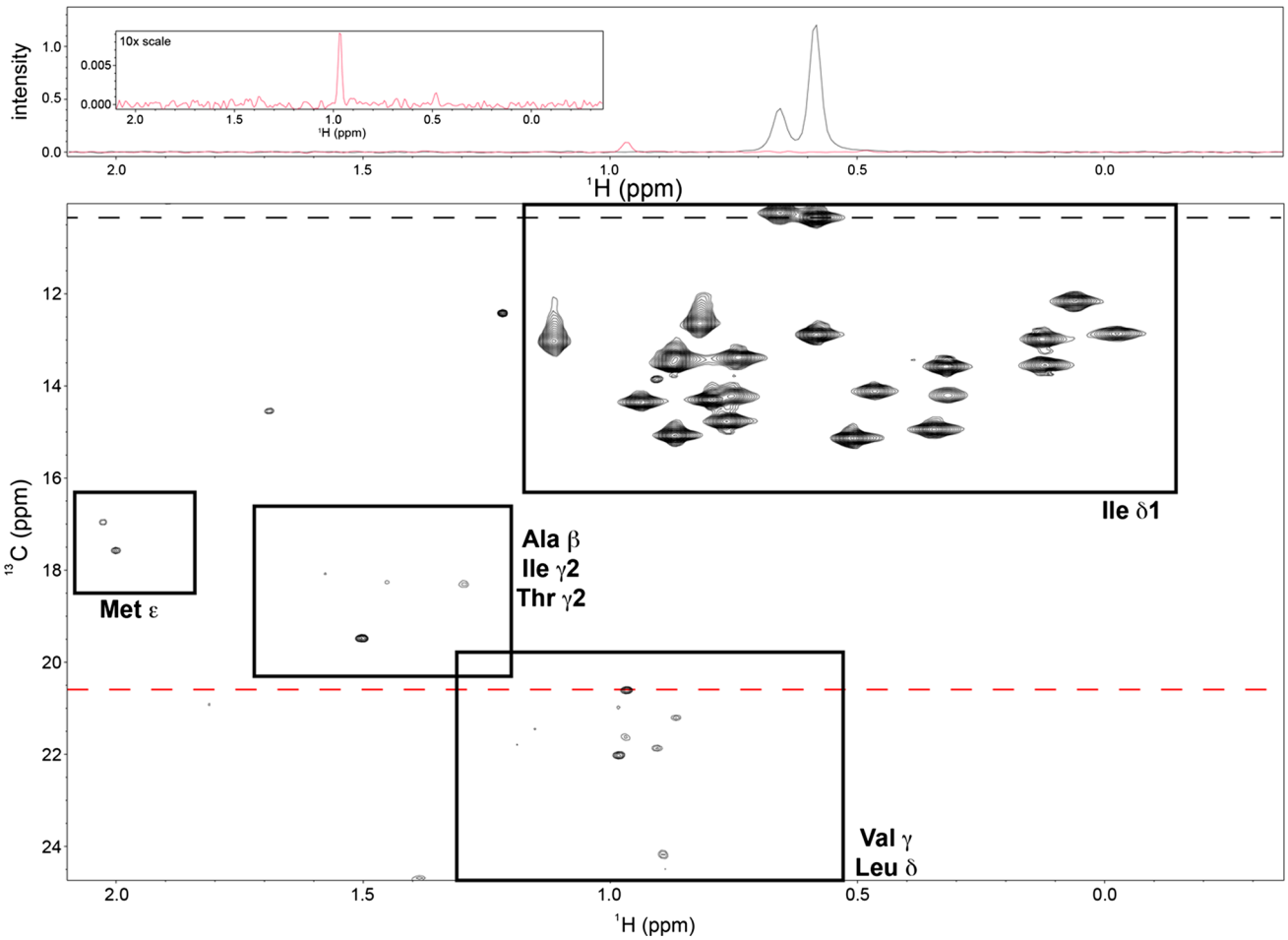

Fig. 3 Expanded view of the ^{1}H-^{13}C HSQC spectrum of isoleucine δ1-methyl labeled maltose binding protein shown in Fig. 2a. *Top panel* shows horizontal slices of the 2D dataset (*bottom panel*), taken at approximately ^{13}C = 10.3 ppm (*black*; Ile δ1) and 20.6 ppm (*red*; Val/Leu) to show representative signal-to-noise in the spectrum for the labeled Ile δ1 methyl groups versus the unlabeled (natural abundance ^{13}C) methyl groups of other amino acids. *Inset* of the *top panel* shows the ^{13}C = 20.6 ppm trace at 10× vertical scale of the surrounding panel to provide a clearer sense of signal-to-noise. Signal-to-noise measurements for all 22 Ile δ1-methyl peaks resulted in an average S/N ratio of 280

polymerizable Drosophila 5C actin (51.5 kDa, 94 % identity to human actin) mutant in *P. pastoris* with mutations that impair the fast growing "barbed-end" of the filament (Zahm et al. 2013). However, the mutant proved toxic, presumably because it interferes with the polymerization of endogenous actin. To solve this problem, we generated a C-terminal fusion to human thymosin β4, an actin binding protein that blocks the intact, slow-growing "pointed-end" and thus ameliorates toxicity (Noguchi et al. 2007). This strategy resulted in high expression levels (10 mg/L of culture) and enabled purification to homogeneity (Supporting Information and Fig. S5).

A representative HMQC (methyl TROSY) spectrum of ^{13}C-Ile-δ1-methyl actin is shown in Fig. 4a. Notably, for a protein with 30 Ile residues, we observe 33 peaks in the ^{1}H-^{13}C spectrum, likely reflecting slow chemical exchange processes at some sites. Taking advantage of the ability to highly deuterate proteins in *P. pastoris*, we repeated expression of *Drosophila* 5C actin in cultures where cells were adapted to D_2O-containing media prior to induction, resulting in 2.5 mg/L of ^{13}C-Ile-δ1-methyl perdeuterated actin. Lines in the ^{1}H-^{13}C HMQC spectra of the deuterated sample were much narrower than those in the HMQC spectrum of non-deuterated actin (Fig. 4a, b).

Future use of TROSY NMR methods to study the dynamics of high-MW mammalian protein complexes and membrane proteins will depend on the tractability of isotope incorporation. We have demonstrated efficient incorporation

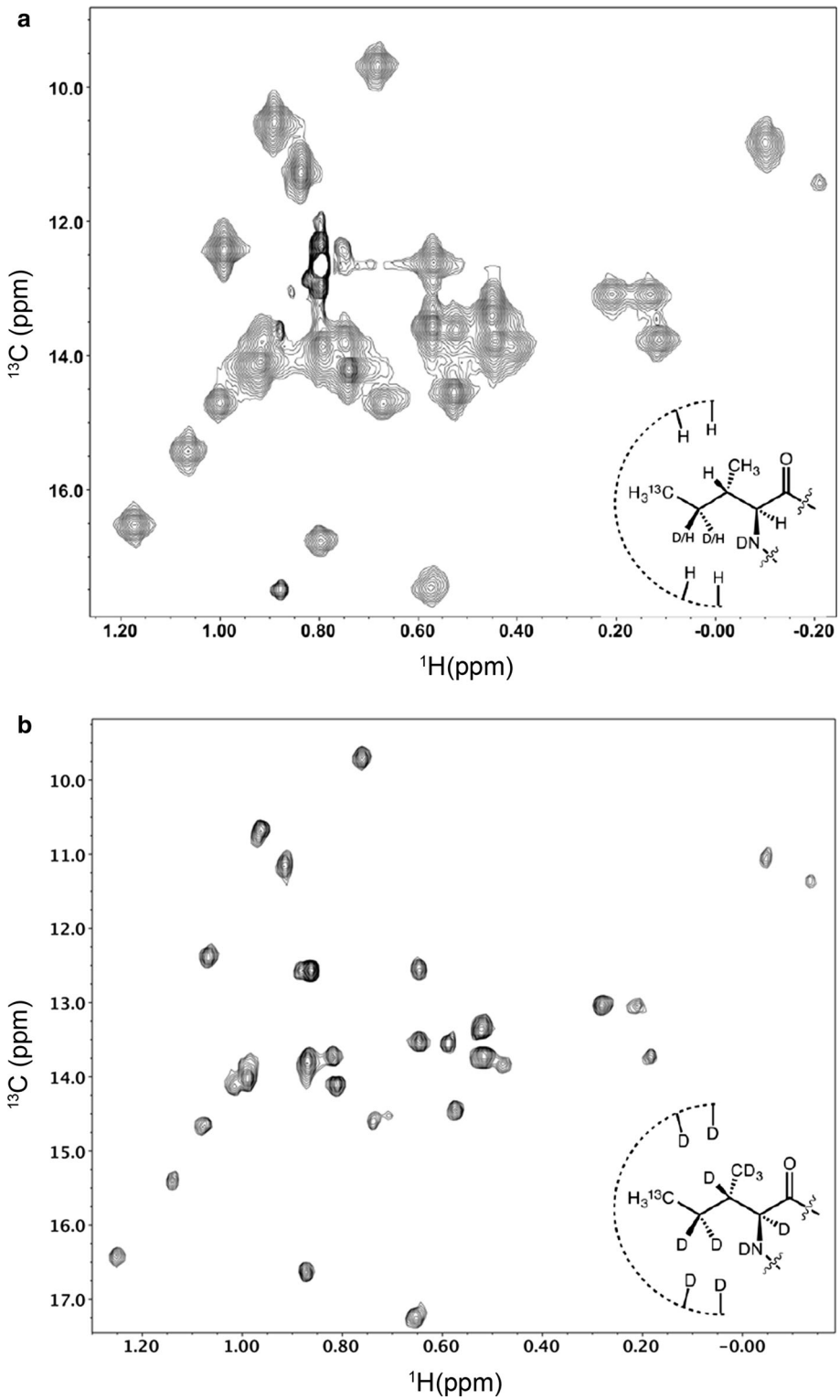

Fig. 4 NMR spectra of *Drosophila* actin labeled and overexpressed in *Pichia pastoris*. **a** ^{1}H-^{13}C HMQC spectrum of ^{13}C-Ile δ1-methyl-labeled actin (180 μM). **b** TROSY-HMQC spectrum of perdeuterated, ^{13}C-Ile δ1-methyl-labeled actin (150 μM). Spectra were recorded at 25 °C on a Varian 800 MHz spectrometer

of ^{13}C at the Ile δ1-methyl groups of proteins expressed in *P. pastoris*, a robust eukaryotic expression host. In conjunction with perdeuteration, we acquired high-quality ^{1}H-^{13}C methyl TROSY spectra on *Drosophila* actin, which were unobtainable before. This development, along with similar approaches using other yeast systems (Miyazawa-Onami et al. 2013), will allow 2D NMR spectroscopy to be applied to many previously intractable proteins.

Acknowledgments Funding was provided by a National Science Foundation Predoctoral Fellowship (Grant No. 1000136529 to L.C.), the Welch Foundation (I-1770 to D.M.R, I-1544 to M.K.R., I-1424 to K.H.G.), the Searle Scholars Program (D.M.R), a Packard Foundation Fellowship (D.M.R), the National Institutes of Health (T32 GM008297 supporting J.Z., R01 GM106239 to K.H.G., R01-GM56322 to M.K.R) and the Howard Hughes Medical Institute (M.K.R.).

References

Cereghino JL, Cregg JM (2000) Heterologous expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol Rev 24(1):45–66

Chen C-Y, Cheng C-H, Chen Y-C, Lee J-C, Chou S-H, Huang W, Chuang W-J (2005) Preparation of amino-acid-type selective isotope labeling of protein expressed in *Pichia pastoris*. Proteins 62(1):279–287

Ecken J, Müller M, Lehman W, Manstein DJ, Penczek PA, Raunser S (2015) Structure of the F-actin–tropomyosin complex. Nature 519(7541):114–117

Fan Y, Shi L, Ladizhansky V, Brown LS (2011) Uniform isotope labeling of a eukaryotic seven-transmembrane helical protein in yeast enables high-resolution solid-state NMR studies in the lipid environment. J Biomol NMR 49(2):151–161

Fujii T, Iwane AH, Yanagida T, Namba K (2010) Direct visualization of secondary structures of F-actin by electron cryomicroscopy. Nature 467(7316):724–728

Gardner KH, Kay LE (1997) Production and incorporation of 15 N, 13C, 2H (1*H*-δ1 methyl) isoleucine into proteins for multidimensional NMR studies. J Am Chem Soc 119(32):7599–7600

Gardner KH, Rosen MK, Kay LE (1997) Global folds of highly deuterated, methyl-protonated proteins by multidimensional NMR. Biochemistry 36(6):1389–1401

Gardner KH, Zhang X, Gehring K, Kay LE (1998) Solution NMR Studies of a 42 KDa *Escherichia coli* maltose binding protein/β-cyclodextrin complex: chemical shift assignments and analysis. J Am Chem Soc 120(45):11738–11748

Gossert AD, Hinniger A, Gutmann S, Jahnke W, Strauss A, Fernández C (2011) A simple protocol for amino acid type selective isotope labeling in insect cells with improved yields and high reproducibility. J Biomol NMR 51(4):449–456

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-labeled proteins. J Biomol NMR 13(4):369–374

Hansen AP, Petros AM, Mazar AP, Pederson TM, Rueter A, Fesik SW (1992) A practical method for uniform isotopic labeling of recombinant proteins in mammalian cells. Biochemistry 31(51):12713–12718

Hino T, Arakawa T, Iwanari H, Yurugi-Kobayashi T, Ikeda-Suno C, Nakada-Nakura Y, Kusano-Arai O, Weyand S, Shimamura T, Nomura N et al (2012) G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody. Nature 482(7384):237–240

Kofuku Y, Ueda T, Okude J, Shiraishi Y, Kondo K, Mizumura T, Suzuki S, Shimada I (2014) Functional dynamics of deuterated β2-adrenergic receptor in lipid bilayers revealed by NMR spectroscopy. Angew Chem Int Ed 53(49):13376–13379

Kudryashov DS, Reisler E (2013) ATP and ADP actin states. Biopolymers 99(4):245–256

Lee J-Y, Urbatsch IL, Senior AE, Wilkens S (2002) Projection structure of *P*-glycoprotein by electron microscopy: evidence for a closed conformation of the nucleotide binding domains. J Biol Chem 277(42):40125–40131

Li Z, Leung W, Yon A, Nguyen J, Perez VC, Vu J, Giang W, Luong LT, Phan T, Salazar KA et al (2010) Secretion and proteolysis of heterologous proteins fused to the *Escherichia coli* maltose binding protein in *Pichia pastoris*. Protein Expr Purif 72(1):113–124

Miyazawa-Onami M, Takeuchi K, Takano T, Sugiki T, Shimada I, Takahashi H (2013) Perdeuteration and methyl-selective (1)H, (13)C-labeling by using a *Kluyveromyces lactis* expression system. J Biomol NMR 57(3):297–304

Morgan WD, Kragt A, Feeney J (2000) Expression of deuterium-isotope-labelled protein in the yeast *Pichia pastoris* for NMR studies. J Biomol NMR 17(4):337–347

Nair UB, Joel PB, Wan Q, Lowey S, Rould MA, Trybus KM (2008) Crystal structures of monomeric actin bound to cytochalasin D. J Mol Biol 384(4):848–864

Noguchi TQP, Kanzaki N, Ueno H, Hirose K, Uyeda TQP (2007) A novel system for expressing toxic actin mutants in dictyostelium and purification and characterization of a dominant lethal yeast actin mutant. J Biol Chem 282(38):27721–27727

Nygaard R, Zou Y, Dror RO, Mildorf TJ, Arlow DH, Manglik A, Pan AC, Liu CW, Fung JJ, Bokoch MP et al (2013) The dynamic process of β2-adrenergic receptor activation. Cell 152(3):532–542

Oda T, Iwasa M, Aihara T, Maéda Y, Narita A (2009) The nature of the globular- to fibrous-actin transition. Nature 457(7228):441–445

Ollerenshaw JE, Tugarinov V, Kay LE (2003) Methyl TROSY: explanation and experimental verification. Magn Reson Chem 41(10):843–852

Otterbein LR, Graceffa P, Dominguez R (2001) The crystal structure of uncomplexed actin in the ADP state. Science 293(5530):708–711

Pervushin K, Riek R, Wider G, Wüthrich K (1997) Transverse relaxation-optimized spectroscopy (TROSY) for NMR studies of aromatic spin systems in ^{13}C-labeled proteins. Proc Natl Acad Sci USA 94(23):12366–12371

Pollard TD, Cooper JA (1986) Actin and actin-binding proteins. A critical evaluation of mechanisms and functions. Annu Rev Biochem 55:987–1035

Rosen MK, Gardner KH, Willis RC, Parris WE, Pawson T, Kay LE (1996) Selective methyl group protonation of perdeuterated proteins. J Mol Biol 263(5):627–636

Rould MA, Wan Q, Joel PB, Lowey S, Trybus KM (2006) Crystal structures of expressed non-polymerizable monomeric actin in the ADP and ATP states. J Biol Chem 281(42):31909–31919

Ruschak AM, Kay LE (2010) Methyl groups as probes of supra-molecular structure, dynamics and function. J Biomol NMR 46(1):75–87

Schmid MF, Sherman MB, Matsudaira P, Chiu W (2004) Structure of the acrosomal bundle. Nature 431(7004):104–107

Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, Abagyan R, Cherezov V, Liu W, Han GW et al (2011) Structure of the human histamine H1 receptor complex with doxepin. Nature 475(7354):65–70

Strauss A, Bitsch F, Fendrich G, Graff P, Knecht R, Meyhack B, Jahnke W (2005) Efficient uniform isotope labeling of Abl kinase expressed in Baculovirus-infected insect cells. J Biomol NMR 31(4):343–349

Strugatsky D, Gottschalk K-E, Goldshleger R, Bibi E, Karlish SJD (2003) Expression of Na + , K + -ATPase in *Pichia pastoris*: analysis of wild type and D369 N mutant proteins by Fe^{2+}-catalyzed oxidative cleavage and molecular modeling. J Biol Chem 278(46):46064–46073

Werner K, Richter C, Klein-Seetharaman J, Schwalbe H (2008) Isotope labeling of mammalian GPCRs in HEK293 cells and characterization of the C-terminus of bovine rhodopsin by high resolution liquid NMR spectroscopy. J Biomol NMR 40(1):49–53

Wider G, Wüthrich K (1999) NMR spectroscopy of large molecules and multimolecular assemblies in solution. Curr Opin Struct Biol 9(5):594–601

Zahm JA, Padrick SB, Chen Z, Pak CW, Yunus AA, Henry L, Tomchick DR, Chen Z, Rosen MK (2013) The bacterial effector VopL organizes actin into filament-like structures. Cell 155(2):423–434