J|A|C|S
COMMUNICATIONS

Published on Web 04/13/2004

Simplification of Protein NOESY Spectra Using Bioorganic Precursor Synthesis and NMR Spectral Editing

Roman Lichtenecker,[†] Martin L. Ludwiczek,[‡] Walther Schmid,*[,†] and Robert Konrat*[,‡]

Institute of Organical Chemistry, Währingerstrasse 38, A-1090 Vienna, Austria, and Institute of Theoretical Chemistry and Molecular Structural Biology, University of Vienna, Rennweg 95b, A-1030 Vienna, Austria

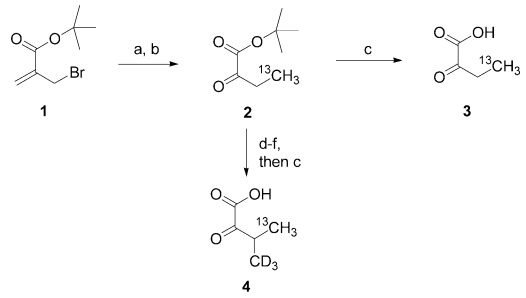
Received January 18, 2004; E-mail: walther.schmid@univie.ac.at; robert.konrat@univie.ac.at

In the past decade, NMR spectroscopy has evolved into an efficient tool for structure determination of proteins in solution.[1,2] Conventional NMR structure determination relies on a large number of nuclear Overhauser enhancement (NOE) restraints. Typically, a large number of restraints per residue (e.g., 15−20) is required to obtain three-dimensional protein structures of reasonable accuracies and precisions. Yet, in practice, the number of obtainable restraints is considerably smaller, predominantly due to severe spectral overlap and because of line broadening. A major advance in this regard has been the introduction of highly deuterated, ^{15}N,^{13}C-labeled protein samples which have attenuated transverse relaxation rates and lead to significant gains in sensitivity and spectral resolution.[3,4] Another important improvement has been the development of the TROSY-detection scheme.[5] Applications include the 110 kDa homooctameric protein, 7,8-dihydroneopterin aldolase,[6] or NMR investigations of the 723-residue monomeric enzyme malate synthase G from *E. coli* (81.4 kDa).[7,8] Recently, NMR methods development has focused on correlation spectroscopy for side chain methyl groups.[9] First, methyl groups have favorable relaxation properties and thus give rise to intense correlation peaks.[10] Second, given their location in hydrophobic cores of proteins and protein interaction interfaces, methyl NOEs are valuable sources for structural information.[11] Third, there are robust techniques for the incorporation of protonated Ile, Leu, and Val methyl groups in highly deuterated ^{13}C,^{15}N-labeled proteins.[12] Finally, they are excellent reporters of protein dynamics and provide unique insight into functionally relevant protein motions.[13]

Building on the success of these experiments, different labeling strategies for methyl groups in proteins were developed.[14] We introduce here a novel synthetic route to precursors for the methyl-bearing amino acids valine, leucine, and isoleucine, respectively. Although the precursor compounds are related to recent work by Kay et al., no synthetic details have been provided by the authors.[9b] Additionally, the particular benefit of our synthetic concept is that we can selectively incorporate side chains with unique ^{13}C,^{12}C,^{2}H,^{1}H-labeling patterns in an otherwise uniformly ^{12}C,^{1}H-labeled (or alternatively ^{13}C,^{2}H-labeled) protein.

In short, the synthetic strategy was as follows (Scheme 1). We started our synthesis of the labeled amino acid precursors **3** and **4** with the easily accessible *tert*-butyl α-bromomethacrylate **1**.[15] Treatment of *tert*-butyl ester **1** with freshly prepared ^{13}C-methyl magnesium iodide followed by ozonolysis of the double bond made α-ketoester **2** accessible. After treatment of derivative **2** with gaseous hydrochloric acid, the selectively ^{13}C-labeled 2-oxo-butanoic acid **3** was obtained in 53% yield (starting from **1** in three steps). However, attempts to directly use the *tert*-butyl 2-oxo-butanoate **2** for the synthesis of ^{2}H- and ^{13}C-labeled 2-oxo-isovalerate failed. Thus, for the introduction of the d_3-methyl group

Scheme 1[a]



[a] Conditions: (a) Mg, $^{13}CH_3I$, Et_2O, room temperature, 2 h, 84%; (b) O_3, CH_2Cl_2, −78 °C, PPh_3, 75%; (c) HCl $_{vap}$, CH_2Cl_2/Et_2O, 83%; (d) $H_2NN(CH_3)_2$, Et_2O, room temperature, 24 h, 95%; (e) LDA, CD_3I, THF, −78 °C, 3 h, 80%; (f) 1 N HCl, THF, Et_2O, room temperature, 1 h, 92%.

present in target compound **4**, we transformed the 2-oxo-derivative by the addition of *N*,*N*-dimethylhydrazine to its corresponding hydrazone,[15] which was further alkylated using d_3-methyl iodide. The target α-ketoisovaleric acid **4** was obtained after the treatment with aqueous and finally gaseous hydrochloric acid in good yield.

As an application, we demonstrate $^{13}CH_3/^{15}N$ NOESY-editing for the quail lipocalin protein Q83. Using differential hybridization techniques, the Q83 cDNA clone was isolated in quail embryo fibroblasts transformed by the v-*myc* oncogene.[16] Recently, the solution structure of the secreted form of Q83 (157 residues) was solved by NMR spectroscopy.[16] Q83 folds into a single globular domain of the lipocalin type. The central part consists of an eight-stranded up-and-down β-barrel flanked by an N-terminal 3_{10} and a C-terminal α-helix, respectively. The geometry of the fold is additionally stabilized by a characteristic (and conserved within the lipocalin protein family) disulfide bridge between Cys59 and Cys152.

Q83 was expressed and purified as described elsewhere.[17] The minimal medium, however, contained 2 g of unlabeled (^{12}C,^{1}H) D-Glucose (CIL) and ^{15}N NH_4Cl (CIL). For selective $^{13}CH_3$-labeling of leucines, valines, and isoleucines, we followed the well-established procedure described by Kay and co-workers. In contrast to their strategy which involves uniformly ^{13}C-labeled precursor compounds, the synthetic scheme of Figure 1 allowed for selective ^{13}C-labeling only at the terminal carbon of the amino acid side chains. Seventy-five milligrams of 2-keto-4-^{13}C-butyrate and 100 mg of 2-keto-3-methyl-d_3-4-^{13}C-butyrate were added to the growth medium. Deuteration of 2-keto-3-methyl-d_3-4-^{13}C-butyrate at position 3 can be achieved following the procedure of Kay and co-workers.[12] The NMR sample was 1 mM in protein, 90% H_2O/10% D_2O, 20 mM phosphate, pH 6.5, 100 mM NaCl, and 1 mM DTT, 26 °C.

Figure 1 shows the first trace of a simultaneous 3D ^{13}C,^{15}N NOESY-HSQC.[17] From inspection of the 1D projection, it can be seen that the precursor compounds α-ketobutyrate and α-keto-

[†] Institute of Organic Chemistry.
[‡] University of Vienna.

10.1021/ja049679n CCC: $27.50

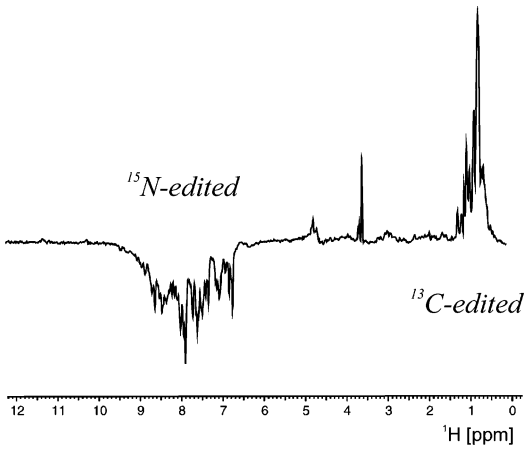


Figure 1. First trace from the simultaneous 2D $^{13}C,^{15}N$-NOESY-HSQC of the {[I(δ1 only), L($^{13}CH_3$,U-^{12}C), V($^{13}CH_3$,U-^{12}C)]} U-[$^{15}N,^{12}C,^{1}H$] sample of Q83. The biosynthetic precursor compounds used for selective labeling of Ile, Leu, and Val are described in Scheme 1.

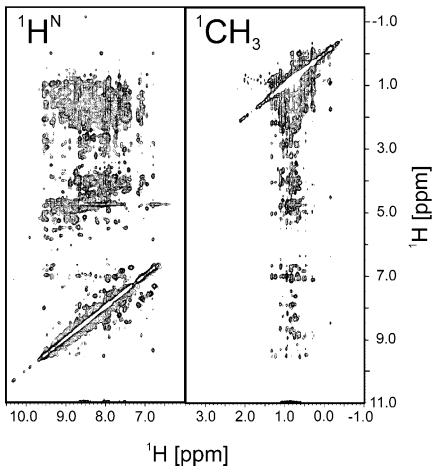


Figure 2. Isotope-editing NOESY using chemically synthesized precursor compounds, (left) ^{15}N- and (right) ^{13}C-filtered. The spectrum was obtained using the simultaneous $^{13}C,^{15}N$ NOESY-HSQC pulse sequence and omitting the heteronuclei evolution period.[17] Experimental parameters were as follows: 256 increments, 32 transients, relaxation delay 2 s; NOESY mising time: 150 ms. The experiments were performed on a Varian Unity Inova 800 MHz spectrometer.

isovalerate are effectively transformed into Val, Leu, and Ile amino acid side chains. Our findings are in accordance with previous results obtained by Kay and co-workers.[12] They have already shown that alternative pathways are of no relevance for these precursor compounds and that the incorporation levels are between 90% and 95%.[12]

The simplified NOESY spectrum is illustrated in Figure 2. Of note is the reduced complexity of the NOESY spectrum in the aliphatic region because of the reduced number of intraresidue methyl−methyl NOEs. This facilitates the assignment of NOE connectivities in this typically overcrowded part of the spectrum. The only observable intraresidue methyl NOEs are as follows: Leu, H_α,H_β;Val, H_α; Ile, H_β,$H_{\gamma 1}$. Note that in D_2O growth medium only interresidue $^{13}CH_3$ methyl NOEs would be observed.

In contrast to existing labeling schemes and applications where only methyl−methyl and methyl−H^N NOEs are observable,[11,14] the new labeling pattern allows for the detection of methyl−methyl and methyl−amide NOEs as well as NOEs between methyl and amide protons and Hα and side chain protons. Discrimination between NOEs involving protons attached to either ^{13}C or ^{15}N and NOEs between protons belonging to different proton subsets (attached to $^{13}C,^{15}N$, or ^{12}C) is straightforward using NMR methodology developed for studies of protein−ligand complexes.[18] Specifically, intra- and intermolecular NOEs can be simultaneously extracted using a recently developed pulse scheme employing adiabatic ^{13}C inversion pulses.[19] An analogous strategy is thus suggested for simultaneous observation of NOEs between $^{13}C/^{15}N-^{13}C/^{15}N$ and $^{13}C/^{15}N-^{12}C$ protons in differentially labeled proteins.

In conclusion, we have developed a novel route to selective $^{13}C,^{1}H$-labeled precursor compounds for Ile, Leu, and Val side chains. The synthesis is very flexible and allows for a large variety of different labeling patterns in the side chains with many applications. As a first application, we demonstrated the simplification of protein NOESY spectra by eliminating intraresidue methyl−methyl NOEs while, at the same time, retaining structurally important NOEs between methyl and amide protons and Hα and side chain protons. The isolated $^{13}CH_3$ groups can also be used as novel probes for side chain motions, and applications involving both $^{1}H-^{1}H$ and $^{1}H-^{13}C$ multiple-quantum coherences are currently underway in our laboratory.

Acknowledgment. We thank Prof. Klaus Bister (Department of Biochemistry, University of Innsbruck) for the Q83 plasmid. This work is supported by the Austrian Science Foundation FWF (P15578 to R.K. and ÖNB 7799 to W.S.), and the WWTF (project LS162 to R.K. and W.S.).

Supporting Information Available: Experimental details of the chemical synthesis and chemical characterization of precursor compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Prestegard, J. H.; Valafar, H.; Glushka, J.; Tian, F. *Biochemistry* **2001**, *40*, 8677−8685.
(2) Wüthrich, K. *J. Biomol. NMR* **2003**, *27*, 13−39.
(3) Gardner, K. H.; Kay, L. E. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 357−406.
(4) Farmer, B. T.; Venters, R. A. In *NMR of Perdeuterated Large Proteins*; Krishna, N. R., Berliner, L. J., Eds.; Kluwer Academic/Plenum Publishers: New York, 1998; Vol. 16, pp 75−120.
(5) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12366−12371.
(6) Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K. *J. Am. Chem. Soc.* **2000**, *122*, 7543−7548.
(7) (a) Tugarinov, V.; Muhandiram, R.; Ayed, A.; Kay, L. E. *J. Am. Chem. Soc.* **2002**, *124*, 10025−10035. (b) Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 10420−10428.
(8) Tugarinov, V.; Kay, L. E. *J. Mol. Biol.* **2003**, *327*, 1121−1133.
(9) (a) Tugarinov, V.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 5701−5706. (b) Tugarinov, V.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 13868−13878.
(10) Gardner, K. H.; Rosen, M. K.; Kay, L. E. *Biochemistry* **1997**, *36*, 1389−1401.
(11) (a) Rosen, M. K.; Gardner, K. H.; Willis, R. C.; Parris, W. E.; Pawson, T.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 627−636. (b) Metzler, W. J.; Wittekind, M.; Goldfarb, V.; Mueller, L.; Farmer, B. T. *J. Am. Chem. Soc.* **1996**, *118*, 6800−6801.
(12) Goto, N. K.; Gardner, K. H.; Mueller, G. A.; Willis, R. C.; Kay, L. E. *J. Biomol. NMR* **1999**, *13*, 369−374.
(13) Mulder, F. A.; Mittermaier, A.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *Nat. Struct. Biol.* **2001**, *8*, 932−935.
(14) (a) Hajduk, P. J.; Augeri, D. J.; Mack, J.; Mendoza, R.; Yang, J.; Betz, S. F.; Fesik, S. W. *J. Am. Chem. Soc.* **2000**, *122*, 7898−7904. (b) Gross, J. D.; Gelev, V. M.; Wagner, G. *J. Biomol. NMR* **2003**, *25*, 235−242.
(15) (a) Giese, B.; Linker, T. *Synthesis* **1992**, 46−48. (b) Villieras, J.; Rambaud, M. *Synthesis* **1982**, 924−926.
(16) (a) Kontaxis, G.; Matt, T.; Schüler, W.; Kräutler, B.; Bister, K.; Konrat, *J. Biomol. NMR* **2000**, *17*, 177−178. (b) Hartl, M.; Matt, T.; Schüler, W.; Siemeister, G.; Kontaxis, G.; Kloiber, K.; Konrat, R.; Bister, K. *J. Mol. Biol.* **2003**, *333*, 33−46.
(17) Pascal, S. M.; Muhandiram, D. R.; Yamazaki, T.; Forman-Kay, J. D.; Kay, L. E. *J. Magn. Reson., Ser. B* **1994**, *103*, 197−201.
(18) Breeze, A. *Prog. Nucl. Magn. Reson. Spectrosc.* **2000**, *36*, 323−372.
(19) Eichmüller, C.; Schüler, W.; Konrat, R.; Kräutler, B. *J. Biomol. NMR* **2001**, *21*, 107−116.

JA049679N