ORGANIC LETTERS

2007
Vol. 9, No. 24
4935−4937

A Simple and Economical Method for the Production of ^{13}C,^{18}O-Labeled Fmoc-Amino Acids with High Levels of Enrichment: Applications to Isotope-Edited IR Studies of Proteins

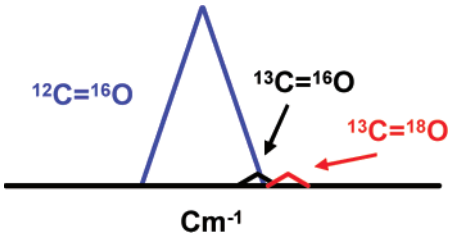
James Marecek,*,† BenBen Song,† Scott Brewer,‡ Jenifer Belyea,‡ R. Brian Dyer,‡ and Daniel P. Raleigh*,†

Department of Chemistry, State University of New York at Stony Brook, Stony Brook, New York 11794-3400, and Physical Chemistry and Applied Spectroscopy Group, Los Alamos National Laboratory, Mail Stop J567, Los Alamos, New Mexico 87545

draleigh@notes.cc.sunysb.edu; james.mareck@stonybrook.edu

Received August 6, 2007

ABSTRACT



Isotope-edited IR of proteins has generated considerable interest. Double labeling with ^{13}C and ^{18}O with high levels of isotopic enrichment is required for residue-specific resolution. Current methods for the preparation of doubly labeled amino acids give modest ^{18}O enrichment, limiting the utility of the approach. We report a simple and economical method for preparing ^{13}C,^{18}O-doubly labeled *N*-(9-fluorenylmethoxycarbonyl)-amino acids with high levels of enrichment for residues that do not require acid-labile side-chain protecting groups.

An ideal probe of protein folding, protein membrane interactions, and protein dynamics would offer site-specific resolution and easily interpretable changes in the spectroscopic signal of the probe, and be minimally invasive.[1] Fluorescence-based experiments are extensively used, but not all proteins contain Trp or Tyr, and their introduction by mutagenesis does not always represent a conservative substitution; furthermore, changes in fluorescence cannot be interpreted in terms of secondary structure. Alternatively, nonnatural fluorophores can be introduced; however, the added groups are usually large and can represent quite drastic perturbations. ESR spin labels studies have also been profitably employed, but again the addition of a spin label can represent a drastic change, and it is hardly clear that such substitutions can safely be considered to be conservative. In principle IR-based methods offer a very attractive probe of protein structure and dynamics. The frequency of the amide-I band is well-known to be sensitive to secondary structure and to backbone solvent interactions. Recently developed two-dimensional infrared (2DIR) methods have generated considerable excitement and promise to extend IR investigations of proteins.[1−3] However the FTIR spectrum of even the smallest proteins consists of an inhomogeneously broadened peak, and site-specific information is not obtained. 2DIR is also compro-

† State University of New York at Stony Brook, Stony Brook.
‡ Los Alamos National Laboratory.

(1) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1985**, *38*, 1811.
(2) Zanni, M.; Gnanakaran, S.; Stenger, J.; Hochstrasser, R. *J. Phys. Chem. B.* **2001**, *105*, 6520.
(3) Mukamel, S. *Ann. Rev. Phys. Chem.* **2000**, *51*, 691.

10.1021/ol701913p CCC: $37.00
Published on Web 10/25/2007

mised by limited resolution. Fortunately, the amide-I band is very sensitive to isotopic substitution at the carbonyl carbon and carbonyl oxygen, since the band can, to a first approximation, be viewed as the sum of individual carbonyl stretching modes. Thus, isotopic substitution will result in a shift of the amide-I band of a labeled site to lower frequencies, opening the door to isotope-edited IR studies. Initial experiments made use of ^{13}C substitution which results in a ca. 40 cm^{-1} shift.[4,5] While useful for the study of small to midsize peptides, the approach suffers from serious drawbacks when applied to proteins or even large peptides. The modest shift induced by a ^{13}C substitution coupled with the 1.1% natural abundance of ^{13}C means that the shifted peak will normally not be baseline resolved and will appear as a shoulder on a large broad band. Recently, $^{13}C{=}{=}^{18}O$ labeling of proteins and peptides has been used to partially circumvent these problems.[6−10] The larger isotope shift induced by the double label, expected to be on the order of 75 cm^{-1} for a purely local oscillator, leads to significantly improved resolution provided that derivatives with a high level of ^{13}C and ^{18}O enrichment are used. This allows true isotope-edited IR studies and allows the full power of 2DIR methods to be exploited. The approach obviously requires ^{13}C,^{18}O-labeled amino acids. While, ^{13}C carbonyl-labeled amino acids are commercially available, the required ^{13}C,^{18}O-labeled variants are not, and less than optimum ^{18}O enrichment can cause major problems.[7] Consider, for example, the results of 75% enrichment with ^{18}O, a typical value obtained using current methods. In this case, a sample of a ^{13}C,^{18}O-labeled 45-residue polypeptide would have a $^{13}C{-}^{16}O$ peak of comparable integrated intensity to the double-labeled peak, resulting in spectral overlap and a significant loss of resolution. Protocols now in use to prepare double-labeled amino acids lead to ^{18}O enrichments on the order of 75 to 80% or less, depending upon the amino acid. Cost is also a critical issue given the expense of $H_2{}^{18}O$ and the large amount of material needed for solid-phase peptide synthesis. Here we report a simple and economical method for preparing ^{13}C,^{18}O-double-labeled *N*-(9-fluorenylmethoxycarbonyl) (FMOC)-amino acids with high levels of ^{18}O enrichment for residues that do not require acid-labile side-chain protecting groups, i.e., Gly, Ala, Val, Ile, Leu, Phe, Trp, and Pro. We demonstrate the utility of the method by preparing ^{13}C,^{18}O-double-labeled FMOC-Ala, -Val, -Ile, -Leu, and -Phe.

The usual approach for the preparation of labeled amino acid derivatives is to incorporate the ^{18}O label into the amino acid, and then perform further manipulations. A number of labeling protocols have been developed, but the most common one is equilibration of the amino acid with enriched water (1 M in HCl) at elevated temperatures for about 24 h.[11,12] We were concerned about loss of label during the preparation of the desired FMOC derivatives from a doubly labeled amino acid since we planned to use 9-fluorenylmethylchloroformate as the reagent for the incorporation of the Fmoc group with water/dioxane as the solvent and sodium carbonate as the base.[13,14] In our hands, we have also found that there are more byproducts generated in these reactions, and the yields are frequently lower (80−90%) than in the case of the preparation of *t*-BOC derivatives. For these reasons, we prepared the FMOC derivative of the ^{13}C-labeled amino acid first and then ran the exchange reaction with ^{18}O-enriched water. Preliminary control experiments indicated that the protected FMOC amino acid was reasonably stable for at least 24 h at 100° in mixtures of dioxane and water which were 0.1 M in HCl. Although there was a slight loss in overall enrichment, we found it more convenient to generate the HCl in situ by reacting acetyl chloride with the ^{18}O-enriched water. This allowed preparation of a precise acid titer without the need for titration and subsequent adjustments. The acetic acid hydrolysis byproduct was innocuous. ^{18}O-Enriched water is quite expensive; thus, the labeling was done in a two-step process. In the first equilibration step the ^{18}O content was raised to about 75−85%, depending on the enrichment level of the medium. A second equilibration brought the final enrichment level to 90−96% as judged by mass spectroscopy. The enrichment medium was usually 3:2 dioxane/$H_2{}^{18}O$ water which was 0.1 M in HCl. The enrichment medium after the second equilibration step was always recovered by a vacuum distillation and recycled back to the first stage since its ^{18}O content is ≥90%. The rate of ^{18}O incorporation was very sensitive to the steric bulk of the amino acid side chain. The reaction was very fast for alanine which required 3 h for complete equilibration and was very slow for valine and isoleucine which required 20−30 h or more for equilibration (Table 1). Recovery after this two-stage ^{18}O-enrichment process, except for isoleucine, was usually 90% of the starting FMOC-amino acid. The recovery for isoleucine was 80%, the losses being caused by decomposition during the exchange reaction. The final level of incorporation after the two-step procedure, was between 93 and 96% except for isoleucine at 90%.

The basic experimental protocol is outlined below. More detailed information is provided in the Supporting Information. A 0.25 M solution of HCl in the ^{18}O-enriched water was prepared by reacting acetyl chloride (196 mg, 2.5 mmol, 180 uL) with 10.0 mL of the water. The ^{13}C-enriched FMOC L-amino acid (10 mmol) was mixed with 15 mL of dioxane in a 100-mL three-neck flask equipped with a reflux condenser and a nitrogen inlet, and the mixture was heated to 100° (bath temperature) at which point the solution was homogeneous. The ^{18}O-enriched water was slowly added. A homogeneous solution usually resulted although partial

(4) Tadesse, L.; Nazarbaghi, R.; Walters, L. *J. Am. Chem. Soc.* **1991**, *113*, 7036.
(5) Decatur, S. *Acc. Chem. Res.* **2006**, *39*, 169.
(6) Torres, J.; Adams, P. D.; Arkin, I. T. *J. Mol. Biol.* **2000**, *300*, 677.
(7) Torres, J.; Kukol, A.; Goodman, J. M.; Arkin, I. T. *Biopolymers* **2001**, *59*, 396.
(8) Torres, J.; Briggs, J. A. G.; Arkin, I. T. *J. Mol. Biol.* **2002**, *316*, 365.
(9) Fang, C.; Wang, J.; Charnley, A. K.; Smith, A. B., III; Decatur, S. M.; Hochstrasser, R. M. *Chem. Phys. Lett.* **2003**, *382*, 586.
(10) Brewer, S. H.; Song, B.; Raleigh, D. P.; Dyer, R. B. *Biochemistry* **2007**, *46*, 3279.
(11) Murphy, R. C.; Clay, K. L. *Methods Enzymol.* **1990**, *193*, 338.
(12) Ponnusamy, E.; Jones, C. R.; Fiat, D. *J. Labelled Compd. Radiopharm.* **1987**, *24*, 773.
(13) Carpino, L. A.; Han, G. Y. *J. Org. Chem.* **1972**, *37*, 3404.
(14) Carpino, L. A. *Acc. Chem. Res.* **1987**, *20*, 401.

Table 1. Reflux Times Used for the Incorporation of ^{18}O into Each Amino Acid, and the Final Level of ^{18}O Enrichment

	time (h) for first exchange[a]	time (h) for second exchange	final ^{18}O enrichment (%)
Ala	3	not used[b]	87[b]
Val	20	20	93
Leu	10	10	94
Ile	34[c]	30[c]	90
Phe	10	10	95−96

[a] The level of ^{18}O labeling of the enrichment medium was 93−94% for the first exchange reaction. A 3:2 mixture of dioxin, ^{18}O water, 0.1 M HCl was used for Val, Leu, Ile, and Phe. A 1:1 mixture was used for Ala. [b] This level of enrichment was acceptable for the application of interest; thus, a second step was not required. [c] Complete equilibration was not achieved. The reaction was stopped to minimize decomposition of the FMOC-amino acid

solution proved to be sufficient. The mixture was stirred under nitrogen at a bath temperature of 100−102° for an appropriate time (Table 1). The reaction was monitored by removing a drop of the solution and evaporating to dryness in vacuo then analyzing by TLC. The enrichment level was checked by mass spectrometry. When the enrichment level was constant, the enrichment medium was recovered by evaporating the mixture in vacuo at room temperature using a short column and collecting the dioxane/water in a liquid-nitrogen-cooled trap. The solid residue was dried under vacuum for several hours and the enrichment process repeated a second time. The crude, enriched FMOC-amino acid was crystallized from a suitable solvent. Measurement of the optical rotation showed that no detectable racemization had occurred.

This simple and affordable approach leads to final ^{18}O-enrichment levels of 93−96% for Val, Phe, and Leu and 90% for Ile after the two-step enrichment process. Only a single enrichment step was required to generate 87% ^{18}O-labeled Ala (Table 1). Recovery after the two-stage ^{18}O-enrichment process, except for isoleucine, was between 90 and 94% of the starting FMOC-amino acid. The recovery for isoleucine was less, 80%, because of decomposition during the exchange reaction.

The power of site-specific $^{13}C{=}^{18}O$-labeling of proteins is illustrated by the temperature-dependent difference FTIR spectrum shown in Figure 1. We prepared a sample of the villin headpiece helical subdomain, HP-36, selectively labeled at Ala-57 via solid-phase peptide synthesis. HP-36 is a 36-residue three-helix protein which has become an extremely popular model system for experimental and computational studies of protein folding. HP-36 folds cooperatively and rapidly in isolation. The helical subdomain is part of the larger villin subdomain, and the numbering system used here corresponds to that of the full length villin headpiece domain. The first residue in the helical subdomain is Leu-42. The $^{13}C{=}^{18}O$-labeled Ala is in a helix and is buried

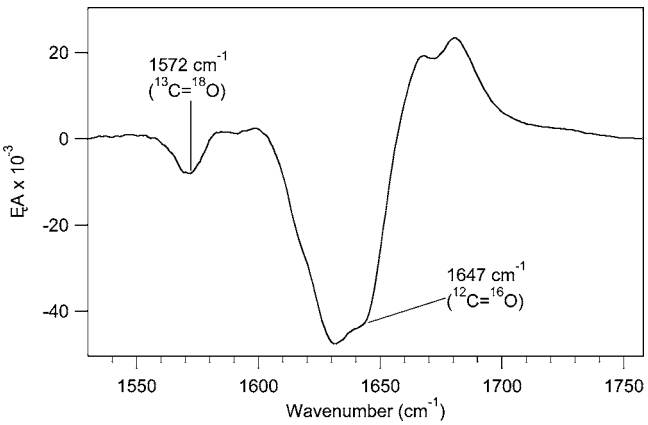


Figure 1. Temperature-dependent difference FTIR spectrum (76 − 16 °C) of $^{13}C{=}^{18}O$ A57 HP36. The peak due to the labeled Ala appears as a negative feature in the difference spectrum at 1572 cm^{-1} due to loss of the helical structure as the sample is heated. The negative feature at 1647 cm^{-1} corresponds to unlabeled ($^{12}C{=}^{16}O$) residues. As the protein unfolds, new peaks appear at 1601 and 1679 cm^{-1}, assigned to the disordered structure peaks for the labeled and unlabeled positions, respectively.

in the hydrophobic core in the folded state, a so-called buried helix position. The figure displays the temperature-dependent difference FTIR spectrum (76 − 16 °C) of thermally unfolded sample minus folded sample. The labeled site gives rise to a negative feature in the difference spectrum at 1572 cm^{-1} due to loss of the helical structure as the sample is heated. The negative feature at 1647 cm^{-1} corresponds to the unlabeled ($^{12}C{=}^{16}O$) residues in a buried helical conformation. As the protein unfolds, new peaks appear at 1601 and 1679 cm^{-1}, assigned to the disordered structure peaks for the labeled and unlabeled positions, respectively. Thus, for both the folded and unfolded conformations, the observed isotope shift for the double label is about 75 cm^{-1}, as expected for a local oscillator model of the C=O stretching vibration. The data illustrates two important points; first, one can easily distinguish a single-labeled position, and second, specific labeling can be used to follow the unfolding reaction. In particular, note that both the helical and disordered structure peaks for the labeled position are discernible in the data.

Acknowledgment. This work was supported by the NIH and NSF via Grants NIH GM70941, NSF MCB-614365 to D.P.R. and NIH GM53640 to R.B.D.

Supporting Information Available: Detailed experimental procedures and compound characterization. This material is available free of charge via the Internet at http://pubs.acs.org

OL701913P