Site-specific ^{19}F NMR chemical shift and side chain relaxation analysis of a membrane protein labeled with an unnatural amino acid

Pan Shi,¹ Hu Wang,² Zhaoyong Xi,² Chaowei Shi,¹
Ying Xiong,^{1*} and Changlin Tian^{1*}

¹National Laboratory for Physical Science at Microscale and School of Life Science, University of Science and Technology of China, Hefei, Anhui, People's Republic of China

²School of Chemistry, University of Science and Technology of China, Hefei, Anhui, People's Republic of China

Received 18 August 2010; Revised 7 October 2010; Accepted 12 October 2010

DOI: 10.1002/pro.545

Published online 15 November 2010 proteinscience.org

Abstract: Site-specific ^{19}F chemical shift and side chain relaxation analysis can be applied on large size proteins. Here, one-dimensional ^{19}F spectra and T_1 , T_2 relaxation data were acquired on a SH3 domain in aqueous buffer containing 60% glycerol, and a nine-transmembrane helices membrane protein diacyl-glycerol kinase (DAGK) in dodecyl phosphocholine (DPC) micelles. The high quality of the data indicates that this method can be applied to site-specifically analyze side chain internal mobility of membrane proteins or large size proteins.

Keywords: site-specific labeling; NMR; chemical shift; side chain relaxation; high viscosity; large size protein; membrane proteins

Introduction

Nuclear magnetic resonance (NMR) spectroscopy can provide a detailed, site-specific description of internal motions in proteins.¹ Most of protein side chain internal motions are analyzed through ^{13}C or ^2H

relaxation. However, acquisition of relaxation data is hampered as the sizes of the proteins increases, mainly due to difficulties of resonance assignment and relative low sensitivity of side chain isotopes.

^{19}F NMR spectroscopy has been widely applied to investigate protein structure change and dynamics, and ^{19}F NMR motional analysis of several large size proteins were reported, including a 100 kDa citrate synthase² and a 210 kDa trimeric pyruvate kinase,³ despite low signal-to-noise ratio and lack of ^{19}F peak assignment.

Recently, an unnatural amino acid, tri-fluoro-methyl-phenylalanine, was introduced to specific sites of a protein for further ^{19}F NMR studies.^{4–6} Briefly, the tfmF was recognized by an artificially

Abbreviations: DAGK, diacylglycerol kinase; DPC, dodecyl phosphocholine; NMR, nuclear magnetic resonance; tfmF p-tri-fluoromethyl-phenylalanine.

Grant sponsor: Chinese Key Research Plan; Grant number: 2006CB910204; Grant sponsor: Chinese National High-Tech Research Grant; Grant number: 2006AA02A321.

*Correspondence to: Ying Xiong and Changlin Tian, School of Life Science, University of Science and Technology of China, Hefei, Anhui 230026, People's Republic of China. Email: yxiong73@ustc.edu.cn or cltian@ustc.edu.cn

Figure 1. ^{19}F chemical shift and T_1 , T_2 relaxation analysis of the Phe7-tfmF-SH3 domain in aqueous buffer. Different ^{19}F chemical shifts (A, D) and T_1 relaxation (B, E), T_2 relaxation values (C, F) are observed in the absence (A, B, C) or presence (D, E, F) of a peptide ligand P868.

evolved tRNA synthetase (RS) and linked to a tRNA containing an amber nonsense anticodon (CUA), catalyzed by the same RS. Then, tfmF can be specifically incorporated into a protein at the amber nonsense codon (TAG) in the protein's coding DNA sequence. Using this site-specific ^{19}F incorporation method, ^{19}F chemical shift analysis was shown to be possible for a large protein.⁴

Here, site-specifically incorporated tfmF and ^{19}F NMR were applied to study side chain internal mobility of large size proteins. High quality ^{19}F chemical shift, T_1 , T_2 relaxation data from samples of human vinexin SH3 domain in aqueous buffer containing 60% glycerol and a nine-transmembrane helix membrane protein diacyl-glycerol kinase (DAGK) in DPC micelles indicate that side chain internal mobility analysis can be achieved for large molecular systems.

Results and Discussion

The tfmF was specifically introduced at Phe7 site of human vinexin SH3 domain for further ^{19}F NMR studies. As the tfmF was site-specifically introduced into Phe7 of SH3, only one peak was observed and assigned to Phe7 in one-dimensional ^{19}F spectrum [Fig. 1(A,D)]. The straightforward resonance assignment makes it convenient for chemical shift and relaxation analysis. With the absence or presence of a proline-rich peptide ligand P868 (GEVPPPPRPP-PEE), which was known to bind to SH3 domain in high affinity,⁷ different ^{19}F chemical shift [Fig. 1(A,D)], T_1 [Fig. 1(B,E)], and T_2 relaxation values [Fig. 1(C,F)] were observed. Observed chemical shift changes are consistent with previous reports

that Phe7 is located in binding pocket for P868 in human vinexin SH3 domain.⁷ About a 1.5-fold increase in T_1 (1496.00 ± 286.71 vs. 1095.70 ± 68.72 ms) and a 10-fold increase in T_2 (1088.04 ± 217.47 vs. 89.24 ± 5.84 ms) were observed upon ligand binding. According to relaxation theory, ratio of T_1/T_2 is a function of internal mobility,^{8–10} and a T_1/T_2 ratio decrease or remarkable T_2 increase indicate a liberated internal motion.¹¹ It has been known that relaxations of ^{19}F spins in tfmF are dominated by chemical shift anisotropy,¹⁰ due to lack of directly bonded protons and consequent dipolar couplings. Taking advantage of fast rotational motion of CF₃- group about the C–C bond to the aromatic ring, ^{19}F T_1 value of free tfmF ($T_1 = 1.47 \pm 0.06$ s) is about 200-fold shorter than 4-F-Phe ($T_1 = 352 \pm 12$ s), in which ^{19}F spin is directly bonded to aromatic ring.¹² Therefore, application of tfmF for ^{19}F NMR chemical shift and relaxation studies is much superior to 4-F-Phe. It is also noticed that measured T_1 , T_2 relaxation values of Phe7-tfmF-SH3 at presence of ligand P868 are in similar level as T_1 , T_2 value of free tfmF ($T_1 = 1470.63 \pm 62.10$ ms, $T_2 = 969.55 \pm 199.38$ ms, data not shown). Explanations for the observations can be that previously restrained Phe7 side chain reorientation motion of human vinexin SH3 domain was liberated, when ligand P868 binds to the binding pocket.

To assess whether ^{19}F NMR side chain mobility studies can be applied for large proteins, samples of SH3 in 60% glycerol were prepared to mimic a protein with short correlation time, for example, a large size protein in aqueous buffer. A pulse field gradient spin echo NMR pulse sequence was applied to

Figure 2. Diffusion coefficient measurements of water (A), ^{19}F chemical shift (B) and T_1 relaxation (C) and T_2 relaxation analysis (D) from the sample of Phe7-tfmF-SH3 in aqueous buffer containing 60% glycerol.

measure water diffusion constant in this sample.^{13,14} In Figure 2A, diffusion coefficient of water in sample of Phe7-tfmF-SH3 in aqueous buffer containing 60% glycerol was measured ($D_s = 2.289 \times 10^{-10} \text{ m}^2/\text{s}$, Fig. 2A). According to Eq. (4) in experimental methods, a small protein (M_w') in high viscosity condition (D_s' , η') can mimic a large size protein in aqueous buffer (M_w , D_s , η). We have known molecular weight of human vinexin SH3 domain (7.4 kDa) and diffusion coefficient of water at 298 K ($2.299 \times 10^{-9} \text{ m}^2/\text{s}$).¹⁵ Therefore, a molecular weight of 7500 kDa [$(2.299/2.289)^3 \times 7.4 \text{ kDa}$] was calculated to have the same correlation time as human vinexin SH3 domain ($M_w = 7.4 \text{ kDa}$) in aqueous buffer containing 60% glycerol.

^{19}F NMR spectrum (Fig. 2B) and T_1 (Fig. 2C), T_2 relaxation data (Fig. 2D) were also collected for human vinexin SH3 domain in aqueous buffer containing 60% glycerol. Despite broad line-width and small T_2 relaxation values, decent spectra and quantitative T_1 , T_2 relaxation values were obtained even in a high viscosity. These data strongly indicate that ^{19}F chemical shift and side chain internal mobility can be analyzed for large size globular proteins with molecular weights up to 7500 kDa, in aqueous buffer.

TfmF was also incorporated at Phe31 site of a membrane protein DAGK (Phe31-tfmF-DAGK) in DPC micelles for further ^{19}F NMR studies. Trimeric DAGK contains total nine transmembrane helices and molecular weight of DAGK/DPC micelles was estimated to be around 100 kDa at 45 °C.¹⁶ ^{19}F

chemical shift, and T_1 , T_2 relaxation analysis of DAGK/DPC were shown in Figure 3. Decent ^{19}F spectrum (Fig. 3A) and quantitative T_1 (Fig. 3B), T_2 relaxation (Fig. 3C) verified that this ^{19}F NMR data can be acquired from a large membrane protein.

In a summary, ^{19}F chemical shift and relaxation analysis of SH3 in 60% glycerol indicate that this method can be applied on large size proteins. ^{19}F NMR data of DAGK/DPC verify that this site-specific side chain chemical shift and internal mobility analysis provide a powerful tool to study molecular mechanisms of large size proteins, especially membrane proteins in detergent micelles.

Materials and Methods

Protein production and ^{19}F site-specific labeling

Plasmid pBAD-His₆-SH3-TAG (a pBAD vector containing SH3 domain with its 7th codon mutated to amber stop codon TAG) and pDule-tfmF (containing DNA sequences coding tRNA_{CUA} and tfmF-specific aminoacyl-tRNA synthetase) were co-transformed into *Escherichia coli* (*E. coli*) host cell TOP10 with the presence of antibiotics. Protein over-expression was achieved with presence of 1 mM ^{19}F -tfmF in rich medium, similar as previous reports.^{4,5} Over-expressed Phe7-tfmF-SH3 was purified using Ni²⁺-NTA affinity and size-exclusive chromatography. Final yield of about 1 mg Phe7-tfmF-SH3 was obtained from every 500 mL culture.

The 31st codon (Phe) of DAGK protein was mutated to amber stop codon and Phe31-tfmF-

Figure 3. One-dimensional ¹⁹F chemical shift (A), T₁ relaxation (B) and T₂ relaxation analysis (C) of a membrane protein DAGK labelled at position 31 with tfmF in DPC micelles (Phe31-tfmF-DAGK/DPC).

DAGK was over-expressed in *E. coli* in similar way as described above. Over-expressed Phe31-tfmF-DAGK was purified in DPC micelles as described previously.¹⁶

¹⁹F NMR

All ¹⁹F NMR spectra were acquired at 298 K, in a Bruker Avance 400 MHz spectrometer. Data were processed with an exponential window function (10 Hz line-broadening) using TopSpin 3.1. ¹⁹F chemical shifts were referenced to an internal compound of trifluoro-acetic acid (TFA, -75.39 ppm).

Protein samples for ¹⁹F NMR analysis were in final concentration of 0.2 mM, including Phe7-tfmF-SH3 or Phe31-tfmF-DAGK/DPC in 500 μL aqueous

buffer of 50 mM Na₂HPO₄/NaH₂PO₄, pH 6.5. Another Phe7-tfmF-SH3 sample was prepared in the same aqueous buffer containing 60% glycerol.

Side chain ¹⁹F T₁ relaxation data were collected with eight delay times (50, 100, 200, 500, 800, 1000, 1500, 2000 ms) using a standard Bruker 1D inverse-recovery pulse sequence. ¹⁹F T₂ relaxation data were collected with 8 delay times (100, 150, 200, 400, 600, 800, 1200, and 1600 ms) using a standard Bruker 1D Carr-Purcell-Meiboom-Gill pulse sequence. Resonance intensities in relaxation experiments were measured and fit to an exponential function.

All ¹⁹F NMR spectra were acquired with 5 s acquisition delays. 64 scans were accumulated for one-dimensional ¹⁹F chemical shift analysis and 1024 scans were accumulated for T₁ and T₂ data acquisition. Therefore, totally 6 min are required for 1D NMR acquisition, while 1.5 h are required to collect each relaxation data point.

Diffusion coefficient measurement and molecular weight calculation

A pulsed field gradient-based longitudinal eddy delay pulse sequence was applied to measure water diffusion coefficient in samples of Phe7-tfmF-SH3 in aqueous buffer containing 60% glycerol. The experiments were conducted at 298 K, in a 500 MHz Bruker Avance spectrometer. A series of 1D data were collected with different delay duration of 13.184, 52.737, 118.66, 210.95, 329.61, 474.64, 646.03, and 843.80 ms. Data were processed using exponential window function and intensities of H₂O signal (4.680 ppm) were measured using Bruker xwinnmr software. Diffusion constant was regressed from intensity decay with different delay time.

The Stokes-Einstein Equation was applied to derive the hydrated radius or the molecular weight of a globular protein:

$$D_s = kT / (6\pi\eta R_s) \quad (1)$$

η – viscosity of the solution R_s – sphere radius of the molecule

$$\text{At a constant temperature: } D'_s \eta' R'_s = D_s \eta R_s \quad (2)$$

Assuming a constant partial specific volume V, protein's molecular weight (M_w) is proportional to cubic of hydrated radius R_s³. Replacing hydrated radius R_s with molecular weight M_w, we can reach another equation:

$$(D'_s \eta')^3 M_w' = (D_s \eta)^3 M_w \quad (3)$$

Therefore, assuming the same correlation time at 298 K, molecular weight of a large size protein in aqueous buffer (M_w, D_sη) can be a mimic of a small protein (M_w') in high viscosity condition (D_s' η')

$$M_w = M'_w (D'_s \eta' / D_s \eta)^3 \quad (4)$$

D'_s and D_s are the water diffusion coefficients in aqueous sample and an aqueous buffer in high viscosity, respectively.

Acknowledgment

Authors are grateful for kind courtesy of plasmid pDule-tfmF from Dr. R.A. Mehl, Department of Chemistry, Franklin and Marshall College, Lancaster, Pennsylvania, USA.

References

1. Lipchick JM, Loria JP (2009) Monitoring molecular interactions by NMR. *Methods Mol Biol* 490: 115–134.
2. Haggie PM, Brindle KM (1999) Mitochondrial citrate synthase is immobilized in vivo. *J Biol Chem* 274: 3941–3945.
3. Williams SP, Haggie PM, Brindle KM (1997) ^{19}F NMR measurements of the rotational mobility of proteins in vivo. *Biophys J* 72: 490–498.
4. Jackson JC, Hammill JT, Mehl RA (2007) Site-specific incorporation of a ^{19}F -amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J Am Chem Soc* 129: 1160–1166.
5. Hammill JT, Miyake-Stoner S, Hazen JL, Jackson JC, Mehl RA (2007) Preparation of site-specifically labeled fluorinated proteins for ^{19}F -NMR structural characterization. *Nat Protoc* 2: 2601–2607.
6. Li C, Wang GF, Wang Y, Creager-Allen R, Lutz EA, Scronce H, Slade KM, Ruf RA, Mehl RA, Pielak GJ (2010) Protein ^{19}F NMR in *Escherichia coli*. *J Am Chem Soc* 132: 321–327.
7. Zhang J, Li X, Yao B, Shen W, Sun H, Xu C, Wu J, Shi Y (2007) Solution structure of the first SH3 domain of human vinexin and its interaction with vinculin peptides. *Biochem Biophys Res Commun* 357: 931–937.
8. Lipari G, Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. II. Analysis of experimental results. *J Am Chem Soc* 104: 4559–4570.
9. Lipari G, Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. I. Theory and range of validity. *J Am Chem Soc* 104: 4546–4559.
10. Grage SL, Durr UH, Afonin S, Mikhailiuk PK, Komarov IV, Ulrich AS (2008) Solid state ^{19}F NMR parameters of fluorine-labeled amino acids. Part II: aliphatic substituents. *J Magn Reson* 191: 16–23.
11. Palmer AG 3rd (1993) Dynamic properties of proteins from NMR spectroscopy. *Curr Opin Biotechnol* 4: 385–391.
12. Durr UH, Grage SL, Witter R, Ulrich AS (2008) Solid state ^{19}F NMR parameters of fluorine-labeled amino acids. I. Aromatic substituents. *J Magn Reson* 191: 7–15.
13. Chou JJ, Baber JL, Bax A (2004) Characterization of phospholipid mixed micelles by translational diffusion. *J Biomol NMR* 29: 299–308.
14. Wu DH, Chen AD, Johnson CS (1995) An improved diffusion-ordered spectroscopy experiment incorporating bipolar-gradient pulses. *J Magn Reson A* 115: 260–264.
15. Hota M, Weingartner H (1991) Calibration in accurate spin-echo self-diffusion measurements using ^1H and less-common nuclei. *J Magn Reson* 92: 115.
16. Oxenoid K, Kim HJ, Jacob J, Sonnichsen FD, Sanders CR (2004) NMR assignments for a helical 40 kDa membrane protein. *J Am Chem Soc* 126: 5048–5049.