Site-specific protein backbone and side-chain NMR chemical shift and relaxation analysis of human vinexin SH3 domain using a genetically encoded $^{15}$N/$^{19}$F-labeled unnatural amino acid

Pan Shi$^a$$^b$$^c$$^1$, Zhaoyong Xi$^c$$^1$, Hu Wang$^c$, Chaowei Shi$^a$$^b$, Ying Xiong$^b$$^*$$^c$, Changlin Tian$^a$$^*$$^b$

$^a$National Laboratory for Physical Science at Microscale, University of Science and Technology of China, Hefei, Anhui 230026, PR China
$^b$School of Life Science, University of Science and Technology of China, Hefei, Anhui 230026, PR China
$^c$School of Chemistry, University of Science and Technology of China, Hefei, Anhui 230026, PR China

A R T I C L E   I N F O

Article history:
Received 7 October 2010
Available online 12 October 2010

Keywords:
Unnatural amino acids
Trifluoromethyl-phenylalanine
Site-specific labeling
Solution NMR
Relaxation analysis
Ligand binding

A B S T R A C T

SH3 is a ubiquitous domain mediating protein–protein interactions. Recent solution NMR structural studies have shown that a proline-rich peptide is capable of binding to the human vinexin SH3 domain. Here, an orthogonal amber tRNA/tRNA synthetase pair for $^{15}$N/$^{19}$F-trifluoromethyl-phenylalanine ($^{15}$N/$^{19}$F-tfmF) has been applied to achieve site-specific labeling of SH3 at three different sites. One-dimensional solution NMR spectra of backbone amide ($^{15}$N)$^{1}$H and side-chain $^{19}$F were obtained for SH3 with three different site-specific labels. Site-specific backbone amide ($^{15}$N)$^{1}$H and side-chain $^{19}$F chemical shift and relaxation analysis of SH3 in the absence or presence of a peptide ligand demonstrated different internal motions upon ligand binding at the three different sites. This site-specific NMR analysis might be very useful for studying large-sized proteins or protein complexes.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Proteins are intrinsically flexible and dynamic systems. These characteristics are critical in allowing proteins to perform various biological functions. Nuclear magnetic resonance (NMR) has been proven to be a unique tool in allowing a detailed, site-specific description of internal motions in proteins. A variety of NMR pulse sequences have been developed to study the internal mobility of the backbone and side chains of a protein [1]. Most protein backbone internal motion analyses have focused on longitudinal (T1) and transverse (T2) relaxation of backbone amide $^{15}$N and dipolar coupling of the $^{15}$N–$^{1}$H bond, while protein side-chain internal motion has been analyzed through relaxation measurements of the nuclei $^{13}$C or $^{2}$H. However, analysis of protein relaxation data for either backbone or side-chain nuclei is greatly hampered by difficulties in peak assignment and rapid signal decay due to slow rotational motion of large-sized proteins.

Unnatural amino acids were first incorporated into a protein using an orthogonal amber tRNA/tRNA synthetase (tRNA/RS) pair about a decade ago [2–4]. In this method, unnatural amino acids can be specifically incorporated into a protein at the amber nonsense codon (TAG) in the protein’s coding DNA sequence [4], which provides a good mechanism for introducing an NMR-active spin at a designated site in a protein. Several different stable-isotope-labeled unnatural amino acids have been incorporated into proteins for chemical shift analysis in 1D and 2D NMR spectra [5–7]. However, to the best of our knowledge, there have not hitherto been any reports on relaxation analysis of proteins with site-specifically incorporated unnatural amino acids with active NMR spins, in particular, with both backbone and side-chain double labeling.

SH3 is a ubiquitous domain mediating protein–protein interactions. The SH3 domain of human vinexin has been extensively studied recently, and it was shown to be bound by a proline-rich peptide (GEVPPPRPPPPEE, P868) at several specific sites with high affinity [8]. Previous data have shown that residues Phe7 and Tyr51 are located in the SH3 binding pocket for peptide ligand P868, while residue Tyr26 is located away from the binding pocket [8]. In the work described herein, a $^{15}$N/$^{19}$F-labeled unnatural amino acid, trifluoromethyl-phenylalanine ($^{15}$N/$^{19}$F-tfmF), has been incorporated at several aromatic residue sites, namely Phe7, Tyr26, or Tyr51. Chemical shift and T1, T2 relaxation analyses of backbone amide ($^{15}$N)$^{1}$H or side-chain $^{19}$F spins at these three sites of SH3 have been conducted in the presence or absence of peptide ligand P868.

2. Materials and methods

2.1. Synthesis of the $^{15}$N-labeled trifluoromethyl-phenylalanine

To synthesize $^{15}$N-labeled trifluoromethyl-phenylalanine, $^{15}$N-labeled diethyl acetaminomalonate was first synthesized as
an intermediate. A mixture of Na$^{15}$NO$_2$ (8.75 g), diethyl malonate (16 mL), and water (0.75 mL) was placed in a three-necked, round-bottomed flask, with toluene (24 mL) as solvent. The stirred mixture was cooled in an ice bath. Glacial acetic acid (10 mL) was then slowly added. The reaction mixture was warmed to 40–50 °C and maintained at this temperature for 6 h. The organic layer was then separated using a separating funnel. An yellow oil (diethyl oximino-malonate) was obtained after removing the solvent. Reaction of diethyl oximino-malonate with glacial acetic acid was initiated by the addition of zinc powder (22 g) and the stirred mixture was kept at 40–50 °C for 4 h.

A mixture of Na$^{15}$NO$_2$ (8.75 g), diethyl malonate (11.0 g) and water (0.75 mL) was placed in a three-necked, round-bottom flask, with toluene (24 mL) as solvent. The stirred mixture was warmed to 50 °C for 1.5 h.

4-Trifluoromethylbenzyl bromide. A solution of diethyl acetamidomalonate (10.0 g) in anhydrous ethanol (125 mL) was added to a sodium ethoxide solution and the mixture was heated at 50 °C for 24 h before the final product of $^{15}$N-labeled trifluoromethyl-phenylalanine (tfmF) was dried. Details of verification of the identity of the product and assessment of its purity can be found in the Supporting Information.

2.2. Constructs

A DNA fragment encoding the human vinexin SH3 domain was PCR-amplified from the previously described expression plasmid [8]. The amplified fragment was inserted into pBAD vector (Invitrogen Co.) between the Ncol and Xhol sites. The protein sequence derived from the plasmid contained a His6-tag attached to the N terminus of the recombinant SH3 domain.

2.3. Incorporation of $^{15}$N/$^{19}$F-labeled trifluoromethyl-phenylalanine into the SH3 domain of human vinexin protein

In plasmid pBAD-His6–SH3, codons corresponding to Phe7, Tyr26, and Tyr51 of the SH3 domain were respectively, mutated to amber stop codon TAG using quick-change site-directed mutagenesis (Stratagene Co.). The pBAD plasmid coding the mutated SH3 domain and plasmid pDule-tfmF (containing DNA sequences coding tRNA$_{\text{UA}}$ and tfmF-specific aminocyl-tRNA synthetase, kindly provided by Dr. R.A. Mehli, Department of Chemistry, Franklin and Marshall College, Pennsylvania, USA) were co-transformed into Escherichia coli host cells TOP10 in the presence of 15 μg/mL tetracycline and 100 μg/mL ampicillin, in a similar manner as reported previously [7,9]. The transformed bacteria were incubated in LB medium overnight, then transferred to fresh 2 × YT medium containing 1 mM $^{15}$N/$^{19}$F-tfmF at 37 °C. Expression of the SH3 protein was induced using 0.2% arabinose when the OD$_{600}$ reached 1.0.

2.4. Protein purification

Cells were harvested by centrifugation at 6000 rpm for 8 min at 4 °C. Cell pastes were suspended in 40 mL of lysis buffer [50 mM Tris–HCl, 500 mM NaCl, 3 mM imidazole, pH 8.0]. Cell suspensions in lysis buffer were probe-sonicated (VC500, Sonics and Materials, Danbury, CT) at a power level of 30%, 2.0 s pulse on and 4.0 s pulse off, for a total of 10 min on ice. The lysate was then centrifuged at 16,000 rpm for 20 min at 4 °C. Pellets from the centrifugation were discarded and the supernatant was mixed with 5 mL of Ni$^{2+}$-NTA resin (Qiagen, Valencia, CA). The mixture was rotated at 4 °C for 20 min before packing onto a gravity-flow column. Impurities were washed out by applying 50 mL of washing buffer [50 mM Tris, 500 mM NaCl, 20 mM imidazole, pH 8.0] to the column. The target proteins were then eluted from the column using elution buffer [50 mM Tris, 500 mM NaCl, 250 mM imidazole, pH 8.0]. Purified recombinant SH3 proteins were analyzed using SDS–PAGE (12%, w/v) and concentrated using an Amicon Ultra 15 mL device (5000 MWCO, Millipore). Two rounds of concentration and dilution in NMR buffer [50 mM NaH$_2$PO$_4$–Na$_2$HPO$_4$, pH 6.5] were applied to the eluted protein sample for imidazole removal and buffer exchange. D$_2$O was added to the SH3 sample to give a 10% (v/v) final concentration before solution NMR analysis.

2.5. Backbone amide $^1$H and side-chain $^{19}$F chemical shift and relaxation data analysis

All one-dimensional $^{15}$N-filtered $^1$H NMR spectra were acquired at 298 K on a Bruker Avance 500 MHz spectrometer equipped with a triple-resonance cryo-probe. A heteronuclear single-quantum correlation (HSQC) pulse sequence was applied to implement $^{15}$N magnetization filtering. The first free-induction decay (FID) data, with 2048 complex points for the HSQC experiment, were collected and processed with an exponential window function (line broadening = 10 Hz) using Bruker data-processing software. Backbone amide $^{15}$N T$_1$ and T$_2$ relaxation data were acquired using a 1D mode standard HSQC-based pulse sequence on the same Bruker 500 MHz spectrometer [10]. Peak intensities of a total of seven longitudinal relaxation durations (61.33, 141.6, 242.0, 362.42, 523.0, 753.82 and 1145.2 ms) were measured and regressed against the durations as a single-exponential function for the T$_1$ data, while peak intensities of six transverse relaxation durations (17.6, 35.2, 52.8, 70.4, 105.6 and 140.8 ms) were measured and regressed for the T$_2$ data of backbone amide $^1$H in tfmF–SH3 using Origin software (OriginLab, Co.).

All one-dimensional $^{19}$F NMR spectra were acquired at 298 K on a Bruker Avance 400 MHz spectrometer equipped with a broadband double-resonance probe and the observation channel was tuned to $^{19}$F (376 MHz). One-dimensional $^{19}$F chemical shift data were acquired with 16384 complex points using a simple 1D pulse, and processed with an exponential window function (line broadening = 10 Hz) using Bruker data-processing software. Side-chain $^{19}$F T$_1$ relaxation data were collected with eight longitudinal relaxation durations (50, 100, 200, 500, 800, 1000, 1500 and 2000 ms) using a standard Bruker one-dimensional inverse-recovery pulse sequence. Side-chain $^{19}$F T$_2$ relaxation data were collected with eight transverse relaxation durations (100, 150, 200, 400, 600, 800, 1200 and 1600 ms) using a standard Bruker 1D Carr–Purcell–Meiboom–Gill (CPMG) pulse sequence. The intensities of the series of peaks were measured and regressed for T$_1$ or T$_2$ data for the side-chain $^{19}$F in $^{15}$N/$^{19}$F-tfmF–SH3.

3. Results and discussion

3.1. Unnatural amino acid tfmF synthesis

The reactions shown in Scheme 1 were carried out to incorporate $^{19}$F into the unnatural amino acids. $^{15}$N spin was introduced at the amide nitrogen of $^{19}$F-trifluoromethyl-phenylalanine ($^{19}$F-tfmF) from $^{15}$N-labeled NaNO$_2$, which provided the original source of $^{15}$N spin (Scheme 1). Before synthesizing $^{15}$N/1$^{19}$F-tfmF, $^{15}$N-labeled diethyl acetamidomalonate (DEAM), an amino acid intermediate, was synthesized (Scheme 1, reaction 1). This $^{15}$N-amide-labeled DEAM could be used as a common amino acid intermediate for the further synthesis of many different natural or unnatural amino acids. Here, 4-trifluoromethylbenzyl bromide...
was added to pre-activated $^{15}\text{N}$-labeled diethyl acetamidomalonate. Consequently, trifluoromethyl-phenylalanine was synthesized with $^{15}\text{N}$ labeling at its backbone amide nitrogen and a trifluoromethyl group attached to the phenylalanine side chain (Scheme 1, reaction 2). Since there is only one neutron difference between $^{15}\text{N}$ and the naturally abundant $^{14}\text{N}$, this minor difference will not affect the size or chemical properties of the unnatural amino acid tfmF. Therefore, tfmF-specific aminoacyl-tRNA synthetase can recognize $^{15}\text{N}$-labeled tfmF and catalyze charging the unnatural amino acid to a specific tRNACUA. Any protein with site-specific $^{15}\text{N}/^{19}\text{F}$-labeled tfmF incorporation will contain both $^{15}\text{N}$ spin in its backbone and $^{19}\text{F}$ spin in a side chain at this specific site.

3.2. Protein expression and purification

Human vinexin SH3 domain containing site-specific $^{15}\text{N}/^{19}\text{F}$-labeled tfmF was over-expressed from a double-plasmid system. Site-specific $^{15}\text{N}/^{19}\text{F}$-tfmF incorporation was accomplished under two specificities: (1) specific recognition of tfmF by tfmF–tRNA synthetase and charging the unnatural amino acid to a specific tRNACUA; (2) complementation between the amber stop codon and the tfmF–tRNA anti-codon sequence. Due to the presence of the stop codon in the SH3 sequence in pBAD plasmid, the SH3 expression yield was not as high as that of the wild-type protein under the same expression conditions; the yield of purified SH3 protein was about 2 mg/L of culture. SDS–PAGE analysis of the purified protein after Ni–NTA affinity chromatography indicated a high purity for further solution NMR studies (data not shown).

3.3. Different chemical shift changes at different sites in the one-dimensional NMR spectra

Since $^{15}\text{N}/^{19}\text{F}$-tfmF was site-specifically introduced at three different sites, only one peak was observed in each $^{15}\text{N}$-filtered $^1\text{H}$ spectrum and $^{19}\text{F}$ spectrum (Figs. 1 and 2). Therefore, resonance assignment is straightforward. In the HSQC-based $^{15}\text{N}$-filtered $^1\text{H}$ one-dimensional spectra (Fig. 1) and directly observed $^{19}\text{F}$ one-dimensional spectra (Fig. 2), the single peaks can be assigned to the $^{15}\text{N}/^{19}\text{F}$-tfmF substituted residue at the respective sites: the Phe7 site in Figs. 1A, B, 2C and D; the Tyr26 site in Figs. 1C, D, 2E and F; and the Tyr51 site in Figs. 1E, F, 2G and H. At the same time, in both the amide $^1\text{H}$ and $^{19}\text{F}$ spectra, pronounced chemical shift changes were observed for $^{15}\text{N}/^{19}\text{F}$-tfmF incorporation at the Phe7 (Figs. 1A, B, 2C and D) and Tyr51 sites (Figs. 1E, F, 2G and H) with the absence or presence of peptide ligand P868, while no chemical shift changes were observed for compound tfmF alone (Fig. 2A and B) or for $^{15}\text{N}/^{19}\text{F}$-tfmF incorporation at the Tyr26 site (Fig. 2E and F). The observed chemical shift changes upon ligand binding are consistent with previous reports that Phe7 and Tyr51 are located in the ligand-binding pocket of SH3, while Tyr26 is located away from the binding pocket [8]. No $^{19}\text{F}$ chemical shift changes were observed for the $^{15}\text{N}/^{19}\text{F}$-tfmF compound in the absence or presence of the peptide ligand, which excluded any non-specific interaction between the peptide and the labeled residues in proteins.

3.4. Different changes of internal motions at different sites upon peptide ligand binding

Fig. 3 shows backbone amide $^{15}\text{N}$ and side-chain $^{19}\text{F}$ longitudinal $T_1$ and transverse $T_2$ relaxation analyses of $^{15}\text{N}/^{19}\text{F}$-tfmF7–SH3 ($^{15}\text{N}/^{19}\text{F}$-tfmF incorporation at the Phe7 site). Upon ligand binding, backbone amide $^{15}\text{N}$ $T_1$ relaxation times at the Phe7 site were

\[ \text{Reaction 1:} \quad \left\{ \begin{array}{c} \text{CH}_3(\text{CO})_2\text{C(OH)}_2 \quad + \quad \text{NaNO}_2 \quad + \quad \text{CH}_3\text{COOH} \\ \text{ZnCl}_2(\text{CO})_2\text{O} \quad + \quad \text{CH}_3\text{COOH} \quad \rightarrow \quad \text{HON=CH(CO)CH}_3 \end{array} \right. \]

\[ \text{Reaction 2:} \quad \left\{ \begin{array}{c} \text{CF}_3 \quad \text{CH}_2\text{Br} \quad + \quad \text{CHCOOEt}_2 \quad \rightarrow \quad \text{CF}_3 \quad \text{CH}_2\text{COOEt}_2 \quad + \quad \text{NAHCO}_3 \quad \rightarrow \quad \text{CF}_3 \quad \text{CH}_2\text{COOEt}_2 \\ \text{HOCO}_2 \quad \text{COOH} \quad \rightarrow \quad \text{CF}_3 \quad \text{CH}_2\text{COOEt}_2 \quad + \quad \text{HOCO}_2 \quad \text{COOH} \end{array} \right. \]
observed to be halved (Fig. 3B cf. D), while no pronounced differences in T1 were observed (Fig. 3A cf. C). At the same time, T2 relaxation times of side-chain 19F at the Phe7 site increased by a factor of 10 (Fig. 3F cf. H), while the T1 relaxation time was observed to increase 1.5-fold (Fig. 3E and G) upon ligand binding.

According to relaxation theory and the spectral density function of 15N and 19F spins, the T1/T2 ratio is a function of internal mobility [11–13]. An increased T1/T2 ratio or pronounced decrease in T2 indicates a restrained internal motion at a specific site, while a decreased T1/T2 ratio or pronounced increase in T2 indicates a liberated internal motion [14]. Therefore, the markedly reduced backbone amide 15N T2 time and increased side-chain 19F T2 time at the Phe7 site indicated that backbone internal motion at the Phe7 site was restrained, while side-chain internal motion at the Phe7 site was liberated, upon binding of the peptide ligand P868 to SH3.

Besides relaxation analysis at the Phe7 site of SH3, site-specific backbone and side-chain T1, T2 relaxation data at the Tyr51 and Tyr26 sites were also collected, and are summarized in Table 1. No pronounced backbone or side-chain T1, T2 relaxation changes were observed at the Tyr26 site, which is consistent with previous reports that Tyr26 is located away from the P868 binding pocket. Relaxation analysis at the Tyr51 site showed a decreased T2 relaxation time and an increased backbone T1/T2 ratio in both the backbone and side-chain, indicating that internal motions in both the backbone and side-chain regions at the Tyr51 site were restrained after peptide ligand binding to human vinexin SH3, which is in marked contrast to the internal motion changes at the Phe7 site.

3.5. Site-specific isotope labeling and relaxation analysis for other proteins

Here, we have achieved site-specific labeling and relaxation analysis of several sites in human vinexin SH3, in the absence or presence of a peptide ligand. The chemical shift data and backbone
Table 1

<table>
<thead>
<tr>
<th>Backbone 15N T1 (ms)</th>
<th>With P668</th>
<th>w/o P668</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe7</td>
<td>6.93</td>
<td>7.03</td>
</tr>
<tr>
<td>Tyr51</td>
<td>7.27</td>
<td>7.68</td>
</tr>
<tr>
<td>Tyr26</td>
<td>8.42</td>
<td>8.42</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Side-chain 19F T1 (ms)</th>
<th>With P668</th>
<th>w/o P668</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe7</td>
<td>1470.6 ± 62.1</td>
<td>n.d.</td>
</tr>
<tr>
<td>Tyr51</td>
<td>1095.7 ± 68.7</td>
<td>1496.0 ± 286.7</td>
</tr>
<tr>
<td>Tyr26</td>
<td>869.98 ± 20.46</td>
<td>910.40 ± 37.44</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Backbone 15N T2 (ms)</th>
<th>With P668</th>
<th>w/o P668</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe7</td>
<td>62.21</td>
<td>n.d.</td>
</tr>
<tr>
<td>Tyr51</td>
<td>61.74</td>
<td>61.94</td>
</tr>
<tr>
<td>Tyr26</td>
<td>62.16</td>
<td>61.96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Side-chain 19F T2 (ms)</th>
<th>With P668</th>
<th>w/o P668</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe7</td>
<td>969.5 ± 199.4</td>
<td>n.d.</td>
</tr>
<tr>
<td>Tyr51</td>
<td>1088.04 ± 217.47</td>
<td>1086.04 ± 217.47</td>
</tr>
<tr>
<td>Tyr26</td>
<td>78.11 ± 3.75</td>
<td>85.47 ± 4.29</td>
</tr>
</tbody>
</table>

and side-chain relaxation times in Table 1 proved to be consistent with the previously reported ligand binding pocket of SH3. Different internal motions were also indicated in the backbone and side chain at the different sites upon ligand binding. These results strongly indicate that the method of site-specific isotope spin incorporation is very useful and reliable for analyzing conformational changes and internal motions at biologically interesting sites of complicated proteins.

Biologically interesting sites of a protein can include active sites of enzymes, ligand binding sites of receptors, the pore-forming region of a channel, or the interface region in protein–protein complexes. To reveal the molecular mechanisms of a complicated protein, analogues of the protein with a series of site mutations (even mutation scanning) or with unnatural amino acid incorporations can be prepared for site-specific dynamics analysis. Since 15N/19F-tfmF is a derivative of phenylalanine, only aromatic residue sites are considered safe for 15N/19F-tfmF incorporation. This does not exclude 15N/19F-tfmF incorporation at other hydrophobic residue sites (e.g., Leu, Ile, Val), provided that residue replacement does not perturb the structure or function of the target protein. To date, there have been reports of other 13C- or 15N-labeled unnatural amino acid incorporations for protein NMR analysis [5,15], but to the best of our knowledge only 15N/19F-tfmF contains two different spins in both the backbone and side chain for solution NMR chemical shift and relaxation analysis.

Normally, to conduct chemical shift and relaxation analyses of a large-sized protein or protein complex, it is prerequisite to have resonance assignments of the protein at hand. Here, single-peak spectra have been ensured by using site-specific labeling, thereby providing a straightforward peak assignment, which is very helpful in that it avoids complicated, time-consuming resonance assignment of a large protein system. At the same time, if we know that there is only one peak, there will be no need to perform too many data accumulations to acquire an NMR spectrum, simply because we can identify a single peak even at a relatively low signal/noise ratio. This is also helpful for speeding up site-specific chemical shift and relaxation analyses of a protein using solution NMR.

In summary, site-specific 15N/19F-labeling has been achieved through genetically incorporating the unnatural amino acid 15N/19F-tfmF in the human vinexin SH3 domain. Site-specific backbone and side-chain chemical shift and relaxation analyses of SH3 have been conducted at several different sites. The observed NMR results are consistent with the previously reported ligand-binding pocket of SH3. Site-specific relaxation analysis data of SH3 indicate different internal motions at different sites upon ligand binding. This method provides a site-by-site mechanism to analyze detailed conformational changes or internal mobility in the backbone or side chain of a complicated protein system in different functional states.

Acknowledgments

The authors are grateful for the kind courtesy of provision of plasmid pDule-tfmF by Dr. R.A. Mehl, Department of Chemistry, Franklin and Marshall College, Lancaster, PA, USA. This work was supported by the Chinese Key Research Plan (2006CB910204) and a Chinese National High-Tech Research Grant (2006AA02A321) to C.L. Tian.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2010.10.046.

References

