Protein Expression and Purification 88 (2013) 196–200

Contents lists available at SciVerse ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Bacterial expression and purification of the amyloidogenic peptide PAPf39 for multidimensional NMR spectroscopy

Aranganathan Shanmuganathan, Anthony C. Bishop, Kinsley C. French, Scott A. McCallum, George I. Makhatadze *

Center for Biotechnology and Interdisciplinary Studies and Department of Biology, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180, USA

ARTICLE INFO

Article history:
Received 30 October 2012
and in revised form 18 December 2012
Available online 11 January 2013

Keywords:
Aggregation-prone
HIV infection
Escherichia coli expression
Atomic force microscopy
Nuclear magnetic resonance spectroscopy
Isotopic labeling

ABSTRACT

PAPf39 is a 39 residue peptide fragment from human prostatic acidic phosphatase that forms amyloid fibrils in semen. These fibrils have been implicated in facilitating HIV transmission. To enable structural studies of PAPf39 by NMR spectroscopy, efficient methods allowing the production of milligram quantities of isotopically labeled peptide are essential. Here, we report the high-yield expression and purification of uniformly ^{13}C- and ^{15}N-labeled PAPf39 peptide, through expression as a fusion to ubiquitin at the N-terminus and an intein at the C-terminus. This allows the study of the PAPf39 monomer conformational ensemble by NMR spectroscopy. To this end, we performed the NMR chemical shift assignment of the PAPf39 peptide in the monomeric state at low pH.

Introduction

Amyloid fibril formation is a hallmark of a number of human diseases. PAPf39,[1] or Prostatic Acidic Phosphatase fragment of 39 residues (corresponding to residues 248–286 in human prostatic acidic phosphatase), is a cationic peptide that forms amyloid fibrils in semen (SEVI) which increase HIV infectivity by up to five orders of magnitude [1]. Understanding the mechanism of PAPf39 fibril formation may provide insights into HIV transmission via semen and lay a foundation for the development of therapeutics against this effect. This approach requires structural information for both the monomeric and fibrillar states of PAPf39 [2,3].

Nuclear magnetic resonance (NMR) spectroscopy is a powerful method for determining the structure and dynamics of biomacromolecules [4–6]. However, modern protein NMR relies on the properties of stable isotopes of carbon (^{13}C) and nitrogen (^{15}N). Current methods of peptide synthesis allow incorporation of isotope enriched amino acids, but synthesis is rather expensive, especially for larger peptides and certain amino acid types. In contrast, isotopic enrichment by bacterial expression in defined media using isotopically labeled metabolites such as ^{13}C glucose and ^{15}N ammonium chloride as the sole source of carbon and nitrogen has been proven as a cost effective strategy to generate milligram quantities of isotopically labeled proteins [4–8]. However, peptides are notoriously difficult to express in bacterial expression systems due to their rapid degradation. Peptides that have an intrinsic propensity to aggregate and, in particular, form amyloid fibrils create additional problems for recombinant expression. To overcome this challenge, various fusion constructs with larger proteins that act as solubility tags and/or protect the peptides from degradation have been used. We have previously expressed soluble peptides as ubiquitin fusion proteins [9–11]. However, when this expression system was used for the amyloidogenic peptide, PAPf39, we observed a significant level of degradation of the peptide from the C-terminus. To alleviate these substantial losses, we attached a self-cleaving intein as a C-terminal fusion. This allowed the high yield expression and easy purification of the PAPf39 peptide free of non-native sequence insertions. The fibril formation of recombinant PAPf39 was characterized by Thioflavin T (ThT) fluorescence assays and atomic force microscopy (AFM) imaging, and NMR chemical shift assignment was performed for the PAPf39 monomer in solution.

Materials and methods

Chemicals and plasmids

All reagents were purchased from Sigma–Aldrich and Fisher Scientific unless stated otherwise. The plasmid pTXB1, restriction

* Corresponding author. Fax: +1 518 279 2955.
E-mail address: makhag@rpi.edu (G.I. Makhatadze).

[1] *Abbreviations used:* PAPf39, a 39 residue peptide fragment from human prostatic acidic phosphatase; GFP, green fluorescent protein; MBP, maltose binding protein; UBQ, ubiquitin; TEV, tobacco etch virus; PAGE, polyacrylamide gel electrophoresis; SEC, size exclusion chromatography; PBS, phosphate buffered saline; Ni–NTA, nickel–nitrilotriacetic acid; TFA, trifluoroacetic acid; 6H, hexahistidine tag.

1046-5928/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.pep.2013.01.003

enzymes, and *Escherichia coli* expression host BL21(DE3) were purchased from New England Biolabs (Ipswich, MA). DNA purification kits and Ni-NTA resin were obtained from Qiagen (Germantown, MD). DNA oligonucleotides were purchased from Integrated DNA technologies (Coralville, IA). ^{15}N ammonium chloride and ^{13}C glucose were purchased from Cambridge Isotope Laboratories (Andover, MA). SDS–PAGE analysis was done using precast Thermo Scientific Precise 4–20% gels in a Tris-HEPES buffer system, which allows all fragments to be resolved on the same gel.

Design of expression construct

Expressing small peptides in *E. coli* is commonly a challenging process because the peptides are susceptible to *E. coli* proteases and/or are toxic to the expression host system. The use of fusion proteins may limit these problems. Our first attempt to express the PAPf39 peptide with ubiquitin attached to the N-terminus of the peptide resulted in C-terminal truncated peptides. Similar effects were observed for other fusion protein constructs such as GFP and MBP. To prevent C-terminal degradation, we created a construct with self-cleavable tag (intein) attached to the C-terminus of the peptide using a pTXB1 vector [12]. The vector diagram of the protein construct is shown in Fig. 1.

The protein construct was designed to have a N-terminal hexahistidine (6H) tag for purification, followed by the ubiquitin (UBQ) sequence for solubility, a TEV protease cleavage site to allow removal of the 6H-UBQ tag [9–11], the PAPf39 sequence, and an intein tag to stop C-terminal degradation of PAPf39. The construct was inserted in a pTXB1 vector using the Xbal and Sap1 restriction enzyme sites. The gene sequence of 6H_UBQ_TEV_PAPf39 (6H-UBQ attached to PAPf39 with a TEV protease cleavage site in between) was amplified from the pGia vector by PCR with the following primers: forward GGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGG and reverse GGTGGTTGCTCTTCCGCAGTACATAATTAATTTTTTATAGGA TGG. The amplified PCR product, and pTXB1 vector were digested with Xba1 and Sap1, and gel purified.

The digested vector was treated with SAP (Shrimp Alkaline Phosphatase) overnight, which was then inactivated by heating at 65 °C. The SAP treated vector and PCR product were ligated using T4 ligase overnight at 4 °C. The ligated products were transformed into DH5α competent cells using electroporation. The transformed cells were plated on LB agar plates with 100 μg/ml of ampicillin and incubated overnight at 37 °C. Individual colonies were picked and grown overnight in LB medium with ampicillin. Plasmid was isolated from the cells followed by sequence verification. The sequence of the verified plasmid used for expression is shown in Fig. 1B, where residues 1–83 correspond to the N-terminal 6H-tagged ubiquitin, residues 84–90 comprise the TEV protease recognition motif (cleavage between Q^{89} and G^{90}), residues 90–128 correspond to the PAPf39 peptide, and residues 129–388 correspond to the intein sequence followed by a chitin binding domain as encoded by the pTXB1 vector.

Recombinant expression of uniformly ^{15}N- and ^{13}C-enriched PAPf39

The pTXB1-6H_UBQ_TEV_PAPf39_INTEIN expression construct was transformed into *E. coli* BL21(DE3) and plated on LB agar plates containing 100 μg/ml of ampicillin. A single colony was picked and grown in 2 ml of LB medium to make a starter culture. From the starter culture, a 100 ml overnight culture (Neidhardt medium [13]) was inoculated. For uniform labeling of ^{15}N and/or ^{13}C, the bacteria were grown at 37 °C in Neidhardt medium containing 0.5 g/liter of $^{15}NH_4Cl$ and 2 g/liter of ^{13}C glucose as the sole nitrogen and carbon sources [13]. When the cell density reached ~0.8 O.D. at 600 nm, expression of the 6H_UBQ_TEV_PAPf39_INTEIN fusion protein was induced by adding 1 mM isopropyl-D-thiogalactopyranoside and incubating at 25 °C for 6 h. Cells were harvested and lysed using a French pressure cell. Soluble 6H_UBQ_TEV_PAPf39_INTEIN fusion protein in the cell extract was purified using Ni-NTA affinity resin under native conditions and eluted in 250 mM imidazole buffer following the manufacturer's protocol.

To remove the intein tag, the pH of the protein solution was adjusted to pH 8.5 using 1 M Tris (pH 8.5) followed by the addition of 125 mM dithiothreitol (DTT). The reaction was incubated at 4 °C for 12 to 24 h and the degree of intein cleavage was assessed by SDS–PAGE. Once cleavage was completed, the DTT was removed by overnight dialysis against four liters of modified TEV protease buffer (50 mM TRIS, 5% glycerol, 300 mM sodium chloride, pH 8 with 0.5 mM EDTA). To remove the 6H-UBQ tag, 1 mg of TEV protease was added for every 40 mg of substrate and the reaction mixture was incubated at 4 °C for up to 24 h. Cleavage of the 6H-UBQ tag was assessed by SDS–PAGE. After TEV cleavage, the pool containing PAPf39, intein, and ubiquitin was titrated with concentrated HCl

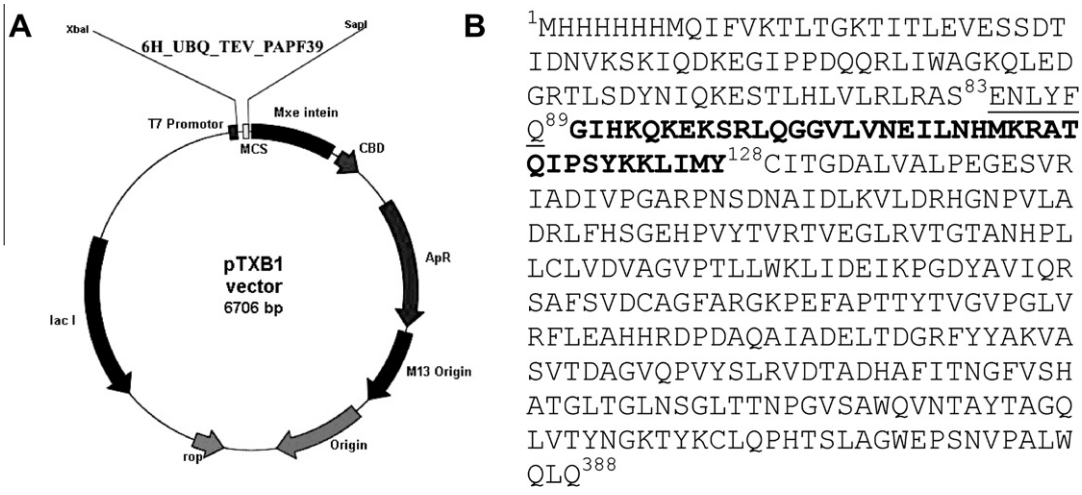

Fig. 1. Expression construct for the PAf39 peptide. Panel A. The hexahistidine tagged-ubiquitin gene with a TEV protease cleavage site followed by the PAPf39 coding sequence was inserted into a pTXB1 construct to yield the pTXB1::6H_UBQ_TEV_PAPf39_INTEIN expression plasmid. Panel B. The sequence of the expressed construct: residues 1–83 correspond to the N-terminal 6H-tagged ubiquitin, residues 84–90 comprise the TEV protease recognition motif (cleavage between Q^{89} and G^{90}), residues 90–128 correspond to the PAPf39 peptide, and residues 129–388 correspond to the intein sequence.