Cloning, Purification, and Structural Characterization by 1D ^{1}H-NMR of the PDZ domain of the Shank3 protein

Meesook Sung*

Kwandong University, College of Medicine, Department of Biochemistry, Gangnung city, 210-701

Received January 19, 2007 / Accepted February 23, 2007

We wished to create a set of small molecular weight PDZ domain ligands that may be used in functional studies on the proteins AF6, PSD-95 and Shank. As a starting point, the Shank3 PDZ domain was cloned, purified, and characterized the structure of Shank3 PDZ domain by 1D ^{1}H-NMR. The chemical shift dispersion of the proton signals indicates that the purified Shank3 PDZ protein is very pure and globally well folded. Currently, we are working on improving the yield of the protein production for complete NMR structural analysis of the Shank3 PDZ domain.

Key words – Shank3 PDZ, 1D ^{1}H-NMR

Introduction

The scaffolding proteins contain multiple domains for protein-protein interactions which are playing in the assembly and recognition of signaling complexes. The Shank family of neuronal proteins consist of three family members, Shank1, Shank2 and Shank3. All shank proteins share a common domain organization consisting of N-terminal ankyrin repeats, SH3, PDZ and SAM motifs, and multiple proline-rich region [1-2]. All of these motifs are potentially involved in protein-protein interactions.

A most promising interaction domain with respect to ligand design is the PDZ domain. PDZ domains are 80-100 amino acids long and characterized by a hydrophobic pocket surrounded by a conserved sequence motif, G-L-G-F. This pocket interacts mostly with the C-terminal residues of peptide ligands, which can be divided into two major classes depending on the corresponding PDZ domains. Class I PDZ domains mainly interact with ligands containing the motif X-(Thr/Ser)-X-Leu (where X represents any amino acids). Class II PDZ domains prefer a C-terminal sequence of the type X-hydrophobic-X-hydrophobic. Other ligands can be divided into minor classes. Like most of the protein domains, PDZ domains show compact globular structure with their 'N' and 'C' termini close to one another in their folded state. The PDZ domain fold typically consists of 6 β-strands (βA - βF) forming two opposing anti-parallel sheets flanked by 2 α-helices [3].

PDZ domains are very important in dynamic organization of the cell. They play a central role in signalling by organizing network of receptors and in targeting selected cellular proteins to multi-protein complexes [4]. Not surprisingly, pathogenic mutations in PDZ domain causes diseases such as Usher-syndrome, Dejerine-Sottas Neuropathy and DiGeorge-Syndrome [5,6].

The elucidation of domain structure is important to understand protein functions and thereby facilitate pharmaceutical drug development [7]. A pre-requisite for structural analysis is the preparation of pure, natively folded protein. NMR has been proven to be a suitable method for evaluating the structural integrity of a protein either by hetero- or homonuclear techniques [8,9]. Especially we were attracted by the fact that a 1D ^{1}H-NMR spectrum already contains enough information to evaluate whether a protein folded. This enabled use of high-throughput structural analysis and screening methods to be implemented for rapid detection of suitable protein expression constructs [10].

One driven force in this work was eventually to understand how PDZ domains bind specifically to a variety of other proteins and to provide a PDZ/ligand specificity relationships. As a starting point, in this present work, we have cloned, purified the Shank3 PDZ domain and applied for 1D NMR. The result implies the Shank3 PDZ protein is very pure and natively well folded.

Materials and Methods

Cloning and Sequencing analysis

DNA fragment encoding residues 635 to 738 of Shank3

*Corresponding author
Tel : +82-33-649-7464, Fax : +82-33-641-1074
E-mail : msung@kwandong.ac.kr

PDZ domain was obtained by PCR amplification using Shank3 protein cDNA as a template (kindly provided by Prof. Dr. W. Birchmeier, Max-Delbruck-center for Molecular Medicine, Berlin, Germany) [11]. The designed PCR primers are as follows: Shank3_PDZ_Forward (5'-GACGACGACAAGATAGAAAACTTGTATTTCCAGGG CGTGGCTATCCTGC GAAAAG-3'); Shank3_PDZ_Reverse (5'-GAGGAGAAGCCCGGTTTACACAGA CACAACCTTCATGA-3'). The underlined sequences are designed for the TEV (tobacco etch virus) recognition site. The amplified PCR product was cloned into the pET-30 expression vector (Novagen, Madison, WI). The DNA insert were completely determined on both strands using a model 377 DNA sequencer (Applied Biosystems, Foster City, CA, USA).

Expression of Shank3 PDZ domain protein

The constructed plasmid was transformed into *E. coli* (DE3) competent cells prepared using the calcium phosphate method [12] and streaked onto plates. A single colony was grown overnight at 30 ℃ in LB medium supplemented with 30 μg/ml kanamycin. Cultures were then diluted in LB medium and grown at 37 ℃ until OD_{600nm} reached at 0.6-1.0. Expression of the protein was initiated by the addition of 1 mM β-D-thiogalactopyranoside(IPTG) (Sigma-Aldrich, St. Luis, MO, USA) and then cultures were grown for 5 hours more at different temperatures (22℃, 30℃ or 37℃) to obtain the appropriate clone. After induction, 1.5 mL culture was centrifuged (9000 rpm, 10 min) and the pellet was then resuspended in 100 μl of sample buffer [20 mM Tris, pH 8.0, 2 mM EDTA, 2% (w/v) SDS, 25% (v/v) glycerin, bromophenol blue] and electrophoresed on 12% SDS-polyacrylamide gel electrophoresis.

Purification of Shank3 PDZ domain protein for NMR

The soluble fraction of the Shank3 PDZ domain was purified by Ni-NTA affinity chromatography. In brief, a 2 liter culture of *E. coli* BL21 (DE3) cells carrying pET30-Shank3 PDZ domain was grown at 30℃ in M9 minimal media supplemented with $^{15}NH_4Cl$ (0.5 g/L) as the sole nitrogen source. All media contained 30 μg/mL kanamycin. Cells were induced and harvested as described above. The cell pellet was resuspended in the buffer containing 50 mM tris hydrochloride, 20 mM $MgCl_2$, 1 mM phenylmethanesulfonic acid, pH 8.0 and treated two rounds of thawing and freezing in dry- ice/ethanol. After incubating with 0.5 mg DNase I for 30 min at 37 ℃, cells were collected and disintegrated using FRENCH press (ThermoSpectronic, Rochester, USA). The resulting homogenate was centrifuged at 12,000x g for 30 min to pellet any insoluble materials. After the filtered supernatant was added to the Ni-NTA resin, the matrix was washed with 50 mM phosphate buffer (pH 8.0) containing 20 mM imidazole and 300 mM NaCl. The His-tagged protein was eluted by 5x volumes of elution buffer containing 250 mM imidazole to the column and concentrated using a Vivaspin concentrator (Vivasicence AG, Hannover, Germany). After His cleavage, a second Ni-NTA column was used to separate the His-tag from the PDZ domain. The factions were pooled and concentrated to a 0.55 mL solution and the buffer was changed to 20 mM phosphate buffer, 50 mM NaCl, pH 6.0.

His tag cleavage

The concentration of protein was estimated using the BCA assay (Pierce, Rockfold, IL, USA). Protein at a concentration of 0.5 mg/mL was incubated with 0.05 mg/mL of each proteases and allowed to incubate at 23℃ for 16 hours. The sample was then run on a SDS gel to check the completion of cleavage.

Determination of protein folding by 1D ^{1}H-NMR spectroscopy.

The 1D ^{1}H-NMR measurement was performed with the purified Shank3 PDZ domain (1.5 mg/0.55 mL, 0.06 mL D_20, 20 mM phosphate, 50 mM NaCl, pH 6.0) at 300K. Ten percents d_6DMSO was added to the protein solution and used as lock signal. Thee spectrometer DRX600 was operated using the program ICON-NMR on top of XWIN-NMR (Bruker, Karlsruhe, Germany).

Results

DNA fragment encoding Shank3 PDZ domain was cloned into a pET-30 expression vector (Novagen, Madison, WI). The PCR primers were designed with TEV (tobacco etch virus) site as a cost-effective alternative enzymatic site. Nucleotide sequencing of the cloned Shank3 PDZ domain (145 amino acids, molecular weight 16 kDa) was carried out to ensure the absence of PCR-induced mutation and the result was translated as shown in Fig. 1. For the highest expression level of the Shank3 PDZ domain, the

1 MHHHHHHSSG LVPRGSGMKE TAAAKFERQH MDSPDLGTDD

41 DDK*1 IENLYFQ *2 GVAILQKRDH EGFGFVLRGA KAETPIEEFT

81 PTPAFPALQY LESVDVEGVA WRAGLRTGDF LIEVNGVNVV

121 KVGHKQVVGL IRQGGNRLVM KVVSV

Fig. 1. Sequencing of the His-tagged Shank3-PDZ domain protein. Total amino acid residues are 145 amino acids. The preceding residues towards the N terminus are referred to as -1. *1 and *2 denote TEV and EK cleavage sites, respectively.

expression pattern was optimized by the variation of IPTG concentration and of temperature. Each *E. coli* cells carrying the Shank3 PDZ domain plasmid was grown in LB media at 30℃, and induced by addition of IPTG (OD_{600} =0.6 and 1.0). As seen in Fig. 2, the protein was well expressed but the yield of soluble protein was generally very low, implying that the Shanks PDZ domain could be aggregated in the inclusion bodies in the cytoplasm of *E. coli*. Nevertheless, the data suggested that the *E. coli* BL21(DE3) strain is effective host for this protein expression. The optimal growth conditions were summarized in the Table 1.

For the nuclear magnetic resonance (NMR) measurement, the Shank3 PDZ protein was expressed in M9 medium with $^{15}NH_4Cl$ and purified as described in the materials and methods. The His tagged Shank3 PDZ domain was treated with proteases (TEV and EK) for the tag cleavage. Thrombin was used as a non-specific protease. As shown in Fig. 3, the His-tag of Shanks PDZ was not removed by TEV protease. The cleavage of His-tag was observed by only EK protease. Thus, the tag-free Shank3 PDZ domain was finally obtained with the sequences of 102 amino acids and molecular weight 11 kDa.

1D ^{1}H-NMR is an established and sensitive method to check for folded protein [13]. Therefore, we tested whether the purified Shank3 PDZ protein is natively folded using 1D ^{1}H-NMR. In Fig. 4, large chemical shift dispersions, especially in the methyl group region between 0 ppm and 0.6 ppm and the amide region downfield of 9 ppm clearly shows that the protein has a well folded global structure. The protein exhibits some beta sheet structure.

Discussion

In this study, the PDZ domain of Shank3 protein was cloned in a modified pET 30 vector containing His-tagged and EK followed by TEV protease recognition sites (K*1IENLYFQG*2VAILQKRD, *1 and *2 denote EK and TEV recognition site, respectively). However, the His-tag free

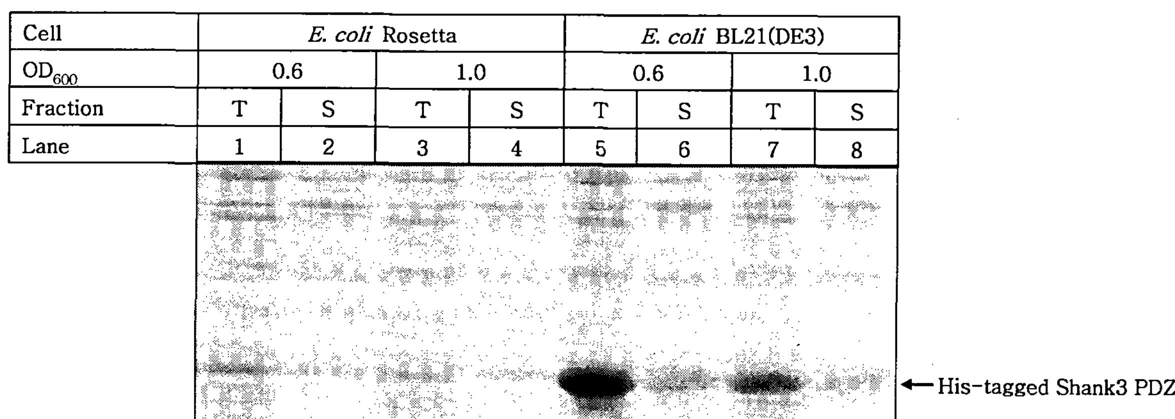

Fig. 2. Comparison of the expression of Shank3-PDZ domain protein in *E. coli* cells. The BL21(lane 1-4) and Rosetta (lane 5-8) cells were expressed as described in the materials and methods. The total cell lysate (T) and soluble (S) fractions were collected at OD_{600nm}= 0.6 and 1.0 and analyzed by 15% SDS-PAGE.

Table 1. Optimized Expression Conditions

Protein	IPTG[μM]	OD600	T[℃]#1	t[h]#2	Cell[g/ℓ]	Protein [mg/ℓ]
Shank3 PDZ	1.0	0.6	30	4-5	~7.5	7-8

#1: Temperature after IPTG induction.
#2: Growth time after IPTG induction.

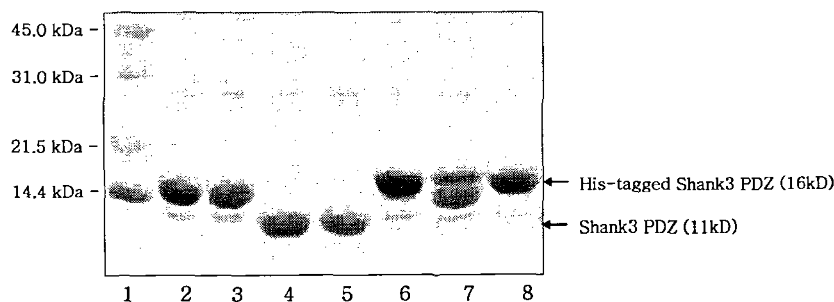

Fig. 3. Treatment of PDZ domain by proteases. The purified His-tagged Shank3-PDZ protein was incubated with each endogenous proteases at different temperatures for 16 hours. Lane 1 (Molecular Marker); Lane 2 (TEV, 15℃); Lane 3 (TEV, 23℃); Lane 4 (EK, 15℃); Lane 5 (EK, 23℃); Lane 6 (TEV, 37℃); Lane 7 (TEV + EK, 23℃); Lane 8 (Thrombin, 23℃).

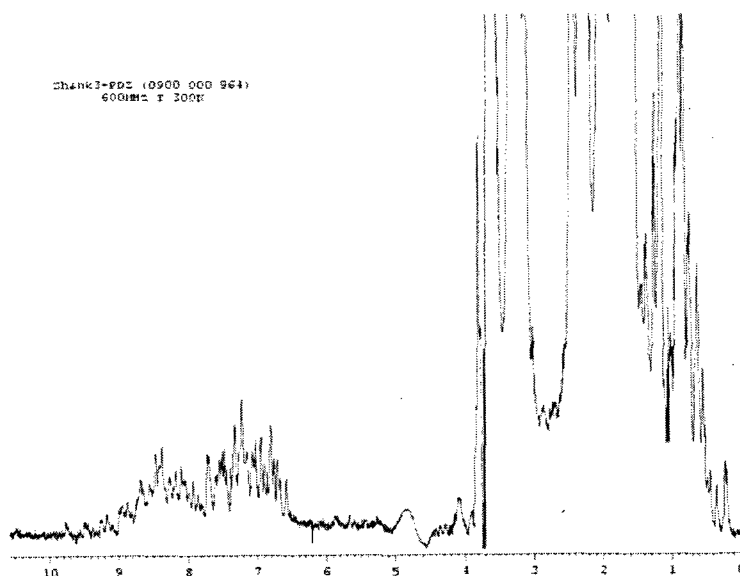

Fig. 4. 1D ^{1}H-NMR spectrum of the Shank3 PDZ domain. The measurement was performed at 0.22 mM in the buffer containing 20 mM phosphate and 50 mM NaCl, pH 6.0 at 300K in the DRX600 spectroscopy (Bruker, Karlsruhe, Germany).

Shank3 PDZ domain was obtained only by EK cleavage, not by TEV. Based on the fact that the conserved N-terminal βA strand of Shank1 PDZ domain spans long 16 residues, which participates dimeric interaction by forming an antiparallel β sheet with βA strand of the other monomer, we assumed that the extra octapeptides are necessary to keep the βA strand and maintain the whole compact structure of protein although the residues are not conserved.

The protein exhibits several amide proton peaks resonating at 9 ppm and alpha protons between 5 and 6 ppm, showing that the protein has beta sheet conformations. The weak intensity of these peaks arise from the low concentration of the NMR sample (0.22 mM). Note that methyl proton peaks (0-1 ppm) are much more intense than the amide and alpha ones because the three equivalent protons of the methyl groups contribute to the signal rather than one proton of the latters. In summary, the 1D NMR spectrum (Fig. 4) clearly shows that the Shank3 PDZ domain sample is very pure and globally well folded. Currently, we are working on improving the yield of the protein production. The complete NMR spectrum will be obtained once enough quantity of Shank3 PDZ domain is prepared from ^{15}N-labeled minimal media.

Acknowledgements

I gratefully thank to Dr. H. Oschkinat, Forchunginstitut fur Molekulare Pharmakologie, Berlin, Germany for supporting all materials. I also thank to Dr. I.J. Byeon, Department of Structural Biology, College of Medicne,

University of Pittsburgh for the critical advice of NMR data. This work was supported by the scientists exchange program between KOSEF and DFG.

References

1. Bockers, T. M., M. R. Kreutz, E. D. Gundelfinger and J. Bockers. 2002. Pro-Sap/Shank postsynaptic density proteins interact with insulin receptor tyrosine kinase substrate IRSp53. *J. Neurochem.* **83**, 1013-1017.
2. Sheng, M. and E. Kim. 2000. The shank family of scaffolding proteins. *J. Cell. Sci.* **113**, 1851-1856.
3. Im, Y. J., J. H. Lee, S. H. Park, S. J. Park, S. Rho, G. B., Kang, E. Kim and S. H. Eom. 2003. Crystal structure of the Shank PDZ-ligand complex reveals a class I PDZ interaction and a novel PDZ-PDZ dimerization. *J. Biol. Chem.* **278**, 48099-48104.
4. Zhang, M. and W. Wang. 2003. Organization of signalling complexes by PDZ-domain scaffolding proteins. *Acc. Chem. Res.* **36**, 530-538.
5. Montell, C. 2000. A PDZ protein ushers in new links. *Nat. Genet.* **26**, 6-7.
6. Boerkoel, C. F., H. Takashima, P. Stankiewicz, C. A. Garcia, S. M. Leber, L. Rhee-Morris and J. R. Lupski. 2001. Periaxin mutations cause recessive Dejerine-Sottas neuropathy. *Am J. Hum Genet* **68**, 325-333.
7. Buchanan, S. G., J. M. Sauder and T. Harris. 2002. The promise of structural genomics in the discovery of new antimicrobial agents. *Current Pharmacetical Design* **8**, 1173-1188.
8. Woestenenk, E. A., M. Hammarstron, T. Hard and H. Berglund. 2003. Screening methods to determine biophysical properties of proteins in structural genomics. *Anal. Biochem.* **318**, 71-79.
9. Gronenborn, A. M. and G. M Clore. 1996. Rapid screening for structural integrity of expressed proteins by heteronuclear NMR spectroscopy *Protein science* **5**, 174-177.
10. Scheich, C., D. Leitner, V. Sievert, M. Leidert, B. Schlegel, B. Simon, I. Letunic, K. Bussow and A. Diehl. 2004. Fast identification of folded human protein domains expressed in E. coli suitable for structural analysis. *BMC Structural Biol.* **4**, 1-9.
11. Scheutz, G., R. Marta, J. Grimm, T. Boeckers, E. Gundelfinger and W. Birchmeier. 2004. The neuronal scaffold protein Shank3 mediates signaling and biological function of the receptor tyrosine kinase Ret in epithelial cells. *J. Cell. Biol.* **167**, 945-952.
12. Sambrook, J. and D. Russell, 2001. Molecular Cloning, third edition Edition, New York, Cold Spring Harbor Laboratory Press 182-184.
13. Rehm T., R. Huber and T. A. Holak. 2002. Application of NMR in structural proteomics: Screening for proteins amenable to structural analysis. *Structure* **10**, 13-18.

초록 : Shank3 PDZ 도메인의 동정, 정제 및 1차 NMR 구조분석

성 미 숙*

(관동대학교 의과대학 생화학교실)

PDZ 도메인을 통하여 여러 단백질과 상호작용하며 신경전달 기전에 관여하는 단백질로 Shank1, Shank2, Shank3, PDS-95, AF있다. 본 연구는 Shank3 PDZ 도메인의 구조를 밝히기 위한 첫 단계로서 Shank3 단백질의 PDZ 도메인을 동정하였고, E. coli에서 발현하여 생성된 단백질을 정제한 후 1차 NMR 구조분석을 시도하였다. 그 결과에 의하면 정제된 Shank3 PDZ 단백질은 순도가 높고 안정적인 접힙(folding)구조를 제시하고있다. 현재 완전한 NMR 구조분석을 위해 좀더 많은 양의 정제된 Shank3 PDZ 단백질을 얻고자 연구하고 있다.