Journal of the Korean Magnetic Resonance Society **2017**, *21*, 50-54
DOI 10.6564/JKMRS.2017.21.2.050

Backbone NMR Assignments and Secondary Structure Determination of a Cupin-family Protein YaiE from *Escherichia coli*

Sung-Hee Lee[1†], Dae-Won Sim[1†], Eun-Hee Kim[2], Ji-Hun Kim[3,*], and Hyung-Sik Won[1,*]

[1]Department of Biotechnology, College of Biomedical and Health Science, Konkuk University, Chungbuk 27478, Korea.
[2]Protein Structure Group, Korea Basic Science Institute, Ochang, Chungbuk 28119, Korea
[3]College of Pharmacy, Chungbuk National University, Cheongju, Chungbuk 28160, Korea.

Received Apr 02, 2017; Revised Apr 24, 2017; Accepted Apr 29, 2017

Abstract Cupin-superfamily proteins represent the most functionally diverse groups of proteins and include a huge number of functionally uncharacterized proteins. Recently, YaiE, a cupin protein from *Escherichia coli* has been suggested to be involved in a novel activity of pyrimidine/purine nucleoside phosphorylase (PPNP). In the present study, we achieved a complete backbone NMR assignments of YaiE, by a series of heteronuclear multidimensional NMR experiments on its [^{13}C/^{15}N]-enriched sample. Subsequently, secondary structure analysis using the assigned chemical shift values identified 10 obvious β-strands and a tentative 3_{10}-helix. Taken all together, the results constitute the first structural characterization of a putative PPNP cupin protein.

Keywords YaiE, cupin, purine/pyrimidine nucleoside phosphorylase, secondary structure, NMR

Introduction

The ubiquitous proteins that share a characteristic β-barrel fold are classified into a cupin superfamily, of which name originated from the Latin term *cupa* for a small barrel.[1] Although typical cupin proteins such as seed storage proteins and some known allergens were originally discovered in higher plants, the members of this protein family were extended to include a wide variety of enzymes and sugar-binding proteins from various bacterial species.[2] The cupin superfamily is now appreciated as the most functionally diverse groups of proteins divided into 61 subfamilies in the Pfam database (http://pfam.xfam.org/). In addition, an appreciable population of the proteins remains as functionally uncharacterized or hypothetical ones. Thus, structure-functional identifications of cupin proteins are generally expected to discover a novel function of the protein family and therefore to provide molecular principles of their functional and evolutional diversification.

Recently, Sévin *et al*. have demonstrated a high-throughput *in vitro* metabolomics approach to discover novel enzymes.[3] In particular, their novel identifications included a potential purine/pyrimidine nucleoside phosphorylase (PPNP) activity for the 94 amino-acid protein YaiE from the *E. coli* K12 strain. PPNPs catalyzes phosphorolysis of nucleosides; *i.e.*, the purine and pyrimidine nucleosides are reacted with phosphates to yield α-D-ribose-1-phosphate and either purine or pyrimidine bases.[4-6] Thus, the PPNP enzymes are involved in cellular nucleotide salvage pathways, which allow the cell to produce nucleotide monophosphates when the *de novo* purine and

† These two authors contributed equally to this work.
* Ccorrespondence to: **Hyung-Sik Won,** Department of Biotechnology, College of Biomedical and Health Science, Konkuk University, Chungju, Chungbuk 27478, Korea, Tel: 82-43-840-3589, E-mail: wonhs@konkuk.ac.kr ; or **Ji-Hun Kim,** College of Pharmacy, Chungbuk National University, Cheongju, Chungbuk 28160, Korea, Tel: 82-43-249-1343, E-mail: nmrjhkim@chungbuk.ac.kr.

pyrimidine biosynthetic pathways are unavailable. In particular, many human pathogens including many parasites lack the *de novo* purine nucleotide biosynthesis pathway. In addition, the purine nucleoside phosphorylase (PNP)-deficiency in human leads to T-cell immunodeficiency. Accordingly, various PNP inhibitors are used and being developed as immunosuppressing medicines and potent therapeutic agents treating malignant tumors and infectious pathogens. Meanwhile, we had earlier noted that YaiE could be categorized into a functionally uncharacterized cupin-family protein in the Pfam database. In spite of the extremely diverse functions of known cupin proteins, the PPNP activity had not been ever reported. Thus, structural identification of the PPNP functionality is required for YaiE, to validate its structural and functional relevance and significance as a PPNP protein. In this respect, we report a preliminary structural characterization of YaiE, using NMR in solution.

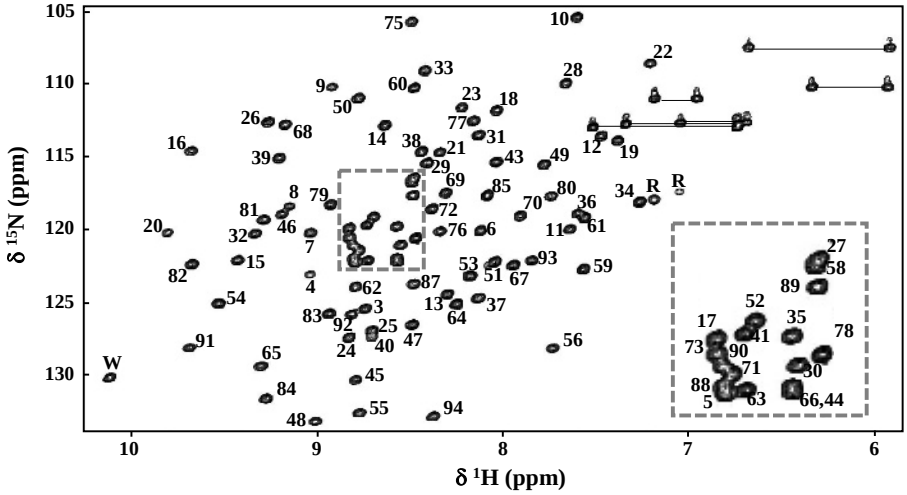

Figure 1. 2D [^{1}H-^{15}N]-HSQC spectrum of YaiE. Each resonance in the spectrum is labeled with the corresponding residue number assigned. Sidechain NH_2 signals from Asn and Gln residues are indicated by horizontal lines. Sidechain NH signals from Trp and Arg residues are marked by 'W' and 'R', respectively. Boxed region was enlarged for clarity.

Experimental Methods

Recombinant plasmid construction- Genomic DNA was extracted from the *E. coli* BL21(DE3)pLysS strain and used as a template for cloning. The ORF corresponding to YaiE-coding region of the genomic DNA was PCR-amplified, using appropriate primers containing 5'-*Nde*I and 3'-*Xho*I restriction sites. The 5' and 3' ends of the PCR products were digested with *Nde*I and *Xho*I restriction endonucleases, respectively, followed by ligation into a pET21a vector restricted by the same enzymes. Since the reverse primer contained a stop codon preceding the *Xho*I restriction site, the resulting translation could be relevant exactly to the native YaiE sequence not harboring the C-terminal His-tag.

Preparation of isotope-enriched protein sample- The constructed plasmids were confirmed via DNA sequencing and transformed into BL21-CodonPlus cells. The transformed cells were then grown at 37 °C in M9 minimal media supplemented with D-glucose-$^{13}C_6$ and ammonium-^{15}N chloride as the sole source of carbon and nitrogen, respectively. The cells were harvested after 5-hrs protein expression, which was induced by adding IPTG to a final concentration of 1 mM when the A_{600} reached about 0.6. Harvested cells were resuspended in 20 ml of lysis buffer (20 mM Tris-HCl, 150 mM NaCl, pH 7.5) per gram of wet cells. The cells were then disrupted by sonication, and from the supernatant, the YaiE was purified via sequential chromatography on

two types of anion-exchange (HiTrap® Q FF and Q HP, GE Healthcare) columns and a gel-filtration (Superdex® 75, GE Healthcare) column. The purified YaiE solution was concentrated by ultrafiltration and buffer-exchanged by diafiltration. The protein concentration was spectrophotometrically deduced using the molar absorptivity at 280 nm, which was predicted from the amino acid sequence.

NMR measurements and analysis- NMR sample contained 1 mM of [^{13}C/^{15}N]-YaiE in PBS at pH 7.4. Conventional 2D [^{1}H-^{15}N]HSQC, 3D HN(CA)NNH, and a series of triple resonance spectra {HNCACB, HN(CO)CACB, HNCO, and HN(CA)CO} were measured on a Bruker Biospin Avance 800 spectrometer at 298 K. NMR data processing, including chemical shift referencing, was performed using NMRPipe/NMRDraw software,[7] followed by spectral visualization and analysis using NMRViewJ program.[8] Sequential assignments were achieved by a typical method primarily based on triple resonance spectra. For the chemical shift index (CSI)[9] analysis, the CSI parameters in the Biological Magnetic Resonance Data Bank (BMRB) were referred to. Backbone dihedral angles of individual amino acid residues were predicted from the assigned chemical shift values, using the TALOS-N program.[10]

Results

Backbone NMR assignments- As represented by 2D [^{1}H-^{15}N]-HSQC in Figure 1, YaiE appeared to adopt a homogeneously well-folded conformation,

Table 1. Sequence-specific backbone NMR assignments of YaiE. Assigned chemical shift values (ppm) of $^{1}H^{N}$, ^{15}N, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and $^{13}C'$ atoms are summarized (n.d., not detected; n.a., not available).

Seq.	$^{1}H^{N}$	^{15}N	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	Seq.	$^{1}H^{N}$	^{15}N	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	Seq.	$^{1}H^{N}$	^{15}N	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$
M 1	n.d.	n.d.	n.d.	n.d.	n.d.	G 33	8.49	109.2	44.7	n.a.	170.5	Y 65	9.39	129.5	58.3	39.5	174.9
L 2	n.d.	n.d.	53.7	43.4	176.6	E 34	7.34	118.2	54.7	31.9	174.7	E 66	8.65	122.0	55.1	31.5	176.3
Q 3	8.82	125.5	55.0	30.9	175.3	Y 35	8.63	119.8	56.7	44.3	174.7	A 67	8.01	122.5	54.5	18.4	177.9
S 4	9.10	123.1	58.6	63.6	173.2	T 36	7.68	119.1	62.2	70.7	172.9	G 68	9.25	112.9	44.8	n.a.	174.5
N 5	8.88	122.2	52.6	42.9	174.0	F 37	8.23	124.9	55.7	42.8	174.1	S 69	8.38	117.6	59.5	65.4	171.5
E 6	8.19	120.2	55.1	33.5	173.4	S 38	8.51	114.7	56.9	65.0	175.1	V 70	7.98	119.1	59.9	35.6	175.6
Y 7	9.11	120.3	55.2	42.4	175.7	T 39	9.27	115.2	59.7	70.4	174.1	F 71	8.85	121.5	55.4	40.4	171.8
F 8	9.21	118.5	59.5	37.1	177.2	A 40	8.80	127.4	52.7	18.2	175.6	N 72	8.45	118.6	52.5	41.8	174.0
S 9	8.99	110.3	58.8	61.8	175.1	E 41	8.78	119.7	56.6	27.5	173.5	V 73	8.91	120.7	58.4	34.5	174.0
G 10	7.68	105.5	45.4	n.a.	174.2	P 42	n.a.	n.d.	62.9	31.9	174.7	P 74	n.a.	n.d.	62.2	32.6	176.3
K 11	7.71	120.0	59.1	33.3	174.3	E 43	8.09	115.5	53.5	33.5	175.0	G 75	8.55	105.8	44.5	n.a.	173.4
V 12	7.55	113.7	60.8	35.6	175.7	E 44	8.65	122.3	55.1	33.8	175.2	H 76	8.42	120.4	56.8	27.3	174.3
K 13	8.37	124.5	55.3	38.3	173.8	M 45	8.86	130.4	53.6	34.3	173.2	S 77	8.24	112.6	57.4	66.3	172.5
S 14	8.71	112.9	56.7	68.3	173.3	T 46	9.27	119.1	61.3	70.5	174.9	E 78	8.53	120.6	55.0	33.0	174.9
I 15	9.50	122.2	59.8	40.9	176.0	V 47	8.56	126.6	63.7	31.1	175.8	F 79	8.99	118.3	55.4	41.5	171.4
G 16	9.75	114.7	44.9	n.a.	173.0	I 48	9.10	133.3	62.4	36.1	175.6	H 80	7.76	118.0	54.1	32.5	173.9
F 17	8.90	119.9	55.9	41.7	171.9	S 49	7.86	115.7	57.7	66.0	171.4	L 81	9.39	119.3	54.1	47.6	176.7
S 18	8.12	111.9	56.9	65.4	173.4	G 50	8.84	111.0	44.8	n.a.	173.4	Q 82	9.74	122.4	56.0	31.3	174.7
S 19	7.45	113.9	57.8	68.0	174.9	A 51	8.13	122.3	52.5	21.6	175.7	V 83	9.02	126.0	62.4	34.8	175.4
S 20	9.84	120.3	62.2	67.9	176.8	L 52	8.76	119.3	53.2	46.1	174.3	A 84	9.35	131.7	53.4	20.2	176.5
S 21	8.43	114.8	60.5	63.3	176.1	N 53	8.24	123.3	52.2	39.4	175.7	E 85	8.15	117.7	54.0	31.3	171.1
T 22	7.29	108.6	62.2	71.8	177.1	V 54	9.59	125.1	61.2	36.1	174.7	P 86	n.a.	n.d.	64.7	31.7	175.5
G 23	8.30	111.7	45.9	n.a.	175.4	L 55	8.87	132.8	53.0	41.0	175.9	T 87	8.56	123.8	63.0	75.2	172.3
R 24	8.91	127.6	58.4	30.6	175.5	L 56	7.77	128.1	53.1	40.9	174.8	S 88	8.87	121.9	56.4	68.0	174.7
A 25	8.78	127.0	51.3	23.3	176.7	P 57	n.a.	n.d.	64.6	31.3	176.9	Y 89	8.56	117.8	56.4	41.5	171.0
S 26	9.34	112.7	57.7	66.2	173.1	D 58	8.55	116.5	55.2	39.9	174.8	L 90	8.87	121.2	53.4	45.9	174.1
V 27	8.56	116.7	58.4	35.1	174.4	A 59	7.64	122.8	51.7	21.7	177.7	C 91	9.77	128.3	56.5	29.1	173.3
G 28	7.73	110.0	45.9	n.a.	170.8	T 60	8.55	110.3	62.6	69.3	174.1	R 92	8.88	125.8	55.0	32.7	174.4
V 29	8.49	115.5	60.0	35.8	173.2	D 61	7.63	119.2	52.6	43.8	175.1	Y 93	7.91	122.2	57.8	38.2	175.2
M 30	8.62	121.1	53.5	36.2	176.0	W 62	8.87	124.1	58.3	29.3	175.7	L 94	8.46	132.9	56.0	42.6	181.7
V 31	8.20	113.5	60.9	32.2	175.5	Q 63	8.80	122.2	54.8	33.1	173.6						
E 32	9.38	120.4	57.6	29.1	175.6	V 64	8.32	125.2	62.3	32.5	175.7						

displaying a well-dispersed NMR spectra with narrow line shapes. For the NMR assignments, peak clusters were built up by engaging individual peaks in the [^{1}H-^{15}N]-HSQC spectrum with their corresponding resonances having the same amide $^{1}H^{N}$ and ^{15}N chemical shifts in HN(CA)NNH and triple resonance spectra. The established peak clusters were then sequentially linked according to ^{13}C chemical shift correlations,[11-13] *i.e.* the pair of HN(CO)CACB and HNCACB spectra provided the sequential $^{13}C^{\alpha}$(i-1)/$^{13}C^{\beta}$(i-1) to $^{13}C^{\alpha}$(i)/$^{13}C^{\beta}$(i) connectivities, while the HNCO-HN(CA)CO pair were employed for the carbonyl $^{13}C'$(i-1) to $^{13}C'$(i) connectivities. In part, the sequential linkage was further supplemented by the HN(CA)NNH spectrum that provided ^{15}N(i-1)-^{15}N(i)-^{15}N(i+1) connectivities. Finally, sequence-specific assignments could be completed for all of the observed backbone amide signals in the [^{1}H-^{15}N]-HSQC spectrum (Figure 1). Sidechain NH resonances from the single Trp and two Arg (aliasing peaks) residues, as well as six pairs of sidechain NH_2 signals from three Gln and three Asn residues, were also clearly distinguished from the backbone amide resonances. The assigned backbone chemical shift values are summarized in Table 1. Generally, NH resonances from proline residues, which lack amide protons attached to their backbone nitrogens, and the N-terminal methionine (M1), of which amino proton undergoes a rapid exchange with the solvent, cannot be detected in the present NMR experiments. Amide NH of the second residue is also known to rarely show up as a signal, due to dynamic fluctuation in N-terminal region and/or post-translational cleavage of the N-terminal methionine. Meanwhile, it was possible to assign the ^{13}C chemical shifts of the second residue (L2) and the four Pro residues, based on the inter-residue ^{13}C correlations with their neighboring residues, whereas their ^{15}N correlations were not detected in the HN(CA)NNH spectrum probably due to poor spectral sensitivity. Therefore, the extent of sequence-specific backbone NMR assignments could be regarded as almost 99% of measureable nuclei.

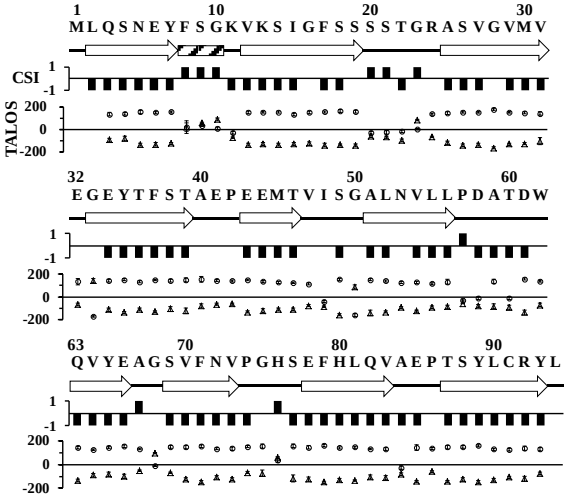

Figure 2. Secondary structure determination of YaiE. Consensus ^{13}C-CSI and TALOS ϕ/φ (circle/triangle with error bars) prediction results are summarized along the amino acid sequence (refer to the main text for details). Determined secondary structure elements are represented by arrows for β-strands and a scratched box for 3_{10}-helix.

Secondary structure determination- Using the assigned chemical shift values, the consensus ^{13}C-CSI was determined by a simple majority (two or three out of three) rule combining individual $^{13}C^{\alpha}$-, $^{13}C^{\beta}$-, and $^{13}C'$-CSIs, where the indexes '1' (0.7 ppm or more downfield/upfield shift against random-coil $^{13}C^{\alpha}$/$^{13}C^{\beta}$ reference values and 0.5 or more downfield shift against $^{13}C'$) and '-1' (the opposite pattern to '1') indicate α-helical and β-stranded tendencies, respectively.[9] TALOS analysis outputs the backbone dihedral ϕ and φ angles predicted statistically with standard deviation.[10] By combining the CSI and TALOS results, 10 obvious β-strands were defined for YaiE, in the regions showing consecutive β tendencies (Figure 2). Additionally, a 3_{10}-helix was tentatively suggestable in the region of F8 to G10. Except for that putative 3_{10}-helix, the secondary structure of YaiE was quite similar to that of a *Helicobacter pylori* cupin protein HP0902, of which monomer (99 amino acid residues) is also comprised of 10 β-strands forming a jelly roll β-sandwich fold.[12-14] Thus, the present results verify that YaiE could be also folded into a minimalistic cupin fold resembling the HP0902 structure, although they share no significant sequence homology.

Discussion

The present NMR results verified in solution that YaiE could belong to the cupin-family of proteins, by showing its predominantly β-stranded topology. Our database searches indicated that YaiE-homologous proteins are mostly unknown in function (data not shown). In contrast, the determined secondary structure of YaiE was quite similar to that of the *H. pylori* protein HP0902. HP0902 has been recently identified as the first example of a cupin protein that is capable of binding to the bacterial endotoxin, lipopolysaccharides (LPS), although its functional significance remained to be investigated.[14] Thus, an LPS-interaction of YaiE would be worthy of examination for discovery of novel functionality of cupin proteins. Furthermore, YaiE has been recently suggested as a novel PPNP enzyme,[3] in spite of no known examples of nucleoside phosphorylase (NP) in cupin-family proteins. To our knowledge, no structural basis has been demonstrated for PPNPs that act on both types of nucleosides, although many structures are available for either purine nucleoside phosphorylases (PNPs) or pyrimidine nucleoside phosphorylases (PyNPs).[4] As YaiE shares no significant sequence homology to known PNPs and PyNPs, structural identification of its molecular interaction with nucleosides and phosphate is required to validate YaiE as a genuine PPNP enzyme. In addition, tertiary and quaternary structural details of YaiE would provide novel information for the molecular basis of NP enzymes. Thus, we expect that the present assignment data would be essentially used for those in-depth approaches to the atomic structure, molecular interaction, and enzymatic functionality of the novel bacterial cupin protein YaiE.

Acknowledgements

This work was supported by Konkuk University. This study made use of NMR machine at Korea Basic Science Institute (KBSI), Ochang, Korea.

References

1. J. M. Dunwell, A. Purvis, and S. Khuri, *Phytochemistry* **65**, 7 (2004)
2. R. Uberto and E.W. Moomaw, *PLoS One* **8**, e74477 (2013)
3. D. C. Sévin, T. Fuhrer, N. Zamboni, and U. Sauer, *Nat. Methods* **14**, 187 (2017)
4. M. J. Pugmire and S. E. Ealick, *Biochem. J.* **361**, 1 (2002)
5. R.G. Silva, J.E. Nunes, F. Canduri, J.C. Borges, L.M. Gava, F.B. Moreno, L.A. Basso, and D.S. Santos, *Curr. Drug Targets* **8**, 413 (2007)
6. G. Liechti and J.B. Goldberg, *J. Bacteriol.* **194**, 839 (2012)
7. F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, *J. Biomol. NMR* **6**, 277 (1995)
8. B. A. Johnson, *Methods Mol. Biol.* **278**, 313 (2004)
9. D. S. Wishart, and B. D. Sykes, *J. Biomol. NMR* **4**, 171 (1994)
10. Y. Shen, and A. Bax, *J. Biomol. NMR* **56**, 227 (2013)
11. Y. S. Lee, W. S. Yoon, I. Chung, K. Y. Chung, H. S. Won, and M. D. Seo, *J. Kor. Magn. Reson. Soc.* **19**, 36 (2015)
12. D. W. Sim, H. C. Ahn, and H. S. Won, *J. Kor. Magn. Reson. Soc.* **13**, 108 (2009)
13. D. W. Sim, Y. S. Lee, J. H. Kim, M. D. Seo, B. J. Lee, and W. S. Won, *BMB Rep.* **42**, 387 (2009)
14. D. W. Sim, J. H. Kim, H. Y. Kim, J. H. Jang, W. C. Lee, E. H. Kim, P. J. Park, K. H. Lee, and H. S. Won, *FEBS Lett.* **590**, 2997 (2016)