Journal of the Korean Magnetic Resonance Society **2009**, *13* 117-125

Backbone ^{1}H, ^{15}N and ^{13}C Resonance Assignment and Secondary Structure Prediction of HP0062 (O24902_HELPY) from *Helicobacter pylori*

Sun-Bok Jang[1], Chao Ma[1], Sung Jean Park[1], Ae-Ran Kwon[2] and Bong-Jin Lee[1*]

[1] *Research Institute of Pharmaceutical Sciences, College of Pharmacy, Seoul National University, San 56-1, Shillim-Dong, Kwanak-Gu, Seoul 151-742, Korea*

[2] *Department of Herbal Skin Care, College of Herbal Bio-Industry, Daegu Haany University, 290, Yugok-Dong, Gyeongsan-Si, Gyeongsangbuk-Do, 712-715, Korea*

Abstract : HP0062 is an 86 residue hypothetical protein from *Helicobacter pylori* strain 26695. HP0062 was identified ESAT-6/WXG100 superfamily protein based on structure and sequence alignment and also contains leucine zipper domain sequence. Here, we report the sequence-specific backbone resonance assignment of HP0062. About 97.7% of all $^{1}H_N$, ^{15}N, $^{13}C_\alpha$, $^{13}C_\beta$ and ^{13}C=O resonances were assigned unambiguously. We could predict the secondary structure of HP0062 by analyzing the deviation of the $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts from their respective random coil values. Secondary structure prediction shows that HP0062 consist of two α-helices. This study is a prerequisite for determining the solution structure of HP0062 and can be used for the study on interaction between HP0062 and DNA and other *Helicobacter pylori* proteins.

Keywords : *Helicobacter pylori,* HP0062, Leucine zipper domain, ESAT-6/WXG100 superfamily, NMR

INTRODUCTION

The identification of the clinical consequences of *Helicobacter pylori* infection is assuredly one of the major discoveries within the past 20 years in medicine. It revolutionized pathogenic understanding and therapeutic concept in gastroenterology. It is a

* To whom correspondence should be addressed. E-mail : lbj@nmr.snu.ac.kr

gram-negative flagellated bacterium that infects half of the world's population and is responsible for majority cases of gastric and duodenal ulcers [1]. *H. pylori* colonizes the human gastric mucosa and persists in the stomach of patients with or without the clinical symptoms for a life time [2]. Studies on *H. pylori* genomes have been focused on attempt to understand pathogenesis, the ability of this *H. pylori* organism to cause disease. Approximately 29% of the loci are in the "pathogeneesis" category of the genome database [3].

The availability of complete genome sequences of the genomes (*H. pylori* strain 26695) could make it possible to focus on the structure of protein, which is likely to be helpful to find its functions, because the function of protein is almost mediated by its structure [2]. As part of our structural genomics effort on *Helicobacter pylori* [4], we could study the solution structure of the HP0062 belonging to the ESAT-6/WXG100 superfamily possessing leucine zipper domain sequence [5]. The HP0062 has 86 amino acids which is identified as hypothetical protein from *Helicobacter pylori* strain 26695 with calculated pI value of 4.6 and molecular weight of 10516.61 Da. The ESAT-6/WXG100 superfamily is composed of about 100 amino acids, exported from the cell with no signal peptides but with Trp-X-Gly motif and has extensive coiled coil topology [6]. As a result of previous studies, HP0062 possess the sequence of leucine zipper domain which is a secondary structural motif found in proteins and creates adhesion forces in parallel α-helices. These proteins are related to DNA-Protein interaction involved in gene expression regulation. Here, we report the sequence-specific backbone resonance assignments and predict the secondary structure of HP0062.

EXPERIMENTAL

Protein Sample Preparation

The predicted ORF of HP0062 was cloned into the pET-21a(+) expression vector and transformed into *Escherichia coli* strain BL21 (DE3) host cells for large-scale protein production. The cells were grown at 37℃ in M9 media that contains $^{15}NH_4Cl$ and $^{13}C_6$-glucose (Cambridge Isotope Laboratories, Inc.) as nitrogen and carbon sources respectively.

When the OD_{600} of cells reached at 0.5, 0.25mM IPTG was added. After further cell growth for 4hrs at 37℃, the cells were harvested. The recombinant HP0062 protein was purified by applying the supernatant from the cell lysate onto a Ni^{2+}-NTA column (Qiagen; 3ml of resin per liter of cell culture) previously equilibrated with the binding buffer [7]. The Fractions containing protein were concentrated to about 1ml and applied to superdex-75(Pharmacia) column that had been equilibrated with the final buffer (50mM sodium phosphate, pH=6.8 , 150mM NaCl, 300mM Glycine, 1mM DTT, 0.1mM PMSF). The NMR sample was ~1mM ^{15}N- and $^{15}N/^{13}C$-labled protein prepared by 90% H_2O/10% D_2O buffer solution.

NMR spectroscopy

All NMR spectra were recorded at 308K on Bruker AVANCE 500, 600 and 900 (equipped with a cryoprobe) spectrometers. The backbone and side chain assignments were performed by using 3D HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HBHA(CO)NH, ^{15}N-TOCSY-HSQC, C(CO)NH-TOCSY spectra. Aromatic ring resonances were assigned by using 3D ^{15}N-NOESY-HSQC (mixing time 100ms), ^{13}C-NOESY-HSQC (mixing time 100ms). Chemical shifts were referenced to DSS externally. NMR spectra were processed using the program NMRPipe/nmrDraw [8] and analyzed with the program NMRView [9]. The secondary structure was predicted from the values using Chemical Shift Index (CSI) [10] and the backbone torsion angles were obtained using TALOS program[11] which use a combination of ^{15}N, $^{13}C_{\alpha}$, $^{13}C_{\beta}$, $^{1}H_{\alpha}$ and ^{13}C=O chemical shift of triplet of adjacent residues.

RESULTS AND DISCUSSION

^{1}H-^{15}N HSQC spectrum of ^{15}N-enriched HP0062 showed good signal dispersion (Fig. 1). Sequence specific assignments for backbone were performed. Among the four Ala residues, Ala43 could not be assigned. Although residue specific labeling experiment was performed using specific ^{15}N source to substitute the Ala-^{14}N residues with Ala-^{15}N to find the Ala43 resonances, the result of this experiment did not help to find the Ala43 peak

(Fig. 2). This residue was not observed due to the rapid exchange of the amide hydrogens. The assignments comprise 97.7 % of all $^{1}H_{N}$, ^{15}N, $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and $^{13}C=O$.

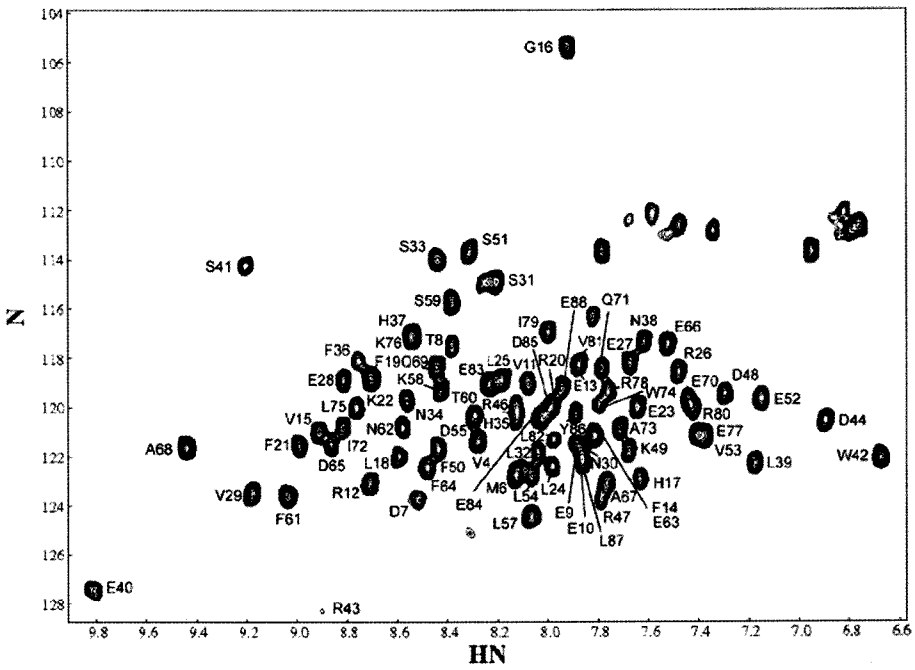

Fig. 1. 2D ^{1}H and ^{15}N TROSY spectrum of HP0062. The each resonance in the spectrum is labeled with assigned amino acid residues. Unassigned peaks are Trp sidecahin and sidechain of Gln and Asn.

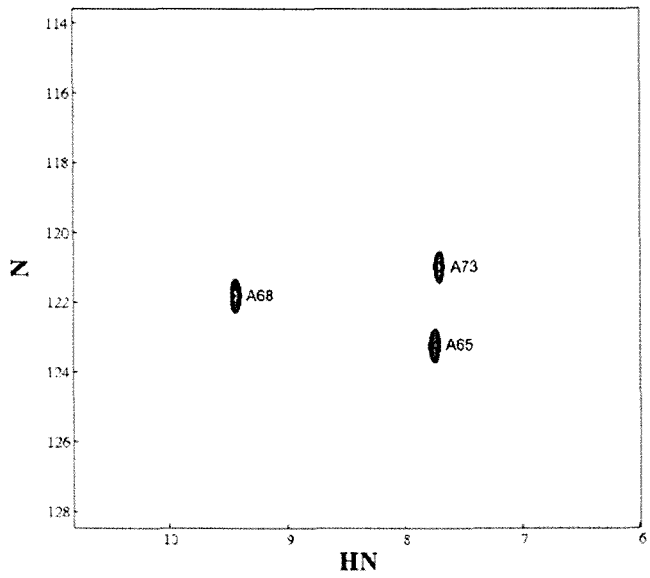

Fig. 2. Residues (Ala) specific ^{15}N labeled TROSY spectrum of HP0062

We used chemical shift difference method between measured values and random-coil values using C_α, C_β and ($\Delta C_\alpha - \Delta C_\beta$) and CSI program for predicting secondary structure of HP0062 [10-12]. Correlations have been observed between C_α and C_β chemical shifts and the local backbone conformation showed similarity with those of a number of proteins. Backbone dihedral angles (Φ,Ψ) are predicted using TALOS program from chemical shifts. Comparing relative random coil chemical shifts, C_α resonances tend to shift upfield in β-sheets and extended strands, and they tend to shift downfield in α-helices. The opposite trend holds for the C_β resonances. Because the C_α and C_β secondary shifts are of similar magnitude and opposite sign for both helices and sheets, subtraction of the C_α and C_β secondary shifts ($\Delta C_\alpha - \Delta C_\beta$) enhances the correlation between the secondary structural elements and the secondary shifts [13-14]. As shown in Fig. 3, examination of ($\Delta C_\alpha - \Delta C_\beta$) plot

indicates the presence of two potentially helical regions. The region of α-helices correspond well to the CSI and TALOS predictions.

Because all backbone amide ($^{1}H_N$ and ^{15}N) resonances were assigned, TROSY spectrum of HP0062 (Fig. 1) can be used to detect the DNA-protein and protein-protein interactions.

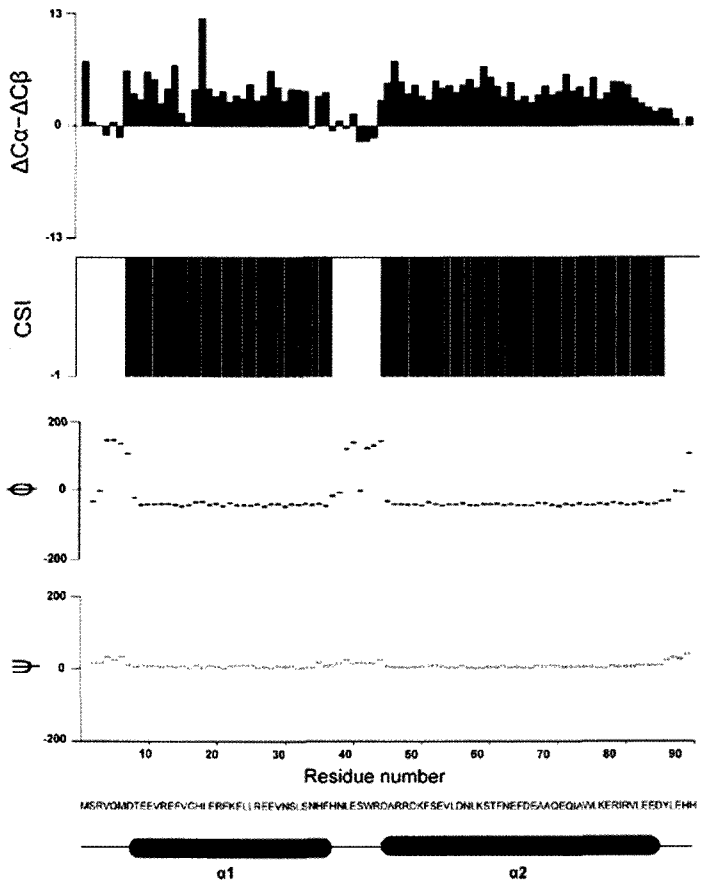

Fig. 3. Summary of backbone resonance assignment of HP0062. Delta values ($\Delta C_\alpha - \Delta C_\beta$) of backbone carbon to random coil chemical shifts were plotted. In the consensus CSI, the values '-1' represents the α-helical tendency. Backbone dihedral angels (phi, psi) were calculated using TALOS. Filled and open circles indicated the predicted phi (Φ) and psi (Ψ) angles, respectively. Predicted secondary structure elements of HP0062 were shown in black boxes for α-helices.

Table 1. Chemical shifts of $^{1}H_{N}$, ^{15}N, ^{13}CO, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ of HP0062. All chemical shifts were referenced to the frequency of the methyl proton resonances of DSS

Residue	HN	N	CO	CA	CB	Residue	HN	N	CO	CA	CB
1MET	ND	ND	ND	ND	ND	46ARG	7.63	120.38	178.04	58.17	28.16
2SER				66.66	64.98	47ARG	7.37	124.27	178.47	58.85	26.33
3ARG	7.52	120.09	175.57	55.64	30.17	48ASP	6.85	119.68	178.53	56.69	38.82
4VAL	7.78	122.81	175.28	61.73	32.40	49LYS	7.24	122.10	178.52	57.73	30.97
5GLN	7.92	122.81	174.83	54.66	29.43	50PHE	7.99	122.23	176.04	60.80	38.01
6MET	7.76	123.19	173.87	56.36	33.45	51SER	7.83	113.82	176.07	60.78	62.86
7ASP	8.11	124.23	176.43	52.53	40.74	52GLU	6.65	119.73	178.53	58.52	28.89
8THR	8.02	118.16	175.45	65.40	67.08	53VAL	6.91	121.74	178.77	65.11	30.67
9GLU	7.48	122.41	178.46	58.58	28.18	54LEU	7.67	123.03	178.07	57.05	40.04
10GLU	7.51	122.43	180.32	58.40	28.72	55ASP	7.82	121.74	178.88	57.00	39.32
11VAL	7.69	119.34	177.16	66.38	30.91	56ASN	7.59	120.74	177.32	55.56	37.58
12ARG	8.27	123.49	179.26	59.99	29.55	57LEU	7.56	124.95	177.66	57.69	40.34
13GLU	7.44	119.36	179.46	58.26	29.01	58LYS	8.24	119.50	178.29	59.84	31.37
14GLU	7.33	121.31	175.52	61.52	39.02	59SER	7.88	116.04	178.34	61.35	62.50
15VAL	8.47	121.44	176.91	66.75	30.51	60THR	7.84	120.76	176.56	66.39	67.51
16GLY	7.45	105.56	175.64	46.52	46.63	61PHE	8.58	123.99	176.68	60.62	36.94
17HIS	7.13	123.07	176.61	56.15	29.75	62ASN	8.14	121.24	177.26	56.03	37.32
18LEU	8.13	122.25	177.60	57.48	40.61	63GLU	7.35	122.16	179.20	58.71	28.73
19PHE	8.24	119.10	178.38	59.83	29.34	64PHE	7.98	122.61	175.26	61.11	38.02
20ARG	7.53	120.29	178.66	58.89	29.54	65ASP	8.38	121.80	177.54	57.34	41.43
21PHE	8.54	121.90	175.64	60.91	39.47	66GLU	7.06	117.65	178.82	58.92	28.88
22LYS	8.26	119.10	177.71	59.13	32.06	67ALA	7.27	123.48	180.64	54.17	18.16
23GLU	7.18	120.28	177.83	58.14	28.76	68ALA	8.98	122.05	178.62	54.35	17.41
24LEU	7.57	122.84	178.01	57.36	41.18	69GLN	8.00	118.73	178.00	58.64	27.80
25LEU	7.72	119.15	177.88	57.23	41.48	70GLU	6.97	120.12	178.65	58.71	28.46
26ARG	6.98	118.78	178.12	59.06	29.16	71GLN	7.30	118.77	178.13	56.75	26.56
27GLU	7.22	118.48	180.01	58.48	28.90	72ILE	8.36	121.15	176.42	65.68	37.45
28GLU	8.39	119.25	180.40	58.16	27.99	73ALA	7.26	121.12	180.16	54.62	17.26
29VAL	8.72	123.98	177.06	65.87	30.25	74TRP	7.33	12014.	177.82	60.92	28.61
30ASN	7.39	121.86	177.01	55.81	37.15	75LEU	8.33	120.40	178.62	57.14	41.27
31SER	7.7	115.05	176.97	60.11	62.75	76LYS	8.10	117.47	179.11	60.20	31.60
32LEU	7.53	122.08	177.85	57.32	40.43	77GLU	6.92	121.66	177.16	57.73	28.05
33SER	7.91	114.24	177.81	61.62	62.98	78ARG	7.29	119.62	178.27	56.78	28.01
34ASN	8.14	120.12	176.57	55.65	37.55	79ILE	7.58	117.43	176.09	64.59	37.31
35HIS	7.66	120.88	177.42	56.98	31.19	80ARG	6.97	120.37	178.65	58.70	28.67
36PHE	8.30	118.35	177.25	60.29	38.81	81VAL	7.46	118.81	178.25	65.08	31.16
37HIS	8.10	117.32	175.14	58.61	28.82	82LEU	7.47	121.62	179.23	56.82	41.05
38ASN	7.17	117.82	174.60	52.64	38.95	83GLU	7.86	119.57	177.99	57.98	28.73
39LEU	6.69	122.49	176.99	54.38	41.16	84GLU	7.56	120.74	177.34	57.89	29.18

40GLU	9.44	127.70	178.16	57.21	30.84	85ASP	7.57	120.70	176.78	55.29	40.65
41SER	8.82	114.73	172.61	59.50	63.66	86TYR	7.48	120.70	176.53	58.78	37.75
42TRP	6.19	122.55	173.63	52.78	29.22	87LEU	7.45	122.18	177.49	55.63	41.09
43ARG	8.48	129.07	173.50	53.94	30.61	88GLU	7.51	120.12	176.42	56.58	29.14
44ASP	6.43	120.80	175.26	52.87	41.14	89HIS	7.46	118.12	174.57	55.74	29.76
45ALA			180.43	54.09	17.80	90HIS	7.39	126.99	180.25	56.64	29.77

*ND; Not Detected **Unit; ppm

Acknowledgements

This work is financially supported by Ministry of Education, Science and Technology (MEST); New Drug Target Discovery; Grant number, 370C-20070095, Innovative Drug Research Center for Metabolic and Inflammatory Disease and 2008 BK21 Project for Medicine, Dentistry and Pharmacy. We thank NCIRF and KBSI for using NMR machines.

REFERENCES

1. Rothenbacher D., Brenner H., *Microbes Infect.* **5(8),** 693-703 (2003)
2. Tomb J. F., White O., Kerlavage A. R., Clayton R. A., Sutton G. G., Fleischmann R. D., Ketchum K. A., Klenk H. P., Gill S., Dougherty B. A., Nelson K., Quackenbush J., Zhou L., Kirkness E. F., Peterson S., Loftus B., Richardson D., Dodson R., Khalak H. G., Glodek A., McKenney K., Fitzegerald L. M., Lee N., Adams M. D., Hickey E. K., Berg D. E., Gocayne J. D., Utterback T. R., Peterson J. D., Kelley J. M., Cotton M. D., Weidman J. M., Fujii C., Bowman C., Watthey L., Wallin E., Hayes W. S., Borodovsky M., Karp P. D., Smith H. O., Fraser C. M., Venter J. C., *Nature* **388(6642),** 539-547 (1997)
3. Suerbaum S., Josenhans C., *Nat Rev Microbiol.* **5(6),** 441-452 (2007)
4. Kim W.J., Lim J.S., Son W.S., Ahn H.C., Lee B.J., *JKMRS.* **12,** 65-73 (2008)
5. Jang S.B., Kwon A.R., Son W.S., Park S.J., Lee B.J., *J. Biochem.* **146(4),** 535-540 (2009)
6. Pallen M.J. T., *Trends Microbiol.* **10(5),** 209-212 (2002)

7. Hong Y.H., Park S.H., Lee K.L., Lee B.J., *JKMRS.* **5,** 46-55 (2001)

8.Delaglio F., Grzesiek S., Vuister G.W., Zhu G., Pfeifer J., Bax A., *J Biomol NMR.* **6(3)**, 277-293 (1995)

9. Johnson B. A., *Methods Mol Biol.* **278**, 313-352 (2004)

10. Wishart D. S., Sykes B. D., *J Biomol NMR.* **4(2)**, 171-180 (1994)

11.Cornilescu G., Delaglio F., Bax A., *J Biomol NMR.* **13(3)**, 289-302 (1999)

12.Metzler W. J., Constantine K. L., Friedrichs M. S., Bell A. J., Ernst E. G., Lavoie T. B., Mueller L., *Biochem.* **32(50)**, 13818-13829 (1993)

13.Wishart D. S., Bigam C. G., Holm A., Hodges R. S., Sykes B. D., *J Biomol NMR.* **5(1)**, 67-81 (1995)

14.Spera S., Bax A., *J. Am. Chem. Soc.* **113(14)**, 5490-5492 (1991)