Journal of Biochemistry and Molecular Biology, Vol. 38, No. 5, September 2005, pp. 591-594

JOURNAL OF Biochemistry and Molecular Biology

Backbone ^{1}H, ^{15}N, and ^{13}C Resonance Assignment of HP1242 from *Helicobacter pylori*

Su-Jin Kang, Sung Jean Park, Seo-Jeong Jung and Bong-Jin Lee*

National Lab. of Membrane Protein Structure (MPS), Research Institute of Pharmaceutical Sciences, College of Pharmacy, Seoul National University, Seoul 151-742, Korea

Received 30 May 2005, Accepted 14 July 2005

One of the small proteins from *Helicobacter pylori*, HP1242, was investigated by the solution nuclear magnetic resonance (NMR) spectroscopy. HP1242 is known as a 76-residue conserved hypothetical protein and its function cannot be identified based on sequence homology. Here, the results of the backbone ^{1}H, ^{15}N, and ^{13}C resonance assignments of the HP1242 are reported using double- and triple-resonance techniques. About 95% of all of the ^{1}HN, ^{15}N, ^{13}CO, $^{13}C\alpha$, and $^{13}C\beta$ resonances that cover 75 non-Proline residues of the 76 residues are clarified through sequential- and specific- assignments. In addition, three helical regions were clearly identified on the basis of the resonance assignments.

Keywords: Coiled-coil, *Helicobacter pylori*, HP1242, NMR, Unknown protein

Introduction

Helicobacter pylori is a gram-negative and spiral shaped bacteria which lives in stomach, and is related with many serious gastric problems, ranging from gastritis to gastric carcinoma or lymphoma (Blaser, 1990; Forman *et al*., 1991; Parsonnet *et al*., 1994). It has a unique way of adapting in the harsh, acidic environment of human stomach and chronically colonizing the epithelium of the stomach. The genome of *H. pylori* has been fully sequenced for two prototype strains (strain 26695 and strain J99) (Jean-F. *et al*., 1997). 1,590 open reading frames (ORFs) were identified in the chromosome of strain 26695. Among the 1,590 ORFs, 499 ORFs have no homologues in other organisms and more have no putative function. Determining 3D structure of these unknown proteins can lead to the inference of the biological function of those and identification of new drug target (Park *et al*., 2003). As a part of our structural genomics on *Helicobacter pylori*, we studied the solution structure of HP1242, one of proteins from *Helicobacter pylori* by using NMR.

The HP1242 gene of *Helicobacter pylori* encodes a 76-residue conserved hypothetical protein from *Helicobacter pylori* strain 26695 with a molecular weight of 9,111Da and a calculated isoelectric point of 6.1. Based on the sequence homology, this protein is classified as the DUF (Domain of Unknown function) 465 family (pfam, http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cig), which has unknown function. These family members are found in several bacterial proteins, and also in the heavy chain of eukaryotic myosin and kinesin, which are predicted to form coiled-coil structures (Jung and Lee, 2004).

Here, we report the sequence-specific backbone resonance assignments of HP1242. This result will facilitate the structure determination of HP1242.

Materials and Methods

For the expression of HP1242, pET-21a was used as a vector system (Novagen, Darmstadt, Germany). The gene product has 6 additional Histidines at the C-terminal to facilitate purification by Ni^{2+}-agarose column (His-bind Resin). The ORF of HP1242 was amplified by PCR with appropriate primers and inserted between NdeI and XhoI cleavage sites. The constructed plasmid was transformed into competent cell, *E. coli* BL21 (DE3) codon+. For NMR experiments, ^{15}N and ^{13}C-uniformly labeled protein was prepared by growing the cells in the isotope-supplemented M9 medium at 37°C. The soluble protein was purified using Ni^{2+}-agarose column (His-bind Resin) and DEAE-sepharose column (Amersham Pharmacia Biotech. Inc., Uppsala, Sweden). The NMR sample containing 0.7 mM of the HP1242 was prepared in 90% H_2O/10% D_2O containing 50 mM NaH_2PO_4/Na_2HPO_4

*To whom correspondence should be addressed.
Tel: 82-2-880-7869; Fax: 82-2-872-3632
E-mail: lbj@nmr.snu.ac.kr

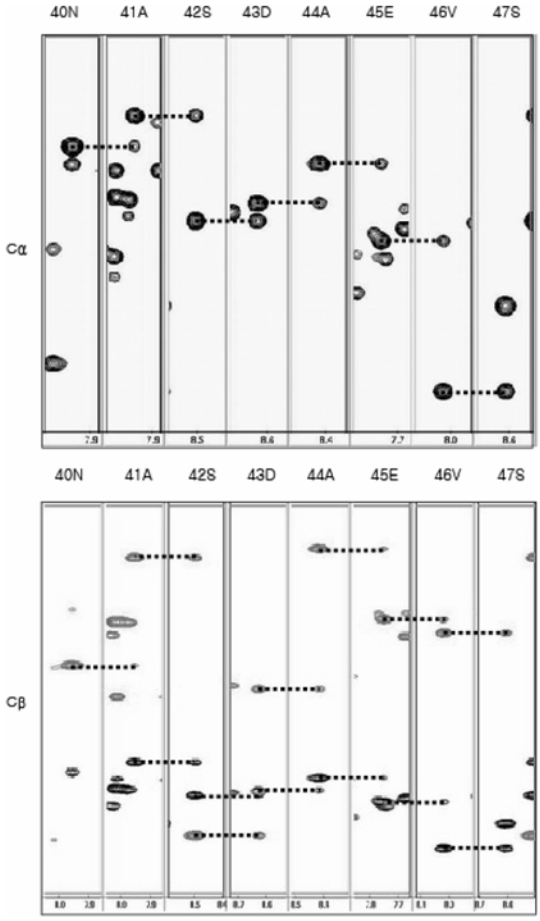

Fig. 1. Strip plot of HP1242 (residues 40-47) (A) Sequential connectivity of Cα carbons is depicted in HNCA spectrum. The dashed lines represent the intermolecular connectivity. (B) Sequential connectivity of Cβ carbons in HNCACB spectrum.

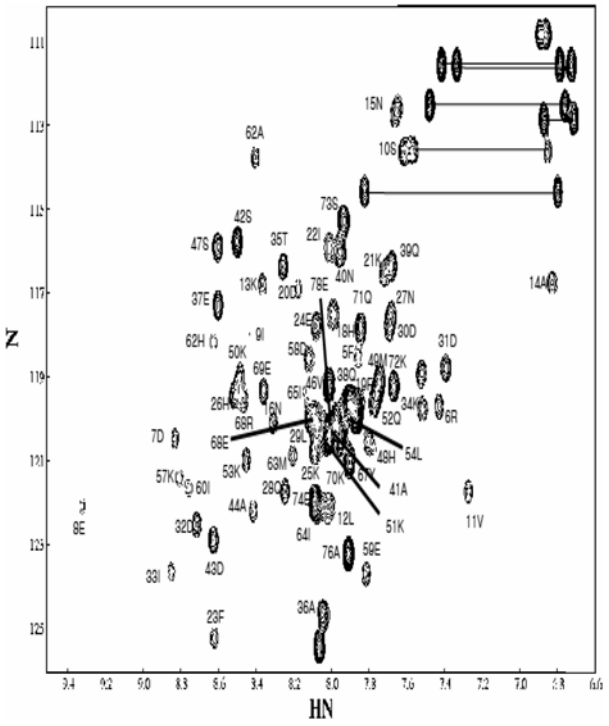

Fig. 3. 2D ^{1}H and ^{15}N HSQC spectrum of HP1242. The each resonance in the spectrum is labeled with the assigned amino acid residues. The dashed lines represent the sidechains of Gln and Asn.

(pH 6.8), 100 mM NaCl and 1 mM EDTA.

All NMR measurements were performed at 303 K on Bruker Avance-600 MHz NMR spectrometer. The experiments recorded on ^{15}N-labelled sample were 2D ^{15}N-HSQC, 3D ^{15}N-TOCSY-HSQC and 3D ^{15}N-HNHA. The experiments recorded on ^{15}N/^{13}C-labelled sample were HNCA, HNCO, HN(CO)CA, HNCACB, CBCA(CO)NH, HNCACO, HCCH-TOCSY, and CCCONH (Bodenhausen and Ruben, 1980; Wittekind and Mueller, 1993; Yamazaki *et al*., 1994; Reid, 1997).

Proton chemical shift was referenced to the methyl signal of 2, 2-dimethylsilapentane-sulfonic acid (DSS) externally. The ^{15}N and ^{13}C chemical shifts were referenced indirectly to DSS (Wishart *et al*., 1995). All spectra were processed using the nmrPipe/nmrDraw software (Delaglio *et al*., 1995), and were analyzed using the program NMRView (Johnson and Blevins, 1994). The sequence-specific resonance assignment was carried out using standard procedures (Wüthrich, 1986; Arseniev *et al*., 1988; Gronenborn *et al*., 1989).

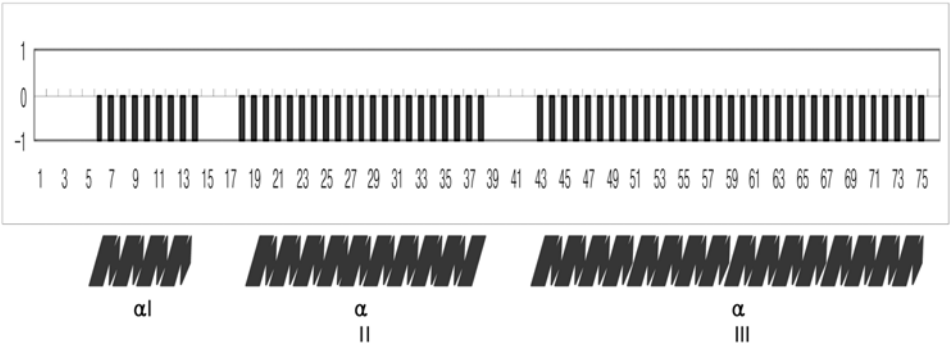

Fig. 2. CSI plot of HP1242 based on resonance assignments. In the consensus CSI, the values '1' and '-1' indicate respectively and α-helix and β-sheet tendancy.

Table 1. Chemical shifts of ^{1}HN, ^{15}N, ^{13}CO, ^{13}Cα and ^{13}Cβ of HP1242. All chemical shifts were referenced to the frequency of the methyl proton resonance of DSS.

Reesidue	HN	N	CO	CA	CB
Met1	*ND	ND	ND	ND	ND
Phe2	ND	ND	ND	ND	ND
His3	ND	ND	ND	56.669	27.315
Glu4	**9.208	119.910	172.730	55.000	26.000
Phe5	7.870	118.423	170.648	52.886	36.710
Arg6	7.438	119.629	176.823	57.432	27.235
Asp7	8.832	120.372	176.159	54.868	37.342
Glu8	9.318	121.970	175.666	58.651	25.925
Ile9	8.431	117.892	173.705	62.967	35.185
Ser10	7.623	113.544	174.065	59.026	59.953
Val11	7.275	121.676	176.296	63.416	29.107
Leu12	8.012	122.034	177.735	55.248	38.162
Lys13	8.367	116.737	175.257	57.894	29.938
Ala14	6.835	116.672	178.255	50.945	16.274
Asn15	7.665	112.707	169.391‘	50.995	38.111
Asn16	8.311	120.014	184.047	47.869	36.968
Pro17			177.873	62.198	29.605
His18	7.991	117.396	172.544	56.932	26.995
Phe19	7.858	119.479	172.518	58.930	36.131
Asp20	8.184	116.842	175.250	55.506	39.578
Lys21	7.723	116.470	180.398	56.742	29.472
Ile22	8.017	115.831	174.255	62.499	34.835
Phe23	8.625	125.172	175.369	59.357	36.703
Glu24	8.089	117.709	178.924	56.447	26.446
Lys25	8.090	120.631	176.331	56.286	29.195
His26	8.514	119.373	171.850	58.810	25.989
Asn27	7.685	117.391	174.546	52.876	34.423
Gln28	8.250	121.684	175.839	55.797	25.622
Leu29	8.125	119.899	176.587	55.232	39.514
Asp30	7.692	117.790	176.522	55.249	39.020
Asp31	7.390	118.723	176.874	54.938	37.996
Asp32	8.720	122.462	179.449	54.891	37.671
Ile33	8.858	123.577	173.610	63.467	35.714
Lys34	7.521	119.692	180.128	57.085	29.254
Thr35	8.264	116.286	180.150	63.676	65.992
Ala36	8.050	124.622	180.372	52.650	15.498
Glu37	8.601	117.213	180.837	56.855	26.832
Gln38	7.902	119.562	174.475	55.382	25.777
Gln39	7.683	116.259	170.719	52.592	25.438
Asn40	7.955	115.933	169.103	51.590	34.390
Ala41	7.954	120.381	174.490	50.042	17.027
Ser42	8.497	115.697	169.336	55.329	61.822
Asp43	8.626	122.856	177.204	54.570	38.152
Ala44	8.415	122.134	180.349	52.541	15.574
Glu45	7.751	119.209	178.251	56.398	26.969
Val46	8.017	119.122	176.180	63.957	29.165
Ser47	8.605	115.839	173.902	59.885	ND
His48	7.767	120.797	176.090	57.117	27.425
Met49	7.740	118.897	176.772	57.356	32.076
Lys50	8.485	118.893	178.645	58.228	29.862
Lys51	8.028	120.445	178.762	57.172	29.500
Gln52	7.770	119.587	176.548	56.083	25.983
Lys53	8.451	120.888	175.354	58.128	29.721
Leu54	7.898	119.666	ND	55.452	39.065
Lys55	7.566	118.834	ND	56.710	29.650
Leu56	ND	ND	176.234	55.188	39.590
Lys57	8.814	121.352	176.389	58.071	29.017
Asp58	8.127	118.503	177.667	54.850	37.644
Glu59	7.810	123.614	177.753	57.209	26.784
Ile60	8.763	121.638	175.462	63.382	35.694
His61	8.635	118.095	174.189	57.702	27.554
Ser62	8.402	113.691	174.202	59.560	ND
Met63	8.213	120.836	178.190	57.673	30.761
Ile64	8.089	122.178	174.428	63.514	35.188
Ile65	8.091	119.960	176.885	61.146	33.566
Glu66	8.075	120.182	177.124	56.868	26.931
Tyr67	7.801	120.515	175.040	58.923	36.060
Arg68	8.465	119.523	179.147	56.991	27.809
Glu69	8.364	119.272	177.955	56.360	26.475
Lys70	7.905	120.977	177.086	56.152	29.334
Gln71	7.843	117.760	175.288	54.750	25.954
Lys72	7.670	119.168	176.089	55.799	29.770
Ser73	7.937	115.217	170.974	57.402	60.624
Glu74	8.093	121.904	174.954	155.056	27.024
Arg75	7.981	119.990	173.410	54.375	27.644
Ala76	7.909	123.168	176.030	50.461	16.045

*ND; not detected
** unit; ppm

Result and Discussion

Nearly complete assignments of HP1242 were achieved. About 95% of ^{1}HN and ^{15}N of the backbone amides of HP1242 (excluding the one Proline residue), and ^{13}CO, ^{13}Cα, and ^{13}Cβ resonances of HP1242 were assigned. Three residues are not observed in a 2D ^{1}H-^{15}N HSQC spectrum; therefore, they could not be assigned (residues 1, 2 and 3). Missing peaks are presumed to be from residues in the intermediate conformational exchange and consequently broadened beyond detection. In addition, residue 56 also could not be assigned, because of overlapping with some peaks. HP1242 has three Leucine residues and six Isoleucine residues, and their chemical shifts are so similar that peaks in the HSQC spectrum are extremely overlapped. We could identify two separated Leucine residues and all six Isoleucine residues, but could not detect one Leucine residue 56.

The sequential connectivity of ^{13}Cα and ^{13}Cβ carbons in the third helical region is shown in Fig. 1. On the basis of resonance assignments, three helical regions were clearly

identified using the CSI program as shown in Fig. 2 (Wishart *et al*., 1995). These correspond to residues R6-A14 (αI), H18-Q38 (αII), and D43-R75 (αIII). The chemical shift values of the ^{1}HN, ^{15}N, ^{13}CO, ^{13}Cα, and ^{13}Cβ resonances are represented in Table 1. Fig. 3 shows the typical ^{1}H-^{15}N HSQC spectrum that is labeled with the assigned amino acid residue. These data will be the first step toward the investigation of 3D structure of HP1242.

Acknowledgments This work was supported by a grant (03-PJ2-PG4-BD02-0001) from ministry of Health & Welfare, R.O.K. This work was also supported in part by National Research Laboratory Program (M1-0412-00-0075) and 2005 BK21 project for Medicine, Dentistry, and Pharmacy.

References

Arseniev, A., Schultze, P., Worgotter, E., Braun, W., Wagner, G., Vasak, M., Kagi, J. H. and Wüthrich, K. (1988) Three-dimensional structure of rabbit liver [Cd7]metallothionein-2a in aqueous solution determined by nuclear magnetic resonance. *J. Mol. Biol.* **201**, 637-657.

Blaser, M. J. (1990) *Helicobacter pylori* and the pathogenesis of gastroduodenal inflammation. *J. Infect. Dis.* **161**, 626-633.

Bodenhausen, G. and Ruben, D. J. (1980) Heteronuclear 2D correlation spectra with doublein-phase transfer steps. *Chem. Phys. Letters* **69**, 185-189.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293.

Forman, D., Newell, D. G., Fullerton, F., Yarnell, J. W., Stacey, A. R., Wald, N. and Sitas, F. (1991) Association between infection with *Helicobacter pylori* and risk of gastric cancer: evidence from a prospective investigation. *Br. Med. J.* **302**, 1302-1305.

Graham, D. Y. and Go, M. F. (1993) *Helicobacter pylori*: current status. *Gastroenterology* **105**, 279-282.

Gronenborn, A. M., Bax, A., Wingfield, P. T. and Clore, M. (1989) A powerful method of sequential proton resonance assignment in proteins using relayed ^{15}N-^{1}H multiple quantum coherence spectroscopy. *FEBS Lett.* **243**, 93-98.

International Agency for Research on Cancer, World Health Organization (1994) Infection with *Helicobacter pylori*. In: Schistosomes, Liver Flukes and *Helicobacter pylori*, *IARC, Lyon. Monogr. Eval. Carcinog. Risks Hum*. **60**, 177-240.

Jean-F., T., Owen, W., Anthony, R. K., Rebecca, A. C., Granger G. S., Robert, D. F., Karen, A. K., Hans, P. K., Steven, G., Brian, A. D., Karen, N., John, Q., Lixin, Z., Ewen, F. K., Anna, G., Keith, M., Lisa, M. F., Norman, L., Mark, D. A., Erin, K. H., Douglas, E. B., Jeanine, M. W., Claire, F., Cheryl, B., Larry, W., Erik, W., William, S. H., Mark, B., Peter, D. K., Hamilton, O. S., Claire, M. F. and Craig, V. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**, 539-547.

Johnson, B. A. and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**, 603-614.

Jung, J. W. and Lee, W. T. (2004) Structure-based functional discovery of proteins: Structural proteomics. *J. Biochem. Mol. Biol.* **37**, 28-34.

Park, S. J., Kim, J. S., Son, W. S., Ahn, H. C. and Lee, B. J. (2003) Backbone ^{1}H, ^{15}N, and ^{13}C resonance assignments of the helicobacter pylori acyl carrier protein. *J. Biochem. Mol. Biol.* **36**, 505-507.

Parsonnet, J., Hansen, S., Rodriguez, L., Gelb, A. B., Warnke, R. A., Jellum, E., Orentreich, N., Vogelman, J. H. and Friedman, G. D. (1994) *Helicobacter pylori* infection and gastric lymphoma. *N. Engl. J. Med.* **330**, 1267-1271.

Reid, D. G. (1997) Introduction to the NMR proteins; in *Protein NMR Techniques*, Reid, D. G. (ed.), pp. 1-28, Humana Press, New Jersey, USA.

Wishart, D. S. and Sykes, B. D. (1994) The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR*. **4**, 171-180.

Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**, 135-140.

Wittekind, M. and Mueller, L. (1993) HNCACB, a highsensitivity 3D NMR experiment to correlate amideproton and nitrogen resonances with the α-carbon and â-carbon resonances in proteins. *J. Magn. Reson. ser.* **101**, 214-217.

Wüthrich, K. (1986) Three-dimensional protein structure by NMR in *NMR of Proteins and Nucleic Acids*. Wüthrich, K. (ed.), pp. 176-199, John Wiley & Sons, New York, USA.

Yamazaki, T., Lee, W., Revington, M., Mattiello, D. L., Dahlquist, F. W., Arrowsmith, C. H. and Kay, L. E. (1994) An HNCA pulse scheme for the backbone assignment of ^{15}N, ^{13}C, ^{2}H labeled proteins: application to a 37-kDa trp repressor-DNA complex. *J. Am. Chem. Soc.* **116**, 6464-6465.