Journal of the Korean Magnetic Resonance Society **2011**, *69-79*

DOI : 10.6564/JKMRS.2011.15.1.069

^{1}H, ^{15}N and ^{13}C resonance assignment and secondary structure prediction of ss-DNA binding protein 12RNP2 precursor, HP0827 from *Helicobacter pylori*

Sun-Bok Jang, Chao Ma, Pathak Chinar Chandan, Do-Hee Kim and Bong-Jin Lee*

Research Institute of Pharmaceutical Sciences, College of Pharmacy, Seoul National University, Kwanak-Ro , Kwanak-Gu, Seoul 151-742, Korea

(Received May 04, 2011; accepted May 31, 2011)

Abstract : HP0827 has two RNP motif which is a very common protein domain involved in recognition of a wide range of ssRNA/DNA.We acquired 3D NMR spectra of HP0827 which shows well dispersed and homogeneous signals which allows us to assign 98% of all $^{1}H_{N}$, ^{15}N, $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and $^{13}C{=}O$ resonances and 90% of all sidechain resonances. The sequence-specific backbone resonance assignment of HP0827 can be used to gain deeper insights into the nucleic acids binding specificity of HP0827 in the future study.
Here, we report secondary structure prediction of HP0827 derived from NMR data. Additionally, ssRNA/DNA binding assay studies was also conducted. This study might provide a clue for exact function of HP0827 based on structure and sequence.

Keywords : HP0827, NMR, *Helicobacter pylori*, RRM, RNP

INTRODUCTION

Helicobacter pylori is one of the most common bacterial pathogens in humans and leading cause of the considerable diseases including peptic ulcer and gastric cancer worldwide. The link between

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

helicobacter pylori and ulcerartion as well as its association with gastric cancer is a serious public health concern.[1] Recently, studies on the molecular genetics of *H. pylori* have focused on finding the exact biological functions of each proteins belonging to *helicobacter pylori* genome to explain the relevance between *H. pylori* and diseases. Because of high prevalence, medical importance and unique genetic phenomena of this bacterial pathogen, there are many researches to find exact function of *H. pylori.*

Approximately 29% of the loci are in the "pathogenesis" category of the genome database.[2] Studies on *H. pylori* genomes have focused on attempt to understand pathogenesis, the ability of *H. pylori* to cause disease.The availability of the complete genome sequences of *H. pylori* could make it possible to do a research from the basic biology of *H. pylori* to the structure of protein. Protein structure is likely to be helpful to explain its functions, because the function of protein is almost related to its structure.[3] As part of our structural genomics effort on *Helicobacter pylori*[4], we could characterize the structural information of HP0827 belonging to the RRM superfamily possessing two RNP motif in the β-sheet region. The HP0827 has 82 amino acids which is identified as ss-DNA binding protein 12RNP2 precursor from *Helicobacter pylori* strain 26695 with calculated pI value of 8.96 and molecular weight of 9.4kDa.

The RRM or ribonucleoprotein (RNP) domain is by far the best-characterized DNA/RNA binding domain and is also most widespread domain which shares the almost same topology and three-dimenstional structure (four-stranded β sheet packed against two α-helices). The highly

conserved aromatic residues within RNP-1 and RNP-2 provide similar stacking interaction forces to RRM proteins.[5] HP0827 also has two RNP motif (RNP1 and RNP2) in the sequence.

In this paper, we report the sequence-specific resonance assignment and the secondary structure of HP0827.

EXPERIMENTAL

Cloning, Expression and NMR sample preparation

The HP0827 gene was amplified using genomic DNA of Helicobacter pylori 26695 and the PCR product was subcloned into the expression vector pET-21a(+) to generate C-terminal His^6-tagged fusion protein. Then, the resulting plasmid was transformed in to *Escherichia coli* strain BL21 (DE3) host cells for large-scale protein production. Uniform labeling of the protein was accomplished by cell growth at 37℃ in M9 media that contains $^{15}NH_4Cl$ and $^{13}C_6$-glucose as the sole nitrogen and carbon sources, respectively. The expressed His-tagged protein was initially purified using Ni^{2+} chelating column, and was further purified using DEAE-Sepharose column. Finally, single protein was further purified by superdex-75 column for high purity in a buffer containing 20mM MES, pH 6.5, 100mM NaCl, 1mM DTT and 0.1mM PMSF. The purity of finally purified HP0827 protein was confirmed by SDS-PAGE. Protein was concentrated up to 1mM for NMR experiments.

NMR spectroscopy

NMR spectra were collected at 35°C using Bruker AVANCE 500, 600 and 900 Mhz NMR spectrometer. 1mM of ^{13}C, ^{15}N-labled HP0827 sample was prepared in a final buffer with 9:1 H_2O:D_2O for locking. The backbone and sidechain assignments were made by three-dimensional heteronuclear correlation experiments: HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HN(CA)CO, HNCO, HBHA(CO)NH, ^{15}N-TOCSY-HSQC and HCCH-TOCSY spectra. Chemical shifts were referenced to DSS externally. NMR data were processed with the NMRPipe/nmrDraw[6] and analyzed with the NMRView[7] and SPARKY program.[8] The secondary structure was predicted from the values using Chemical Shift Index (CSI)[9] and the backbone torsion angles were obtained using TALOS program[10] 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. In addition, we used chemical shift difference method between measured values and random-coil values using C_{α}, C_{β} and ($\Delta C_{\alpha} - \Delta C_{\beta}$) for predicting secondary structure of HP0827.

Protein-Nucleic acids binding experiments

The gel electrophoresis mobility assay (EMSA) was performed using a 6-nucleotide ssDNA (5′-ATTGCA-3′) and ssRNA (5′-AUUGCA-3′) to detect protein complexes with nucleic acids.[11] The protein sample and the ssDNA/RNA stock solution were all prepared in the same buffer. Agarose gel

(2.0%) was used to show the band shifting of protein-nucleic acids complexes. The assays was performed with small protein and ssDNA/ssRNA in concentrations (0.2 μM) and small (10 μl) sample volume.

RESULTS AND DISCUSSION

^{1}H-^{15}N TROSY spectrum of ^{15}N-enriched HP0827 showed good signal dispersion (Fig. 1). Sequence specific assignments for backbone were performed. The assignments comprise 98 % of all $^{1}H_{N}$, ^{15}N, $^{13}C_{\alpha,}$ $^{13}C_{\beta}$ and $^{13}C=O$.

Secondary structural informations were identified on the basis of chemical shifts which are the most accessible and easily measured quantities in all of NMR spectroscopy. The chemical shift has long been recognized as being a sensitive measure of molecular conformation, backbone dihedral angle, hydrogen bond interactions, backbone dynamics and ring-flip rates.[12] The Chemical Shift Index (CSI) has become a standard method for determining secondary structure of protein by characterizing deviation of chemical shifts of certain nuclei in amino acid relative to their random coil values.[9] In addition, TALOS program was also used to obtain backbone dihedral amgles (Φ□Ψ) information using chemical shift values.[10]

Analysis of the chemical shift difference of $13C_{\alpha}$ and $13C_{\beta}$ between measured values and random-coil values shows the correlation with result of CSI and TALOS program. As shown in Fig. 2, examination of ($\Delta C_{\alpha}-\Delta C_{\beta}$) plot indicates the presence of potentially 2 α-helices and 4 β-sheets . The region of α-helices and β-sheets correspond well to the CSI and TALOS predictions.

RRM family proteins are well known to interact with ssDNA/RNA using RNP motif and positive charged residues.[13] We checked nucleic acids binding affinity of HP0827 using EMSA (Electrophoresis Mobility Shift Assay) method. EMSA is a rapid and sensitive method to detect protein-nucleic acid interaction.[14] In general, protein-nucleic acid complexes migrate more slowly than the corresponding free nucleic acids. So, we conducted the binding affinity experiment between nucleic acids and HP0827 by using EMSA.

Continuous variation electrophoresis mobility shift assay result supports the HP0827-ssDNA/RNA complex (Fig. 3). In futher study, ITC experiment and NMR HSQC titration experiment can be used to verify exact nucleic acids sequence , binding sites and modes. Because, all backbone amide ($^{1}H_{N}$ and ^{15}N) resonances were assigned, TROSY spectrum of HP0827 (Fig. 1) can be used to find the exact binding site of protein with nucleic acid.

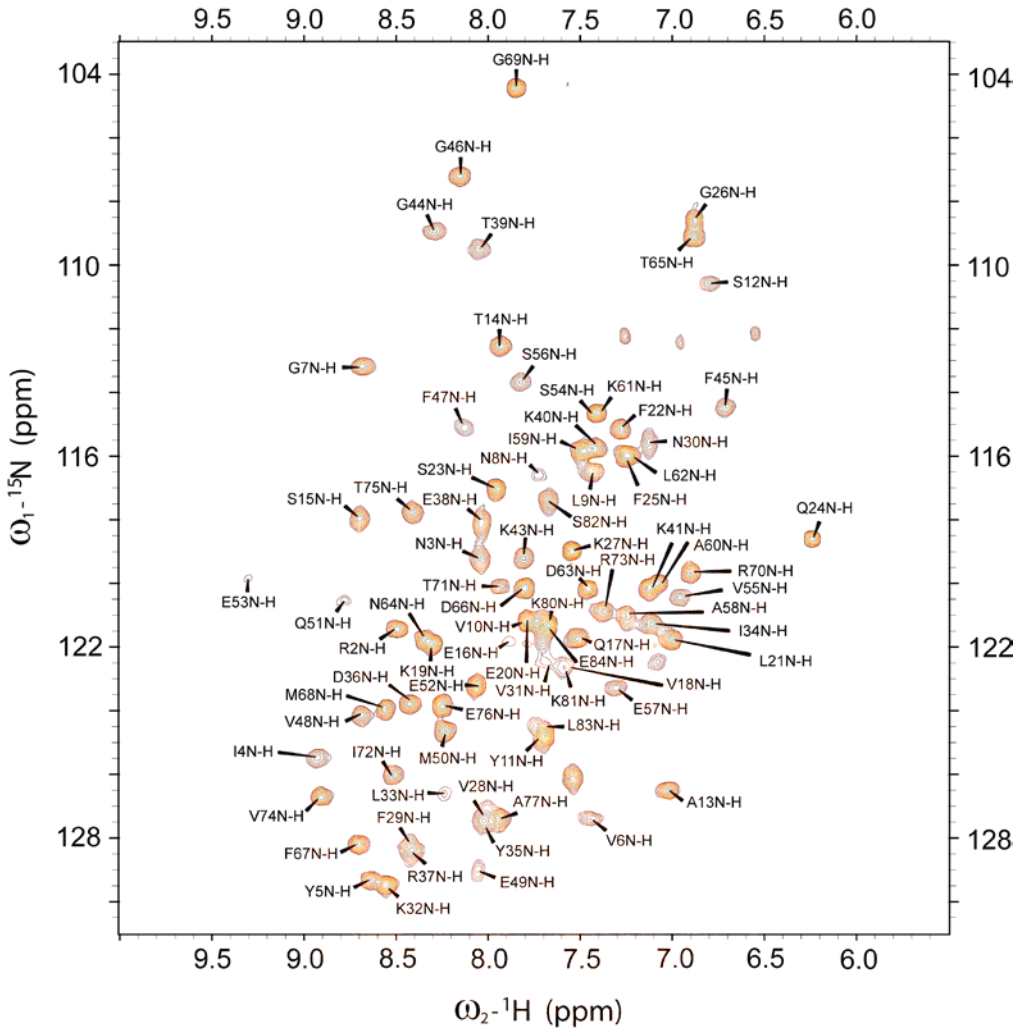

Figure.1. 2D ^{1}H and ^{15}N TROSY spectrum of HP0827. The each resonance in the spectrum is labeled with assigned amino acid residues.

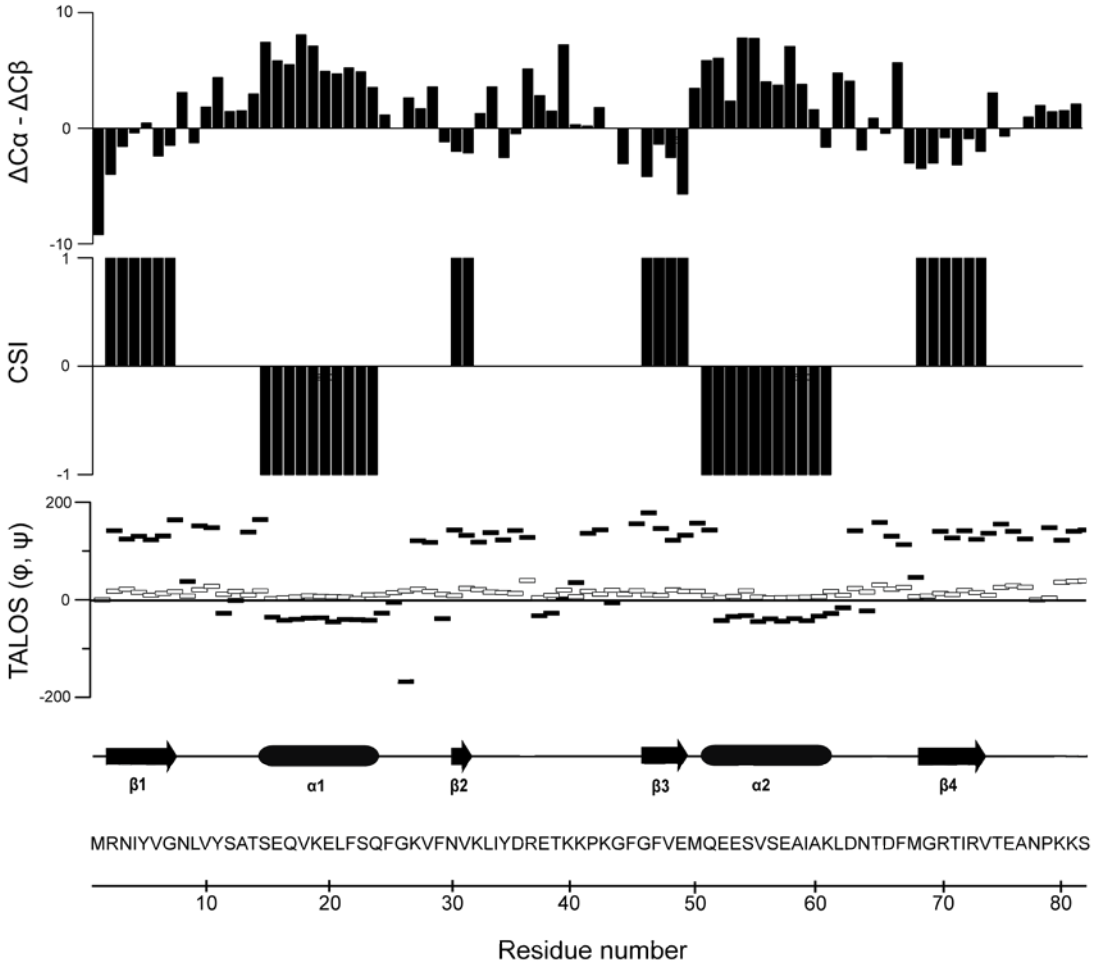

Figure.2. Summary of backbone resonance assignment of HP0827. Delta values ($\Delta C_\alpha - \Delta C_\beta$) of backbone carbon to random coil chemical shifts were plotted. In the consensus CSI, the values '-1' represents the α-helical tendency while '1' represents the opposite pattern (β-strand). Backbone dihedral angels (phi, psi) were calculated using TALOS. Filled and open rectangles indicated the predicted phi (Φ) and psi (Ψ) angles, respectively. Predicted secondary structure elements of HP0827 were shown in black boxes for α-helices and arrows for β-strand.

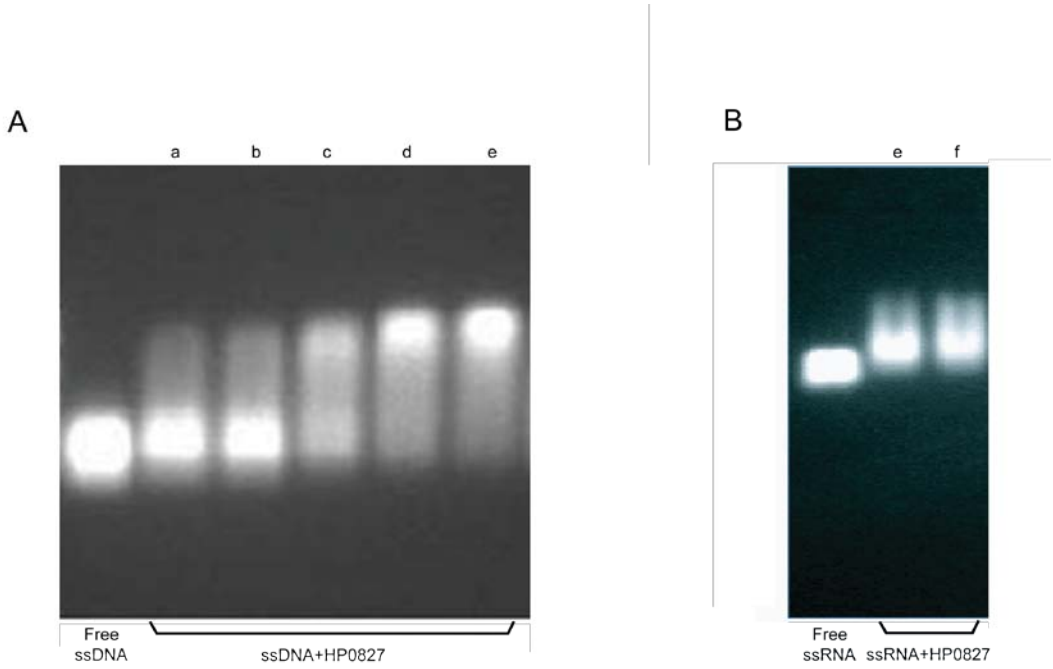

Figure.3. Titration of 6-bp ssDNA/RNA fragment with HP0827 protein.

Electrophoresis were carried out with 2.0 % agarose gel. All samples contained 0.2 μM A:ssDNA/B: ssRNA, respectively. The binding buffer is pH 6.5, 20mM MES, 100mM NaCl, 1mM DTT, EDTA and NaN_3.

A. As increasing the concentration of protein, the binding quantity of ssDNA to protein was increased. Sample a-e contained in 0.2, 0.4, 0.6, 0.8 and 1.0 μM protein, respectively.

B. The migration of ssRNA band was observed when HP0827 was added to ssRNA. Sample e-f contained 0.2 and 1.0 μM protein, respectively.

Table.1. Chemical shifts of $^{1}H_{N}$, ^{15}N, ^{13}CO, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ of HP0827. 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	176.23	53.92	42.20	43LYS	7.79	119.21	55.59	55.59	32.20
2ARG	8.49	121.50	173.34	53.86	34.05	44GLY	8.42	109.40	172.30	45.61	
3ASN	8.02	119.35	174.27	52.23	41.20	45PHE	6.74	114.61	172.45	54.64	41.20
4ILE	8.90	125.33	174.89	60.45	40.84	46GLY	8.15	107.20	168.71	44.54	
5TYR	8.65	129.33	172.38	56.83	38.88	47PHE	8.14	115.45	174.80	55.57	43.24
6VAL	7.51	127.53	172.47	58.12	33.31	48VAL	8.70	123.99	173.51	60.46	34.64
7GLY	8.72	113.28	172.84	43.03		49GLU	8.04	128.99	173.37	54.93	32.45
8THR	7.78	116.82	172.82	53.26	37.58	50MET	8.21	124.63	173.71	54.45	39.23
9LEU	7.46	116.20	178.77	53.38	43.23	51GLN	8.80	120.67	178.16	56.89	28.56
10VAL	7.85	121.81	177.33	61.35	32.34	52GLU	8.06	123.24	177.11	59.84	29.60
11TYR	7.84	125.05	175.86	58.89	37.02	53GLU	9.26	119.71	177.60	59.43	28.37
12SER	6.86	110.83	174.12	57.42	62.90	54SER	7.43	115.96	173.49	60.11	64.68
13ALA	7.04	126.54	177.13	52.31	18.41	55VAL	6.94	120.37	176.36	66.52	31.52
14THR	7.98	112.81	175.62	58.64	71.61	56SER	7.79	113.79	176.60	62.68	61.83
15SER	8.71	118.02	175.70	62.37	61.87	57GLU	7.32	123.45	176.19	58.28	29.26
16GLU	7.86	121.82	178.94	59.60	28.88	58ALA	7.08	121.44	178.63	53.93	17.82
17GLN	7.38	120.95	179.62	58.61	28.22	59ILE	7.47	115.91	177.64	65.36	38.29
18VAL	7.52	121.88	176.78	66.50	31.23	60ALA	7.09	120.07	180.38	54.32	18.13
19LYS	8.31	122.08	178.88	60.51	31.79	61LYS	7.42	114.76	177.49	56.23	33.01
20GLU	7.79	121.25	178.32	59.21	29.30	62LEU	7.22	116.14	178.59	54.07	44.31
21LEU	6.98	121.89	178.32	57.72	41.62	63ASP	7.47	120.24	176.57	57.22	41.16
22PHE	7.27	115.34	176.98	61.47	39.75	64ASN	8.35	121.81	174.51	54.38	37.72
23SER	7.96	117.08	177.04	60.77	62.80	65THR	6.86	109.05	172.24	59.95	71.82
24GLN	6.22	118.64	175.66	57.04	28.60	66ASP	7.81	120.30	176.42	54.14	41.98
25PHE	7.25	116.11	174.24	57.20	39.56	67PHE	8.72	128.33	173.95	56.63	40.56
26GLY	6.86	108.93	169.79	43.78		68MET	8.57	124.04	174.25	56.39	29.83
27LYS	7.54	119.06	175.74	56.70	32.49	69GLY	7.83	104.59	174.18	45.09	
28VAL	8.01	127.39	175.59	61.26	32.38	70ARG	6.87	119.67	174.29	54.12	33.80
29PHE	8.41	128.28	175.25	59.38	39.31	71THR	7.99	120.23	175.07	62.02	68.30
30ASN	7.12	115.87	172.61	53.33	41.90	72ILE	8.54	126.40	174.58	61.11	40.29
31VAL	7.62	122.40	173.46	61.52	34.52	73ARG	7.23	120.97	174.54	51.84	31.22
32LYS	8.56	129.59	174.19	54.23	34.78	74VAL	8.92	126.80	174.03	60.61	34.31
33LYS	8.22	126.63	175.22	53.72	41.06	75THR	8.45	117.79	172.59	59.38	71.38
34ILE	7.04	121.73	174.76	58.30	34.72	76GLU	8.27	23.83	176.45	57.84	29.79
35TYR	8.01	127.39	174.76	55.73	40.76	77ALA	7.90	127.39	176.18	51.27	19.55
36ASP	8.43	123.87	177.34	54.30	43.46	78ASN	ND	ND	ND	ND	ND
37ARG	8.42	128.27	176.50	58.77	29.85	79PRO			176.96	63.27	31.72
38GLU	8.03	118.29	178.06	57.97	30.16	80LYS	7.49	122.29	176.37	55.87	32.31
39THR	8.04	109.49	175.37	61.88	70.40	81LYS	7.49	122.29	175.97	55.87	32.87
40LYS	7.46	115.96	174.91	57.48	28.68	82SER	7.70	117.43	174.36	58.00	63.38
41LYS	7.11	120.37	174.31	54.02	32.14	83LEU	7.69	124.96	176.99	55.22	41.72
42PRO			176.68	62.83	32.05						

*ND; Not Detected **Unit; ppm

Acknowledgements

This work was supported by the Innovative Drug Research Center for Metabolic and Inflammatory Disease of the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korean Government (MOST); Korea Healthcare Technology R&D Project, Ministry for Health, Welfare & Family Affairs, Republic of Korea [A092006] and 2010 BK21 Project for Medicine, Dentistry and Pharmacy. We thank NCIRF and KBSI for using NMR machines.

REFERENCES

1. A. De Luca, and G. Laquinto, *Cancer Lett.* **213**, 1. (2004).
2. S. Suerbaum, and C. Josenhans, *Nat Rev Microbiol.* **5,** 441. (2007).
3. S. Renfrey, and J. Featherstone, *Nat Rev Drug Discov.* **1**, 175. (2002).
4. S. B. Jang, M. Chao, S. J. Park, A. R. Kwon, and B. J. Lee, *JKMRS.* **13,** 117. (2009).
5. J. Dinq, M. K. Hayashi, Y. Zhang, L. Manche, A. R. Krainer, and R. M. Xu, *Genes Dev.* **13**, 1102. (1999).
6. F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, *J Biomol NMR.* **6**, 277. (1995).
7. B. A. Johnson, *Methods Mol Biol.* **278**, 313. (2004).
8. T. D. Glddard, and D. G. Kneller, *SPARKY 3*, University of California, San Fancisco
9. D. S. Wishart, and B. D. Sykes, *J Biomol NMR.* **4**, 171. (1994).
10. G. Cornilescu, F. Delaglio, and A. Bax, *J Biomol NMR.* **13**, 289. (1999).
11. C. Maris, C. Dominguez, and F. Allain, *FEBS Journal.* **272**, 2118. (2005).
12. L. Szilagyi, *Prog Nucl Magn Reson Spectrosc.* **27**, 325. (1995).
13. C. Antoine, J. Sandrine, B. Natalya, D. Cyril, S. Stefan, and H. A. Frederic, *Nat Struct Mol Biol.* **18**, 443. (2011).
14. M. H. Lance, and G. F. Michael, *Nature Protocols* **2**, 1849. (2007).