Journal of the Korean Magnetic Resonance Society **2016**, *20*, 36-40
DOI 10.6564/JKMRS.2016.20.2.036

Secondary structure analysis of MRA1997 from *Mycobacterium tuberculosis* and characterization of DNA binding property

Hyo Jung Kim[1], Ki-Young Lee[1], Yena Kim[1], Ae-Ran Kwon[2], and Bong-Jin Lee[1]*

[1]Research Institute of Pharmaceutical sciences, College of Pharmacy, Seoul National University, Gwanak-gu, Seoul 151-742, Korea
[2]Department of Herbal Skin Care, College of Herbal Bio-industry, Daegu Haany University, Gyeongsan 712-715, Korea

Received Mar 19, 2016; Revised Apr 25, 2016; Accepted May 7, 2016

Abstract MRA1997 is a highly conserved protein from mycobacterial strains. However, no structural and functional information is associated with it. Thus, to obtain details about structure and function of this protein, we have utilized NMR spectroscopy. The recombinant MRA1997 was highly purified and its DNA binding mode was characterized. The tertiary structure of MRA1997 was modeled on the basis of our NMR chemical shift data combined with the webserver CS23D. The binding of MRA1997 with DNA was first monitored by electrophoresis mobility shift assays. The residues involved in DNA binding are identified using NMR chemical shift perturbation experiments. Based on our study, we suggest that MRA1997 interacts with DNA and may play an important role in *Mycobacterium tuberculosis* physiology.

Keywords *Mycobacterium tuberculosis*, MRA1997, NMR, EMSA

Introduction

Mycobacterium tuberculosis is a chronic infectious disease that is responsible for the death of more than 1.5 million people annually. One third of the world's population is infected with *tuberculosis*, and new infections occur at a rate of one per second.[1] *M. tuberculosis* is carried in aerosolized droplets and spread to a person or animal. The primary infection is almost asymptomatic, but infections in patients who are immunosuppressed or immunocompromised commonly result in death.[2] Even in immunocompetent individuals, *M. tuberculosis* organisms can reactivate at a later date and causes chronic pulmonary inflammation.[3] Investigation *of M. tuberculosis* over many decades has been attributed to the advent of live attenuated vaccines and antibiotics.[4] However, the widespread use of vaccine and antibiotics lead to the emergence of resistant strains. Drug-resistant strains of *M. tuberculosis* have evolved due to the sequential accumulation of mutations under pressure of anti-tuberculosis drugs.[5] MDR TB (multidrug-resistant tuberculosis) is defined as resistance to at least two *tuberculosis* drugs, isoniazid and rifampicin. XDR TB (extensively drug-resistant tuberculosis) shows additional resistant to any fluoroquinolone, and to any of the second-line antibiotics.[6] Because of these

* Address correspondence to: **Bong-Jin Lee,** Research Institute of Pharmaceutical sciences, College of Pharmacy, S eoul National University, Gwanak-gu, Seoul 151-742, Korea, Tel: +82-2-880-7869; Fax: +82-2-875-3632; E-mail: lbj @nmr.snu.ac.kr

bacterial resistances, the prognosis for tuberculosis is poor. Therefore understanding the detailed mechanism of bacterial pathogenic strategies of *M. tuberculosis* remains a major goal.

In the present study, the structural and functional studies of MRA1997 were conducted. MRA1997 is a hypothetical protein from *M. tuberculosis* H37Rv. MRA1997 is conserved in most mycobacterial strains but its structure and function remains unknown. Since the conservation of a gene implies that it is essential to cell viability, a highly conserved MRA1997 is predicted to have a certain functional value. Herein, we for the first time show that MRA1997 can interact with DNA. The interaction of MRA1997 with DNA was monitored by electrophoresis mobility shift assays. The MRA1997 residues affected by DNA binding are identified using NMR chemical shift perturbation experiments. This structural and functional research regarding highly conserved mycobacterial protein will promote a better understanding of bacterial physiology.

Experimental Methods

Cloning, protein expression, and purification- The predicted ORF of MRA1997 was amplified from *M. tuberculosis* H37Rv genomic DNA using standard PCR methods. The forward and reverse oligonucleotide primers designed using the published genome sequence were 5'-CGCCATATGACCGACT CTGTGGTCGTCCGCG and 5'-CCGCTCGAGACG ACTCAGACGGAAACGCTTG, respectively, where the bases underlined represent the *Nde*I and *Xho*I restriction enzyme cleavage sites. The amplified DNA was inserted into the *Nde*I/*Xho*I-digested pET-21a(+) expression vector (Novagen Inc.). The resulting construct contains eight nonnative residues remains attached at the C-terminus (LEHHHHHH) that facilitate protein purification. The accuracy of the cloning was confirmed by DNA sequencing. The resulting expression plasmid was then transformed into E.coli BL21(DE3) competent cells (Novagen Inc.). Uniformly ^{15}N-labeled [U-^{15}N] MRA1997 protein was expressed in *E. coli* BL21(DE3) strain, and grown in M9 medium supplemented with ^{15}N-NH_4Cl and/or ^{13}C-glucose. Recombinant protein expression was induced by addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and cells were grown for additional 4 hours at 37 °C. The cells were harvested by centrifugation and resuspended in 50 mM Tris-HCl, pH 7.5, and 500 mM NaCl buffer. Cells were lysed by sonication at 4 °C and the supernatant was loaded to Ni^{2+}-NTA (Ni^{2+}-nitrilotriacetate) affinity column (Qiagen; 3 ml of resin per liter of cell culture) previously equilibrated with the same buffer. The column was washed extensively with wash buffer (50 mM Tris-HCl, pH 7.5, 500 mM NaCl, and 50 mM imidazole); then the bound protein was eluted with elution buffer (50 mM Tris-HCl, pH 7.5, 500 mM NaCl, and 500 mM imidazole) until there was no detectable absorbance at 280 nm in the eluent. Purified protein was judged to be over 95% by SDS-PAGE. The protein solution was concentrated using 3,000 Da molecular-mass cut-off spin columns (Millipore Inc.). The protein concentration was estimated by measuring the absorbance at 280 nm, employing the calculated extinction coefficient of 1,490 $M^{-1}cm^{-1}$ (Swiss-Prot; http://www.expasy.org).

NMR spectroscopy- Protein concentrations used for NMR titration experiments was 0.5 mM. NMR buffers included 20 mM Sodium phosphate (pH 6.0), 100 mM NaCl, and 1 mM DTT in 90% H2O and 10% D2O (v/v). All of NMR spectra were recorded at 298K on JEOL ECA 600 MHz NMR spectrometers. The δ ppm values of backbone N and HN resonances of MRA1997 were assigned from the data obtained in an earlier NMR study[7], which shows 88.3% assigned chemical shifts in the [U-15N] 2D-spectrum of MRA1997. The NMR chemical shift perturbation experiments were performed by the addition of DNA to protein. The final protein:DNA ratio was 1:5, and the disappearance and/or shift of the peaks were monitored. All NMR experiments were processed with NMNRPipe[8] and analyzed by NMRviewJ.[9]

Electrophoretic mobility shift assays (EMSA)- To

visualize the binding of MRA1997 with DNA, electrophoretic mobility shift assays were conducted. DNA and MRA1997 were prepared in a binding buffer consisting of 20 mM Tris HCl pH 7.5, and 150 mM NaCl. Varying amount of MRA1997 protein were mixed with DNA, and incubated for 30 min at 4 °C. The total binding reaction was loaded onto 1% agarose gel in 0.5X TBE buffer for electrophoresis, and the results were visualized using a Gel Doc software (Bio-Rad Inc.).

Results and Discussion

The tertiary structure of MRA1997 was predicted using CS23D[10] webserver. The predicted structure of MRA1997 consists of three antiparallel β-strands located at N-terminus, one α-helix, and C-terminus random coiled. The structure was submitted to DALI server to search for the structural homologous. The DALI algorithm revealed that MRA1997 is structurally similar to sentrin-specific proteases, TATA binding proteins, transcription initiation factor, and DNA-directed RNA polymerase subunit. Since most structural homologs are involved in nucleic acid binding, MRA1997 is also expected to have a binding property with nucleic acids (DNA and/or RNA).

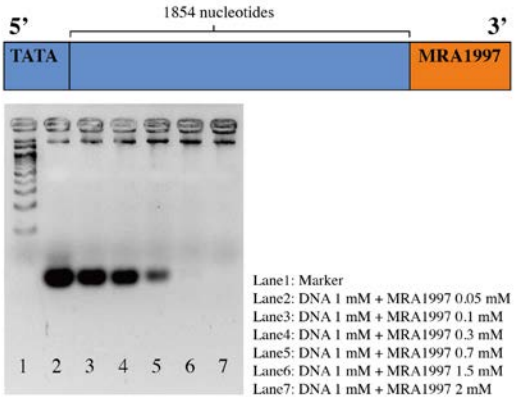

Figure 2. The gel electrophoresis mobility shift assay (EMSA). The schematic representation shows the location of TATA box and the possible promoter region. The EMSA study was performed using increasing amount of protein, and the DNA: protein ratios were used as 1:0.05; 1:0.1, 1:0.3, 1:0.7, 1:1.5 and 1:2.

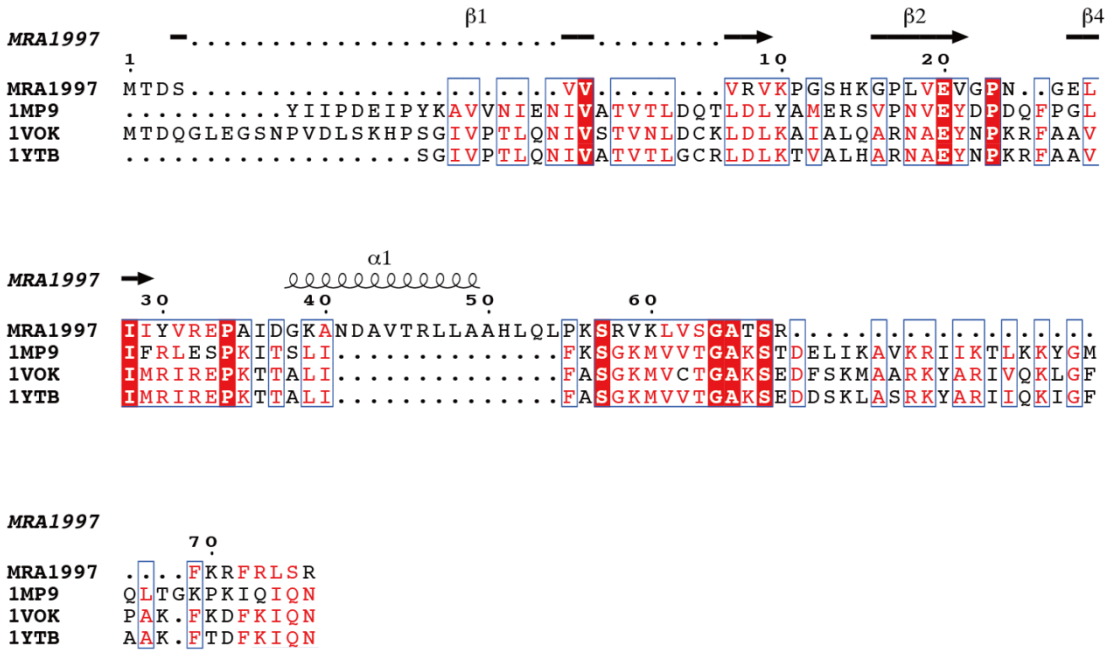

Figure 1. Sequence alignment of MRA1997 with TATA binding proteins. The PDB codes of TATA binding proteins are shown as 1MP9, 1VOK, and 1YTB. The secondary structural topology of MRA1997 is indicated on the top of its amino acid sequence.

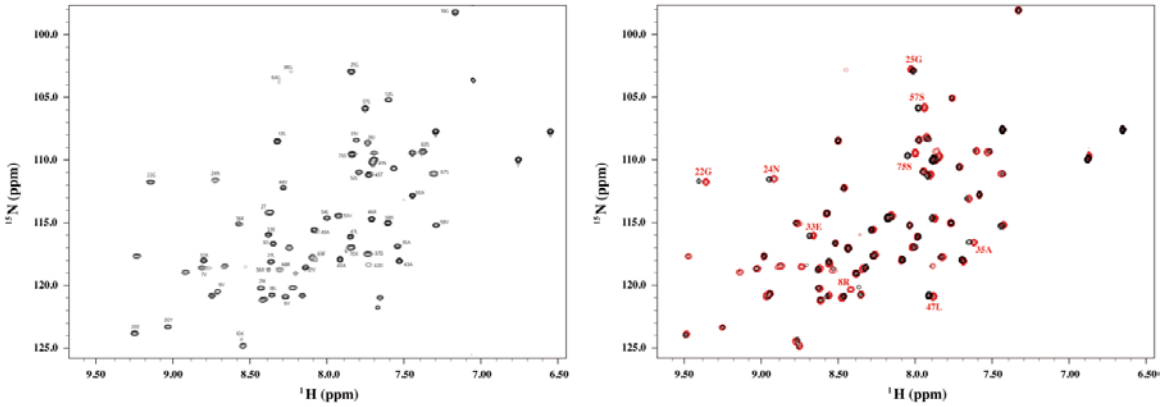

Figure 3. NMR analysis of MRA1997. *Left* [U-^{1}H-^{15}N] 2D-HSQC spectrum of MRA1997. *Right* overlaid [U-^{1}H-^{15}N] 2D-HSQC spectra of MRA1997 in the absence and presence of DNA, which are colored in black and red, respectively. The residues showing significant chemical shift changes upon DNA binding are labeled.

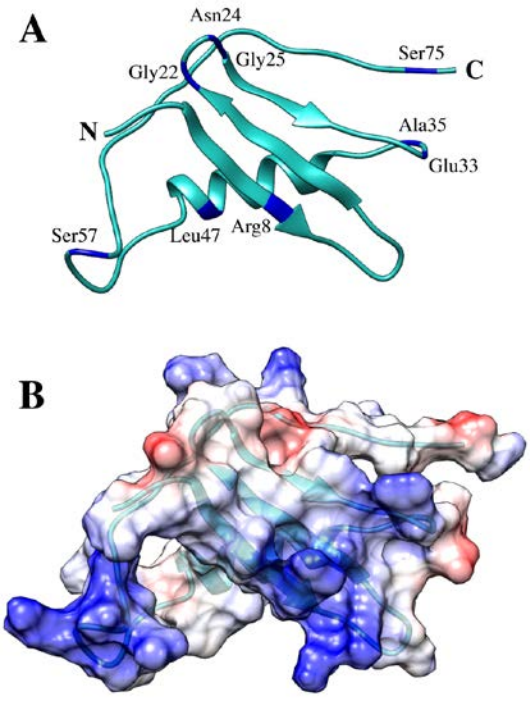

Figure 4. The chemical shift mapping onto the modeled tertiary structure of MRA1997. (A) Mapping of the most perturbed residues upon DNA binding onto the ribbon structure model of MRA1997. (B) The electrostatic surface model of MRA1997 calculated with UCSF Chimera.[14] Surfaces are colored between -10 (red) and +10 (blue) kcal/mol·e.

Further search revealed that the sentrin-specific proteases possess a cysteine residue that act as a functionally important nucleophile in the catalytic activity.[11] As MRA1997 does not contain such cysteine residue, the possibility of that MRA1997 might belong to TATA binding proteins was evaluated. The sequence alignment with TATA binding proteins was done with Clustal Omega webserver tool and viewed using ESPript webserver tool.[12,13] (Figure 1).

Several DNA constructs including TATA sequence that forms palindromic structure were tested. Figure 2 shows the EMSA study of MRA1997 using 25 bp ssDNA (TATACAGCGGACCGATTGTCGGCAA). As a result, free DNA bands gradually disappeared, and protein bound DNA bands gradually appeared as concentrations of protein increased. Furthermore, the MRA1997 residues involved in DNA binding were investigated by NMR titration experiments. The comparison between the two spectra in the absence and presence of DNA showed that the most perturbed peaks upon DNA binding correspond to Arg8, Gly22, Asn24, Gly25, Glu33, Ala35, Leu47, Ser57, and Ser75 (Figure 3).

These residues are mapped onto the modeled structure of MRA1997 as shown in Figure 4A. These residues are distributed in a relatively broad range, suggesting the possibility that (i) the DNA binding

site of MRA1997 is relatively broad and (ii) the conformational changes can arise upon DNA binding. These possibilities can be supported by the fact that small-sized proteins like MRA1997 are relatively difficult to adopt defined nucleotide binding sites and to have restricted motions due to complex networks of individual secondary structure. The structure is also represented by electrostatic surface model (Fig. 4B).

In summary, we identified the DNA binding property of MRA1997 using EMSA and NMR titration experiments and predicted the structure of MRA1997 using chemical shift assignments. Considering the fact that DNA interacting proteins regulate the expression of many genes and thereby play an important role in cellular signaling process, we suggest that MRA1997 might be involved in regulation of mycobacterial physiology through DNA interaction. The insight into the relationship between the cellular role and DNA bound structures of MRA 1997 awaits further investigation.

Acknowledgements

This study was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2015R1A2A1A05001894 & 2014K1A3A1A19067618). This study was also supported by BK21 Plus Project for Medicine, Dentistry, and Pharmacy.

References

1. P. B. Pavlinac, *Int. J. Tuverc. Lung Dis.* **20**, 895 (2016)
2. F. A. Krsulovic, *PloS one* **11**, e0153710 (2016)
3. R. H. Chisholm, *The Proceedings of The Royal Society* **283**, (2016)
4. A. Jain, *Int. J. Tuverc. Lung Dis.* **20**, 870 (2016)
5. J. Tang, *Tuberculosis* **98**, 30 (2016)
6. J. W. Wilson, *Mayo Clinic proceedings* **91**, 482 (2016)
7. H. J. Kim, *Journal of the Korean Magnetic Resonance Society* **19**, 49 (2015)
8. F. Delaglio, *Journal of biomolecular NMR* **6**, 277 (1995)
9. B. A. Johnson, *Methods in molecular biology* **278**, 313 (2004)
10. D. S. Wishart, *Nucleic acids research* **36**, 496 (2008)
11. L, N. Shen, *The Biochemical Journal* **397**, 279 (2006)
12. X. Robert, *Nucleic acids research* **42**, W320
13. P. Gouet, *Bioinformatics* **15**, 305
14. E. F. Patterson, *J. Comput. Chem.* **25**, 1605 (2004)