藥學碩士 學位論文

Structural and Functional Study on DNA binding of MazE2 from *Mycobacterium tuberculosis*

*Mycobacterium tuberculosis*에서

유래한 **MazE2**의 **DNA binding**에 대한

구조적, 기능적 연구

2017年 8月

서울대학교 대학원

약학과 물리약학전공

박 성 현

Abstract

Structural and Functional Study on DNA binding of MazE2 from *Mycobacterium tuberculosis*

Park Seong-Hyun, physical pharmacy, the graduate school,
Seoul National University

Mycobacterium tuberculosis was first discovered by Robert Koch in 1882 and is a human infectious strain that causes lung disease through the respiratory tract. *M. tuberculosis* differs from ordinary bacteria in that it has thick membranes on the cell surface and does not stain Gram stain. However, when stained with Ziehl-Neelsen, it is difficult to decolorize by acid, alcohol and boiling, and this property is called acid-fast. Gram staining is negative, but they do not have external membranes and are classified as acid-fast gram-positive bacteria. Treatment of patients infected with *M. tuberculosis* is mainly medication. Isoniazid, rifampin, and pyrazinamide are used, and streptomycin is also used.

Mycobacterium tuberculosis has the largest number of Toxin-Antitoxin system pairs among bacteria. The Toxin-Antitoxin system is classified into six types according to antitoxin properties. The target protein Rv0660c of this study is MazE, an antitoxin of MazEF system, which is one of Type2 Toxin-Antitoxin system.

The Toxin-Antitoxin system plays a role in inhibiting the growth of bacteria or leading to death by external stimuli-activated toxins. Antitoxin normally binds to toxin and inhibits the activity of toxin. but it breaks down in extreme situations. Thus, dissociated toxin induces growth inhibition and death. This also induces latency and lowers susceptibility to antibiotics.

We overexpress the N1-44 construct of Rv0660c to characterize the tertiary structure of this Rv0660c protein. His-tagged Rv0660c protein was purified using Immobilized Metal Affinity Chromatography (IMAC), and the His-tag was removed through thrombin cutting to further increase the protein purity. As a result, the crystal could be made and the structure could be obtained with a high resolution of 1.69 Å.

The DNA binding experiments of Rv0660c were carried out by NMR and EMSA. Rv0660c was labeled with 13C and 15N isotopes. Heteronuclear multidimensional NMR spectra were measured and backbone assignments were made through HNCO, HNCA, HNCACO, HNCOCA, HNCOCACB and HNCACB spectra. The TALOS program was used to predict the secondary structure and identify the parts which interact DNA through DNA titration. We confirmed the binding of Rv0660c with DNA in vitro from Electrophoretic Mobility Shift Assay (EMSA).

Key words: SBDD(Structure Based Drug Design), toxin-antitoxin system, antitoxin, x-ray crystallography, NMR(Nuclear Magnetic Resonance), EMSA(Electrophoretic Mobility Shift Assay)

Student number: 2015-21873

Contents

List of Figures

Abbreviations

3D	Three dimensional
E. coli	*Escherichia coli*
M. tuberculosis	*Mycobacterium tuberculosis*
IPTG	Isopropyl-β-D-thiogalactopyranoside
LB	Luria Bertani
M9	Minimal 9
NMR	Nuclear Magnetic Resonance
OD	Optical Density
PAGE	Poly Acryl amide Gel Electrophoresis
SDS	Sodium Dodecyl Sulfate
PCR	Polymerase Chain Reaction
ppm	Part Per Million
RPM	Revolutions Per Minute
TB	Tuberculosis
IMAC	Immobilized Metal ion Affinity Chromatography
HSQC	Hetero nuclear Single Quantum Coherence spectroscopy
TALOS	Torsion Angle Likelihood Obtained from shift and Sequence Similarity

Ⅰ. Introduction

1.1 Structure Based Drug Design (SBDD)

SBDD is the way to develop new drugs. In general, designing drug molecules while looking at the stereo structure of the target protein is called structure-based drug design. To do this, we analyze the 3D structure through X-ray crystal structure analysis or NMR.

SBDD is far superior in cost and time to traditional drug development methods that improve the compounds discovered accidentally from natural products such as herbal medicines. The traditional method depends on the intuition and experience of the researchers and requires a lot of time and effort. It takes 3-5 years to lead discovery. However, if the three-dimensional structure of the disease target protein is identified, you can shorten up to 2-3 years and save a lot of money.

In addition, since SBDD is a method of designing drugs in a rational manner, it is possible to develop a drug having a high activity and low side effect. Since it is a method to develop a drug structurally most suitable for a target corresponding to a therapeutic purpose considering an adverse effect from the beginning, the probability of entering and passing the clinical trial phase is high and the probability of unexpected side effects occurring in post marketing can be very low.

1.2 Characteristics of *Mycobacterium tuberculosis*

Tuberculosis has been a terrible infectious disease that has suffered a great deal of life around the world, long since it has given much suffering to humankind and has been called 'white pest' in the 19th century.

However, now that tuberculosis is in a country with poor sanitation as well as malnutrition, tuberculosis is considered a 'forgotten disease' or a 'past disease'. However, tuberculosis is a disease that is never forgotten and is still a disease that is threatening the health of the world.

It was discovered by Robert Koch in 1882 as a major cause of tuberculosis. It grows much slower than other bacteria and can survive long in dry conditions due to the fat-rich cell walls. Tuberculosis is most often caused by airborne infection (droplet infection). [Issar Smith., 2003]

It is resistant to strong acids and alkalis, but it is vulnerable to heat and sunlight. If exposed to direct sunlight, it will die in a few minutes. It is a thin and long bacterium, and both ends are dull, round, straight or slightly curved. It does not form sporophytes, the cavernous membranes, and has no motility. [Issar Smith., 2003]

1.3 Epidemiology of *Mycobacterium tuberculosis*

According to World Health Organization (WHO), around 5000 patients die every day from tuberculosis. It is estimated that the number of new tuberculosis cases per year is about 10 million, and the death toll is about 1.8 million.

According to the World Health Organization (WHO), the number of TB patients in Republic of Korea by 2015 is about 40,000, which is about 80 per 100,000 people. It is the lowest among OECD member countries. The rate of tuberculosis incidence and mortality rate is the number one, and the incidence of multi-drug-resistant tuberculosis, which is particularly problematic, is also the number one.

It is the immune TB that aggravates the global TB crisis. If you are resistant to the first drug known to be the most effective (rifampicin, isoniazid), it is judged to be multi-drug-resistant tuberculosis (MDR-TB). In addition, when resistant to secondary drugs, it becomes a more complex form of Extensively drug-resistant tuberculosis (XDR-TB), also known as super tuberculosis. With this relationship, it is becoming more important than ever to develop new drugs for resistant tuberculosis.

1.4 Toxin– antitoxin System

We are mainly studying Toxin-antitoxin System as a target for new drug development. The TA system is found in almost all bacteria and some fungi. [Yoshihiro Yamaguchi, Jung-Ho Park, and Masayori Inouye., 2011] Normally, cognate antitoxin against toxin and toxin are complexed and toxin is not activated. In a stress condition such as nutrition starvation and antibiotics, the less stable antitoxin is rapidly degraded and a single, free active toxin targets the essential processing steps of cells such as DNA replication and cell wall synthesis to prevent the growth of damaged cells, it leads to cell death and lowers susceptibility to antibiotics. [Kenn Gerdes and Etienne Maisonneuve., 2012]

TA systems are divided into six groups according to the mechanism of TA gene regulation and the characteristics of antitoxin. Among them, Type 2 is the antitoxin protein which directly binds to toxin protein and inhibits the toxicity of toxin. In *M. tuberculosis*, the most abundant family is the type 2 TA system in which 67 out of 79 TA belongs.

And *M. tuberculosis* has the most diverse Toxin-Antitoxin systems among bacteria. The greatest feature of *M. tuberculosis* as a germ is its ability to survive for long periods of time in a host-resistant, non-multiplying state, and then wake up again later to cause disease. One of the notable areas of attention to demonstrate this capability is that it has a vast number of toxin-antitoxin systems, presuming that the activation of the TA system will form the pathogenesis and the possibility of TA with other bacteria is being studied. [Ambre Sala, Patricia Bordes and Pierre Genevaux., 2014]

1.5 Characteristics of Rv0660c

Rv0660c is a Type 2 toxin-antitoxin system of *Mycobacterium tuberculosis* with Rv0659c. Rv0659c is a toxin known as MazF2 and Rv0660c is an antitoxin known as MazE2. Rv0660c is 81 amino acids with a molecular weight of 9kDa (9169) and theoretical PI is 4.4.

At first, full-length experiments failed to form crystals because they were well resolved during the purification process. Therefore, we excluded the flexible c-terminal part, which was well disassembled, and tried to divide the N-terminal part into several parts by looking at the expected secondary-

structure through J-pred. The length of the sample that succeeded in the crystallization was 1-44.

1.6 Purpose of the study

The purpose of this study is to identify the tertiary structure of the N1-44 construct of Rv0660c through x-ray crystallography and to obtain information on DNA binding through NMR and EMSA experiments.

To accomplish this goal, the N-terminal parts of Rv0660c were selected and cloned. Cloned samples were transformed into competent cells for expression and then overexpressed using *E. coli*. Over-expressed proteins were able to undergo solubility testing and purification. The purified sample was crystallized through crystallization and the tertiary structure was obtained with a high resolution of 1.69 Å.

Various NMR and EMSA experiments were performed to obtain information on DNA binding. Particularly, at this time, we could confirm which parts of the tertiary structure bind DNA in comparison with the structure obtained through crystal.

Ⅱ. Materials and Methods

2.1 Materials

2.1.1 Reagents

The genomic DNA of *Mycobacterium tuberculosis* H37Rv was purchased from BIONEER. Expression host *E. coli* codon plus(DE3) and pET-28a (+) expression vector were bought from Novagen Inc. (Darmstadt, Germany). Sequencing service was from Cosmogenetech Service. PCR premix kits were purchased from INTRON Biotechnology Inc. PCR primer bought from BIONEER. T4 DNA ligase, T4 DNA ligase buffer were purchased from TaKaRa Inc. BSA, Restriction endonucleases, and buffer4 were purchased from New England Biolabs(NEB). Isopropyl β-D-1-thiogalactopyranoside (IPTG) and LB media were purchased from Calbiochem (Nottingham, UK). Kanamycin was bought from Biosesang. Vitamins solution and other reagents to make buffer solutions and media for M9 were bought from SIGMA (St. Louis, USA). $^{15}NH_4Cl$ and ^{13}C-glucose for isotope labeling were bought from Cambridge Isotope (Andover,MA,USA). Ni^{2+}-affinity column resin which is HisTrap HP column was from GE Healthcare Life Sciences (Little Chalfont, United Kingdom). All reagents and chemicals were purchased from certified vendors in analytical or biotechnical grade for the reliability of results.

2.1.2 Apparatus

PCR reaction was carried out by Perkin-Elmer PCR system 9600 (Perkin-Elmer, USA). The concentration of protein was measured using NanoDrop ND-1000 from Coleman Technologies Inc. (Wilmington, DE, USA). Cell lysis was performed by the sonic oscillator, sonifier450 designed by Branson Ultrasonic Corporation (Connecticut, USA). J2-MC and the fraction collector were bought from Bio-Rad Laboratories Inc. (California, USA). Centricon, Centriprep were bought from Millipore Corporation (Massachusetts, USA). Ni^{2+}-affinity Chromatography was conducted by HisTrap HP column in fast protein liquid chromatography (FPLC) (AKTA prime, GE Healthcare, USA). Size Exclusion Chromatography was conducted by superdex 75 16/200 column fast protein liquid chromatography (FPLC) (AKTA, GE Health care, USA). NMR tubes were prepared from Shigemi Inc. (Tokyo, Japan). All NMR spectra were recorded on JEOL 600MHz NMR spectrometer (JEOL Ltd., Tokyo, Japan). All NMR measurements were analyzed by Bruker NMR systems 600MHz (California, USA). All NMR data were processed by NMRpipe, NMRDraw and NMRView from Silicon Graphic Inc. (California, USA).

2.2 Methods

2.2.1 Cloning of target protein

The N1-44 of Rv0660c gene was amplified by PCR from *M. tuberculosis* H37Rv genomic DNA. PCR was carried out under general procedure. The sense and antisense oligonucleotide primers contained the restriction enzyme sites, NdeI and XhoI, respectively. Consecutive six histidine residues were tagged to the carboxyl terminus of protein in order to increase the efficiency of purification. The amplified target protein was cloned into the expression vector pET-28a(+) (Novagen) that was digested by both NdeI and XhoI (Figure 1). After ligation process between truncated PCR inserts and truncated vectors, the recombinant plasmids were transformed to competent cell, *E. coli* DH5α. After sequence analysis of the recombinant plasmids, they were transformed to *E. coli* cells of BL21(DE3), Codon plus(DE3), C41 and Rosetta2(DE3) for protein over expression test with IPTG.

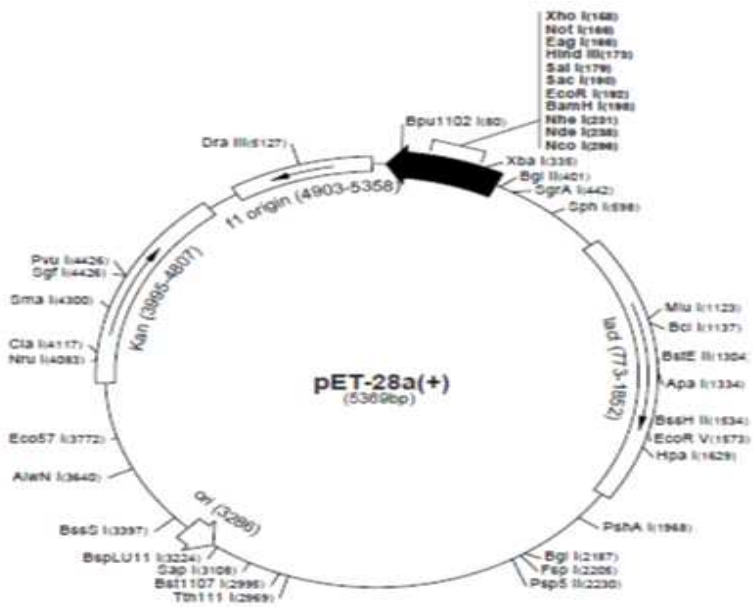

Figure 1. pET-28a (+) cloning vector.

2.2.2 Protein Over-expression and purification

Each of the transformed BL21(DE3), Codon plus(DE3), C41 and Rosetta2(DE3) cells was grown in 5 ml of LB media with 5 μl of kanamycin (30μg/ml) at 37℃ and a shaking speed of 180 RPM overnight. Next day, they were inoculated into 200 ml of the fresh autoclaved LB media each containing 200μl of kanamycin and incubated at 37℃ and a shaking speed of 180 RPM until the value of Optical Density (O.D.) at 600 nm was approximately 0.5. 200μl from each of the cultured cells was obtained to micro centrifuge tubes. They were centrifuged for 5 minutes at 13,000 RPM to save pellet only for later analysis. 200μl of isopropyl-β-D-1-thiogalactopyranoside (IPTG, 0.5μg/ml of final concentration) was added to each of the cultured cells to induce expression. After 4 hours, we took 200μl from each of the induced cells and removed supernatants after centrifugation. The pellet fractions for both pre-samples and post induced samples were suspended in 50μl of distilled water and boiled for 10 minutes.

We ran SDS-PAGE electrophoresis to observe the presence of induced band. Overexpressed cells were harvested by a centrifugation (Beckham J2-MC) at 4℃ and 10,000 RPM for 10 minutes. The cell pellet was suspended in 30 ml of the lysis buffer (20mM Tris-HCl (pH 7.8) and 500mM NaCl) at 4℃. The suspended cells were disrupted by an ultrasonic homogenizer on ice bath for total 10 minutes (pulse on 1 second, pulse off 5 seconds) at 30% amplitude. The lysate was clarified by centrifugation at 18,000 RPM for 1 hour at 4℃. The pellet extraction and supernatant were tested on SDS-PAGE to confirm protein's solubility.

For NMR spectroscopy, ^{15}N-labeled and $^{15}N,^{13}C$-labeled protein sample was prepared following the procedure above. The seed strain was inoculated in 5ml of LB media containing kanamycin(30μg/ml) at 37℃ with shaking at-180 RPM overnight. They were inoculated in 10ml of M9 media without isotope labeled carbon and nitrogen sources for 9 hours at 37℃. The cells were scared up in 1L of M9 minimal media supplemented with 1g ^{13}C-glucose and 1g ^{15}N-NH_4Cl. After sonication, the supernatant was filtered by 0.45μm syringe filter and loaded into a Ni2+ affinity column, which had been previously equilibrated with the basic buffer (20mM Tris-HCl (pH7.8) and 500mM NaCl). Then, the column was washed with the same buffer at gravity flow. The target protein was eluted by a stepwise concentration gradient of imidazole from 50 to 500mM.

The elution fractions were checked visually by SDS-PAGE electrophoresis and concentrated by ultrafiltration using an Amicon Ultra-15 centrifugal filter unit (Millipore,USA) up to 2ml. Concentrated samples were cleaved with Thrombin overnight at 4°C to remove His-tag and then loaded onto IMAC to purify only Rv0660c without His-tag. The solution was then concentrated until the concentration of target protein becomes approximately 0.2mM. All the purified fractions were analyzed by SDS-PAGE electrophoresis and concentrated by ultrafiltration using an Amicon Ultra-15 centrifugal filter unit (Millipore, USA).

2.2.3 Crystallization

Crystallization conditions were searched by the sitting drop vapor-

diffusion method at 20℃ using screening kits from Hampton Research Inc. (Crystal Screen I, II, Index I, II) and from Emerald Biosystems Inc. (Wizard I, II, III, and IV). The crystallization drops were made by mixing the protein solution and the reservoir solution in one-on-one ratio. The well-made crystals appeared in- 2-3 days in optimized reservoir solution consisting of 0.1 M Tris pH 8.5, 0.3 M Magnesium formate dihydrate. By optimizing this buffer condition, we could get a better crystal. The crystal was vitrified using the cryoprotectant solution consisting of the reservoir solution supplemented with 20% (v/v) glycerol. Crystals were soaked in the cryoprotectant solution for a few seconds before being frozen in liquid nitrogen.

2.2.4. X-ray data collection and structure determination

X-ray diffraction data was collected using synchrotron radiation on ADSC Q315r detector at beamline PAL-5C (SBII) (Pohang, South Korea) at λ=0.97944. Raw data was processed and scaled using HKL2000 program package. [Otwinowski and Minor., 2002] The structure was determined by molecular replacement using the phase program in CCP4 suite of programs and Phenix. [Adams et *al*., 2010] Coot was used for manual model. [Emsley and Cowtan., 2004]

2.3 Structural and Functional studies by NMR spectroscopy

2.3.1 NMR data collection

NMR spectra were collected by JEOL 600 MHz NMR spectrometer equipped with a triple resonance probe and xyz-pulsed field gradient unit. 2D-^{15}N HSQC at 298K(25°C) and 303K (30°C) were recorded using 1mM concentration of ^{15}N-labeled Rv0660c protein in 20mM Tris-HCl(pH 7.8), 500mM NaCl and 10 % D_2O. 3D NMR spectra for backbone assignments (HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HN(CA)CO and HNCO) were measured using 1mM of ^{15}N, ^{13}C-labeled protein at 298K(25°C) in same buffer condition above. All NMR datasets were processed in NMRPipe [Delaglio *et al*., 1995] and analyzed in NMR View Program. [Johnson BA and Belvins RA., 1994]

2.3.2 Backbone assignment

2D ^{1}H-^{15}N HSQC was first step for structural studies. It was measured at 298K(25°C) and 303K (30°C) in order to find the proper temperature for 3D NMR experiments. ^{1}H chemical shift values were referenced to DSS and ^{15}N chemical shift values were referenced indirectly by multiplying the ^{1}H-

carrier frequency. Each peak in 2D-HSQC spectrum should correspond to an amino acid residue in the sequence of N1-44 of Rv0660c. HNCA, HN(CO)CA, HNCO, HN(CA)CO, HNCACB and CBCA(CO)NH spectra were recorded to assign chemical shift values for $^{1}H^{N}$, ^{15}N, ^{13}CO, $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ of a protein backbone. The phase was correctly adjusted in NMRPipe and the processing script with optimized parameter was executed. [Delaglio*et al*., 1995] The sequence-specific resonance assignments were performed using standard procedures in NMR View. [Johnson BA and Belvins RA., 1994]

2.3.3 Secondary structure prediction based on TALOS

The secondary structure of Rv0660c was predicted by TALOS sever based on assigned chemical shift values of Rv0660c. TALOS program is a protein`s secondary structure prediction program. We input assigned chemical shift values of the protein backbone ($^{1}H^{N}$, ^{15}N, ^{13}CO, $^{13}C^{\alpha}$ and $^{13}C^{\beta}$) and N1-44 of Rv0660c protein sequence information into TALOS web server. It calculated phi and psi backbone torsion angles and provided secondary structure information with the Ramachandran plot where the residue is probable to reside in.

2.3.4 DNA synthesis and preparation

NMR spectroscopy also can be used to exam structural perturbations upon complex formation. It provides information about chemical shift changes of amino acid residues involved in the interaction protein and DNA. We synthesized a 20-mer double stranded oligonucleotide sequence from BIONEER to study function of DNA binding. The sequence is shown in Figure 2.

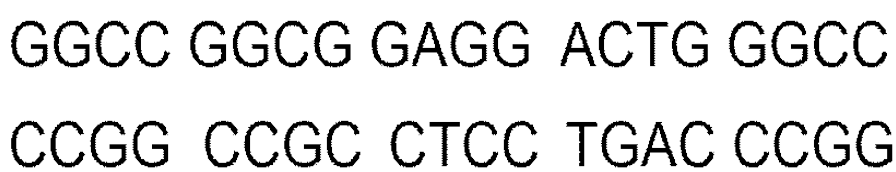

Figure 2. The DNA oligonucleotide sequence for NMR titration experiment.

First, 2D ^{1}H-^{15}N HSQC for N1-44 of Rv0660c protein was measured at 298K(25°C) as the reference. The concentration of N1-44 of Rv0660c protein was 1mM in 20mM Tris-HCl(pH 7.8), 500mM NaCl and 10 % D_2O. And then, 2D ^{1}H-^{15}N HSQC spectra were recorded with a gradient of protein:DNA=1:0 to 1:0.2, 1:0.4, 1:0.6. All these NMR spectra were also processed and analyzed by NMRPipe and NMRView.

2.4 Electrophoretic Mobility Shift Assay (EMSA)

The 5′-biotinylated DNA oligonucleotides used for EMSA contained-

complementary single strand DNA having sequence 5′-GGCCGGCGGAGGACTGGGCC -3′ and 5′- CCGGCCGCCTCCTGACCCGG -3′. Binding reaction mixtures (20μl) containing 10μM N1-44 of Rv0660c protein and increasing amounts of DNA(10, 50, 100, 200μM) in 20mM Tris-HCl(pH7.8), 500mM NaCl were incubated at 25 °C for 10 min. After reaction at 25 °C for 10 min, the mixtures were then loaded on a 10 % polyacrylamide gel in 0.5×TBE buffer at 100V for non-denaturing electrophoresis and transferred to a nylon membrane for detection using the Light-shift EMSA kit (Thermo Fisher Scientific, Inc.) and analyzed using Molecular Image Gel Doc TM XR+ System with Image Lab TM software. (Bio-Rad Laboratories, Inc.)

Ⅲ. Result

3.1 Protein preparation

3.1.1 Cloning, overexpression and purification

N1-44 of Rv0660c was transcribed under control of the T7 promoter and *lac* operator in pET-28a vector. The protein was highly expressed in *E. coli* codon plus(DE3) at 37°C incubation and fully soluble. The result of-

expression is given in Figure 3. Ni^{2}+ ions which immobilize by forming covalent bonds with his-tag recombinant proteins are commonly used. We used the Ni^{2}+ affinity column for the first purification process. The purified proteins were eluted in the presence of 100–350mM imidazole. Figure4 shows SDS-PAGE of the first purification result. Subsequently, purified $Rv0660c^{1-44}$ was cut overnight with Thrombin at 4°C to remove His-tag. Figure5 shows SDS-PAGE of $Rv0660c^{1-44}$ with His-tag removed. The complete $Rv0660c^{1-44}$ with His-tag removed can be obtained by loading again on the Ni2+ affinity column. After purification, The final Concentration of $Rv0660c^{1-44}$ was about 1mM.

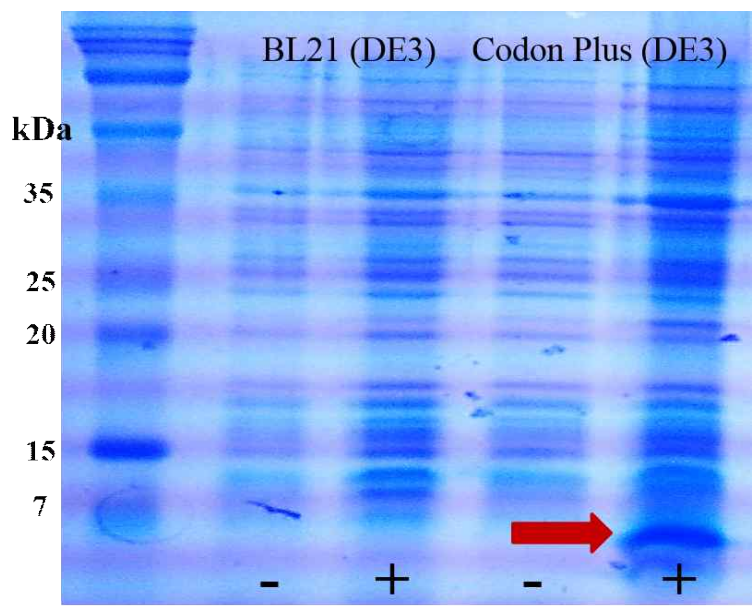

Figure 3. SDS-PAGE result of over expression test of $Rv0660c^{1-44}$.

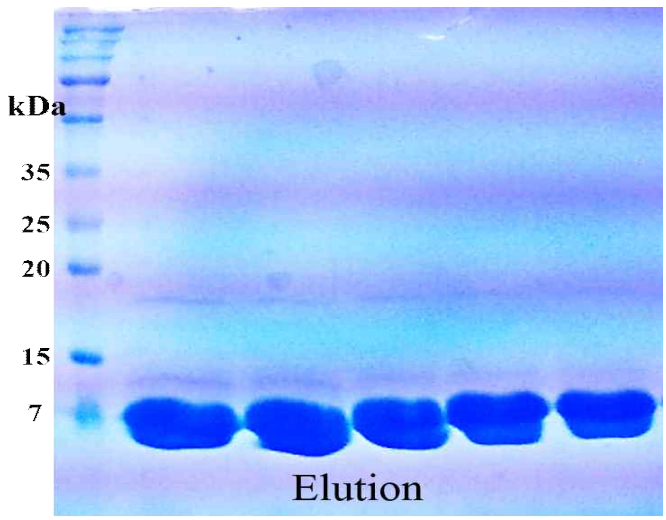

Figure 4. SDA PAGE result of purification of Rv0660c$^{1-44}$ by using Ni^{2}+ affinity-column.

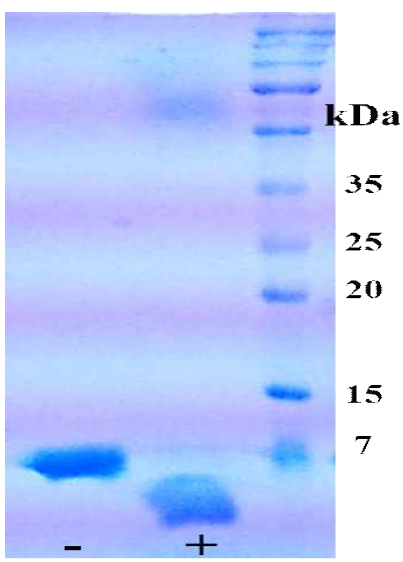

Figure 5. SDA PAGE result of final purification which is thrombin cutting of-Rv0660c$^{1-44}$.

3.1.2 Crystallization

The crystals of Rv0660c$^{1-44}$ made by sitting-drop vapor diffusion are shown in figure 6. crystals which were formed in 0.1 M Tris pH 8.5, 0.3 M Magnesium formate dihydrate could diffract X-ray well. Crystals of Rv0660c$^{1-44}$ look like columnar shape. When we optimize the buffer conditions for this crystal hit, we can get a crystal of a much bigger and better shape and the shape look like octahedron. The big and sharp crystal of N1-44 of Rv0660c is shown in figure 7.

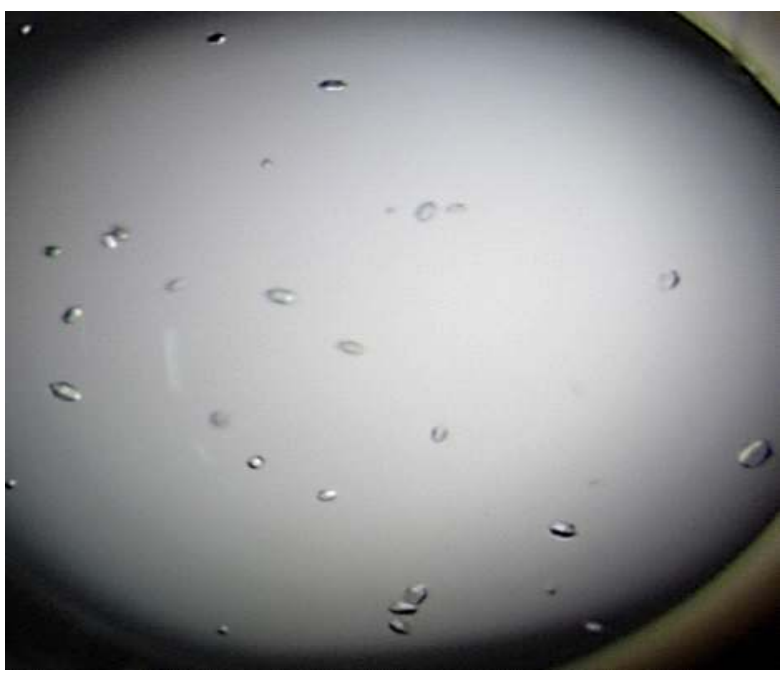

Figure 6. The crystals of Rv0660c$^{1-44}$.

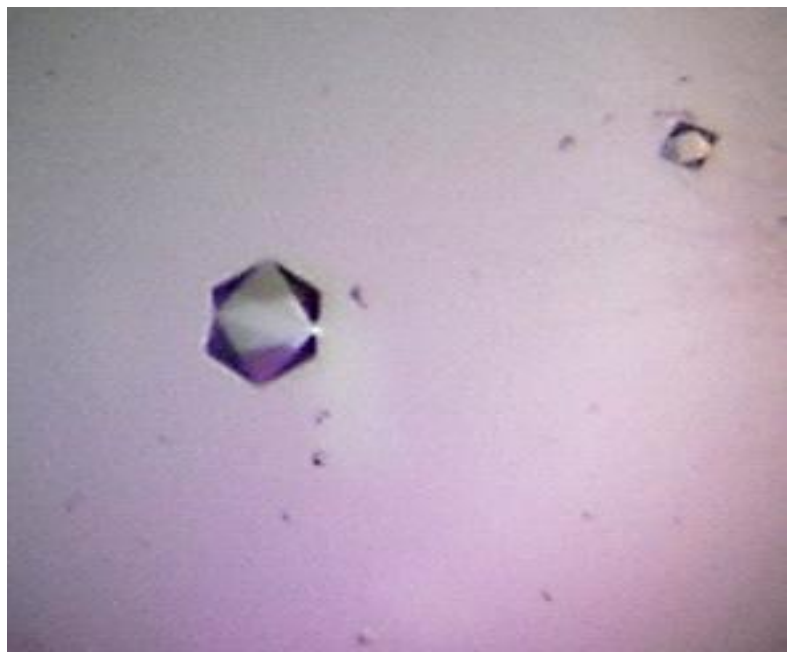

Figure 7. The optimized crystals of Rv0660c$^{1-44}$.

3.2 Crystal structure of Rv0660c$^{1-44}$

Figure 8. shows the crystal structure of the Rv0660c$^{1-44}$. Two monomers each consisting of one β-sheet and two α-helix form antiparallel homodimers and it seems that it has a ribbon-Helix-Helix motif as a DNA binding domain.

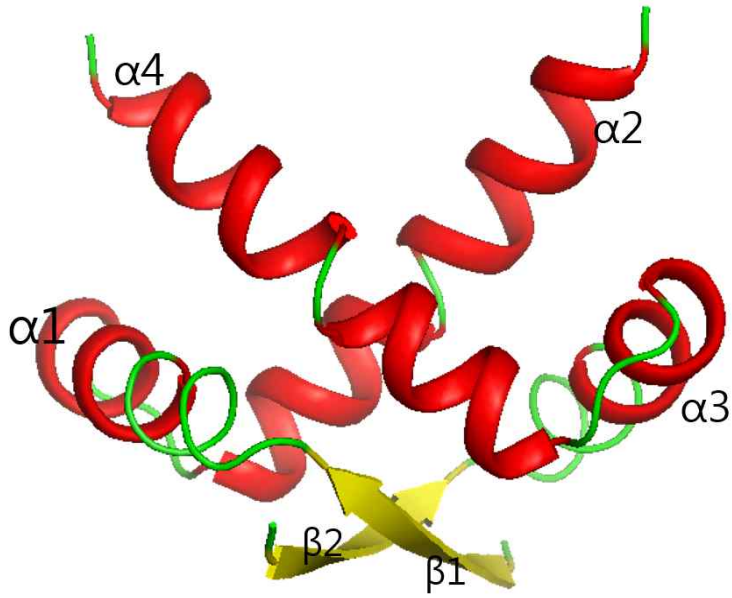

Figure 8. The crystal structure of Rv0660c$^{1-44}$ dimer.

3.3 NMR studies of ^{15}N or ^{15}N-^{13}C labeled Rv0660c$^{1-44}$

3.2.1 2D ^{1}H-^{15}N HSQC

Figure 9 shows ^{1}H-^{15}N HSQC spectrum of Rv0660c$^{1-44}$ at the optimized buffer condition. Since temperature also affects rate of folding in the tertiary structure of protein, we observed 2D ^{1}H-^{15}N HSQC spectra at 298K (25℃) and 303K (30℃). There are more cross peaks with high intensity at 298K (25℃) and the peaks were spread out very well. It means Rv0660c$^{1-44}$ protein is folded in the most stable form at 298K(25℃). The concentration of protein is 1 mM in 20mM Tris-HCl (pH 7.8) and 500mM NaCl.

3.2.2 Sequential backbone assignment

For the backbone assignments of Rv0660c$^{1-44}$, six NMR spectra of ^{15}N-^{13}C labeled Rv0660c$^{1-44}$ were carried out; HNCA, HN(CO)CA, HNCO, HN(CA)CO, HNCACB and HN(CO)CACB.

HN(CO)CA, HN(CO)CACB and HNCO correlate inter-residue and HNCA, HNCACB and HN(CA)CO correlate both inter and intra-residue. HN(CO)CA that gives $^{1}H^{N}$(i), ^{15}N(i) and $^{13}C^{\alpha}$(i-1) chemical shifts and HNCA that gives chemical shifts of $^{1}H^{N}$(i), ^{15}N(i), $^{13}C^{\alpha}$(i-1) and $^{13}C^{\alpha}$(i), were used to assign C $^{\alpha}$ resonance signals. HNCACB that provides chemical shift data of $^{1}H^{N}$(i), ^{15}N(i), $^{13}C^{\alpha}$(i), $^{13}C^{\beta}$(i), $^{13}C^{\alpha}$(i-1) and $^{13}C^{\beta}$(i-1) was used, in tandem with HN(CO)CACB that correlates $^{1}H^{N}$(i), ^{15}N(i), $^{13}C^{\alpha}$(i-1) and $^{13}C^{\beta}$(i-1), to identify C^{β} resonance signals. HN(CA)CO that gives $^{1}H^{N}$(i), ^{15}N(i), ^{13}C(i) and ^{13}C(i-1) chemical shifts and HNCO that provides chemical shift data of $^{1}H^{N}$(i), ^{15}N(i) and ^{13}CO(i-1) were combined to determine ^{13}C chemical shifts.

Approximately 95% of backbone assignment was completed. Figure 9 shows a 2D ^{1}H-^{15}N HSQC spectrum with assigned peaks of Rv0660c$^{1-44}$ at the optimal buffer condition.

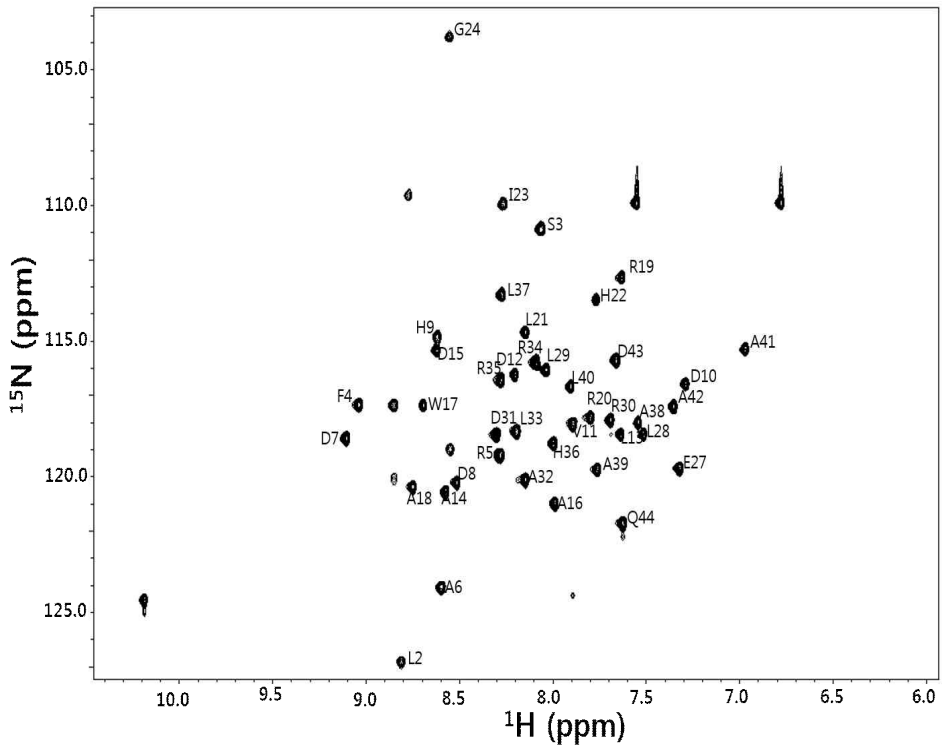

Figure 9. 2D ^{1}H-^{15}N HSQC spectrum of Rv0660c$^{1-44}$. Numbers indicates sequence-numbers and capital letters are single-letter codes for amino acids.

3.3.3 Predicted secondary structure of Rv0660c$^{1-44}$

Using the acquired chemical shift value and other information from NMR spectrum, we requested TALOS server to predict secondary structure.

Basically, the meaning of '-1' is the α-helix tendency and the meaning of '1'means the β-strand tendency of the atom of the residue. Meaning of '0' is the chemical shift within the reference value range. Figure 10 shows Secondary structure prediction based on TALOS. And, as shown in the figure 10, it can be seen that the structure of one β-sheet and two α -helix is the same as the tertiary structure identified by X-ray crystallography.

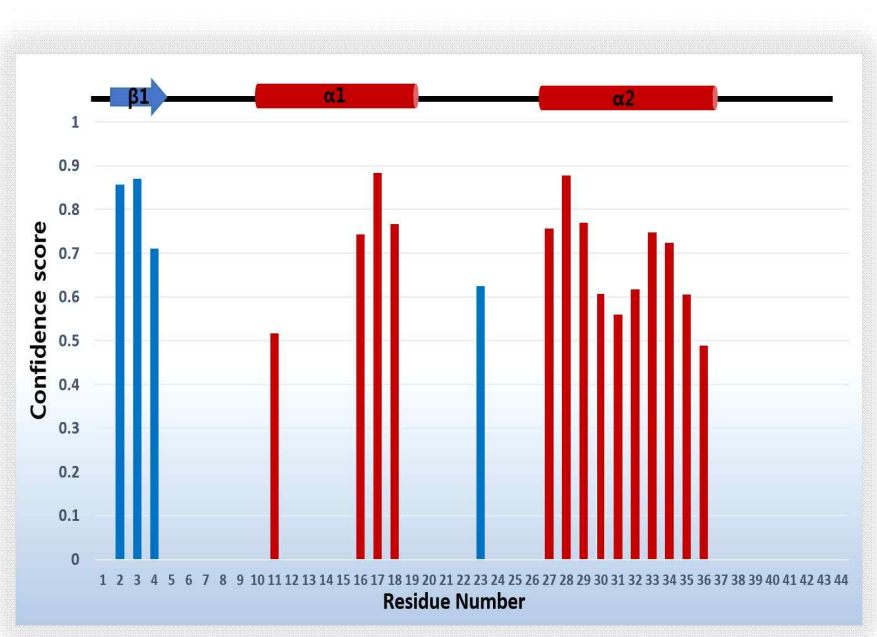

Figure 10. Secondary structure prediction of Rv0660c$^{1\text{-}44}$ based on TALOS.

3.4 Functional study

3.4.1 Protein-DNA binding

In order to identify the property of DNA binding and determine the binding site of $Rv0660c^{1-44}$, We performed a NMR protein-DNA titration experiment. The DNA sequence was a double stranded and a palindrome sequence located in promoter region of $Rv0660c^{1-44}$. Overlap of HSQC cross peak patterns of free and DNA-bound protein showed that several peaks were shifted (Figure 11). And, as you know, figure 2 shows The DNA sequence that we deigned.

By plotting HSQC by concentration of protein and DNA, you can see where the chemical shifts have changed so you can see where the interaction occurs when binding to DNA. It can be seen that the position of the residues of Peak's movement is shown in red in the crystal model (Figure 12).

The following figure13 is a graph of Intensity loss. When HSQC is photographed through DNA titration, it can be understood that the DNA will directly bind to the portion where the intensity loss is severe. The residues that are expected to bind directly are shown in purple on the Crystal model (figure 14).

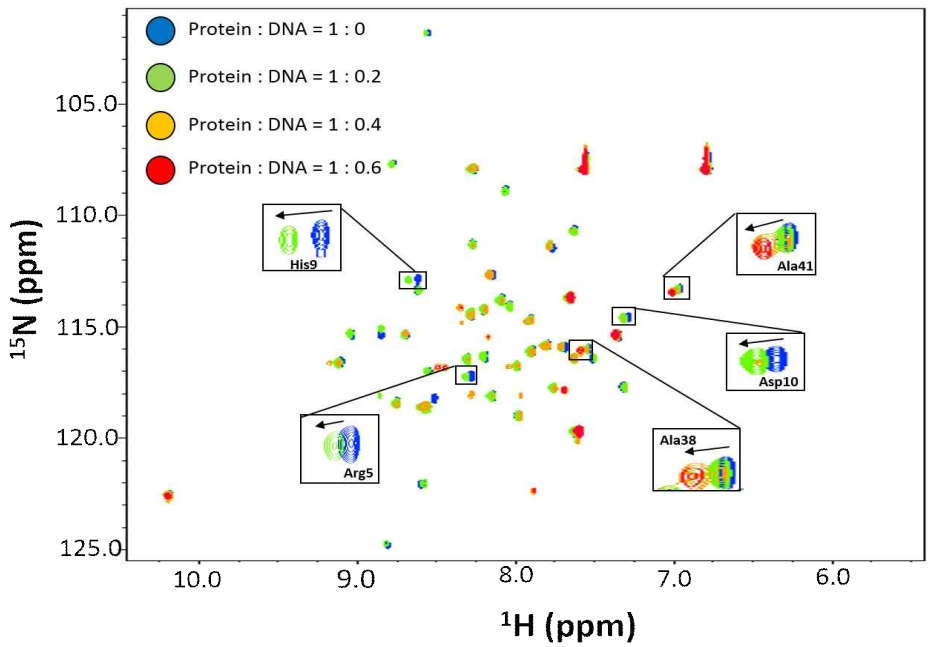

Figure 11. The shifted peaks of NMR titration experiments.

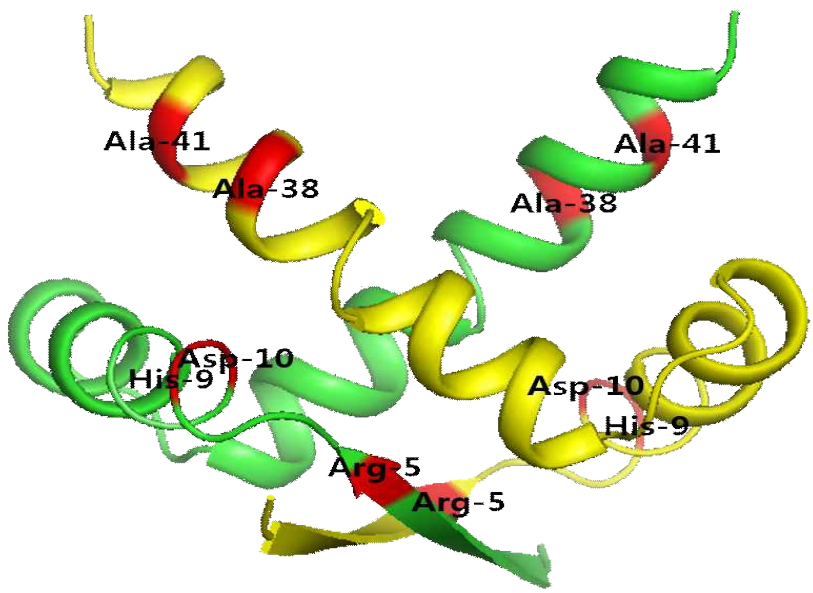

Figure 12. The crystal structure of Rv0660c$^{1-44}$ within formation on DNA binding.

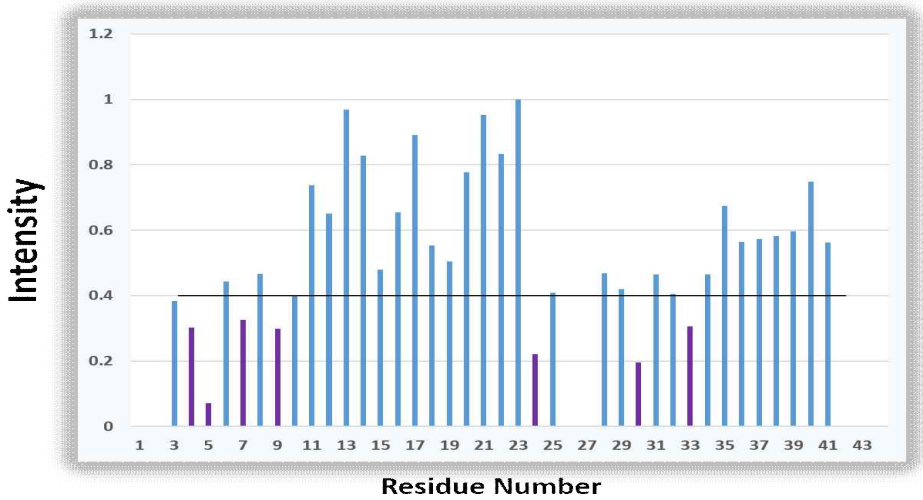

Figure 13. The intensity loss graph of Rv0660c$^{1-44}$ due to DNA binding.

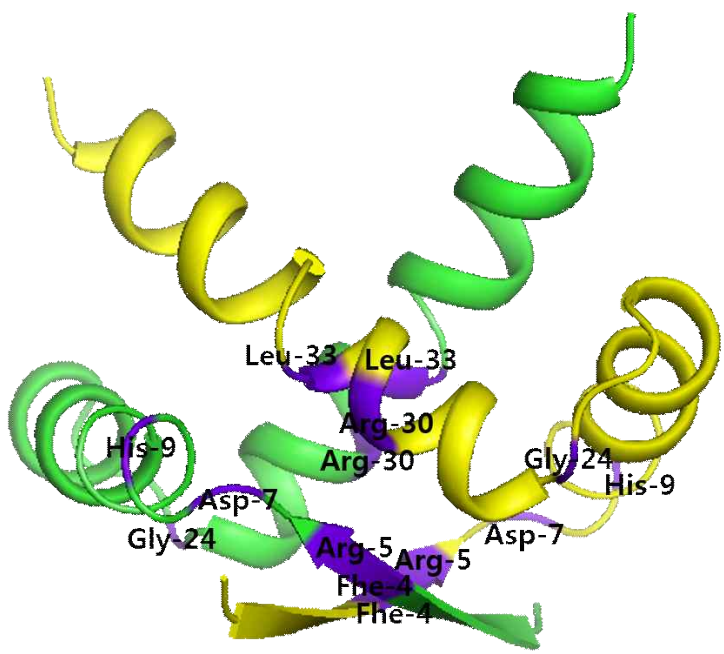

Figure 14. The crystal structure of Rv0660c$^{1-44}$ within formation on direct DAN-binding.

3.4.2 Comprehensive result of Protein-DNA binding

Thus, Rv0660c, antitoxin, can be seen that DNA will bind the side chains of the residues of the purple moiety seen in the Ribbon model of the following structure (figure 15). In the typical RHH motif of DNA-binding proteins, residues such as Lysine and Arginine, which have a particularly positive charge, are known to bind directly to the DNA backbone (figure 16). In the electrostatic charge model of the following surface, it can be seen that the part to be bound has a positive charge and that Rv0660c interacts with DNA as in the DNA binding model of RHH motif (figure 17).

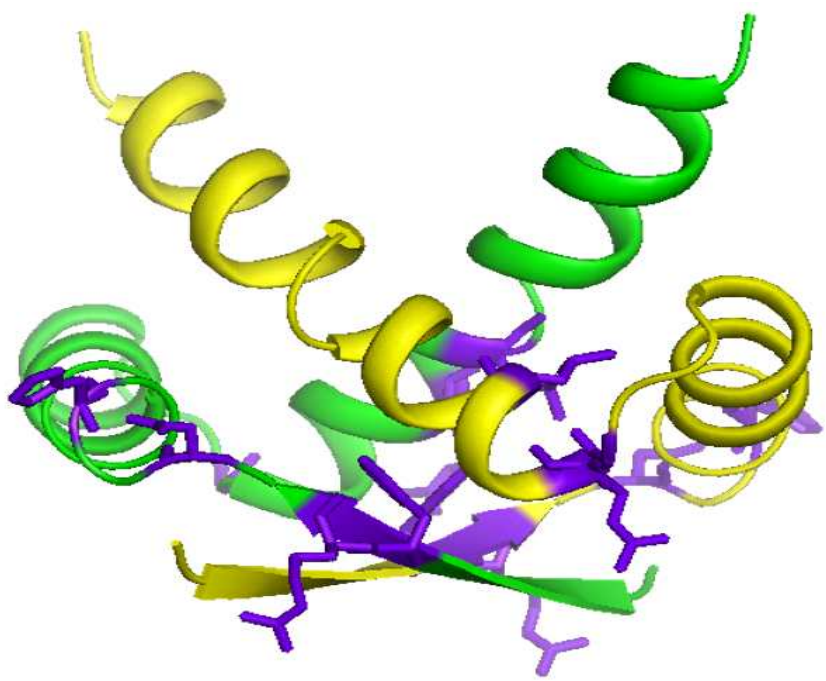

Figure 15. The ribbon model of Rv0660c[1-44].

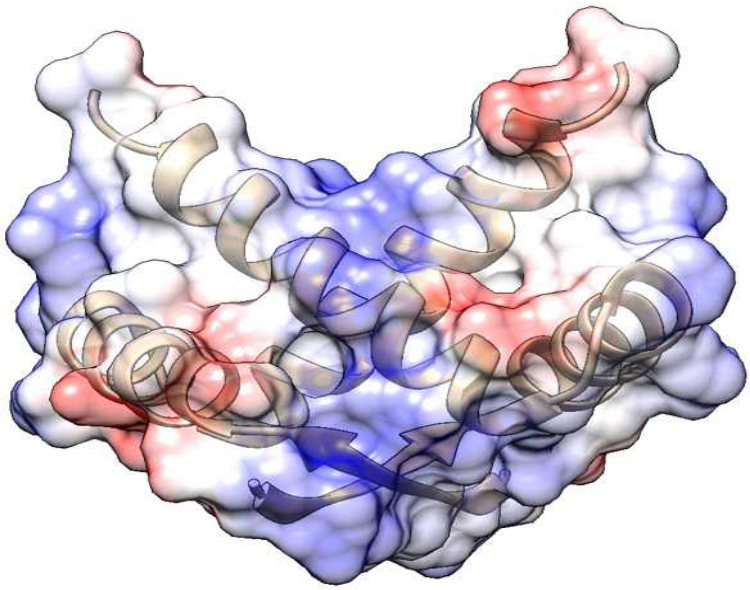

Figure 16. The electrostatic charge model of Rv0660c[1-44].

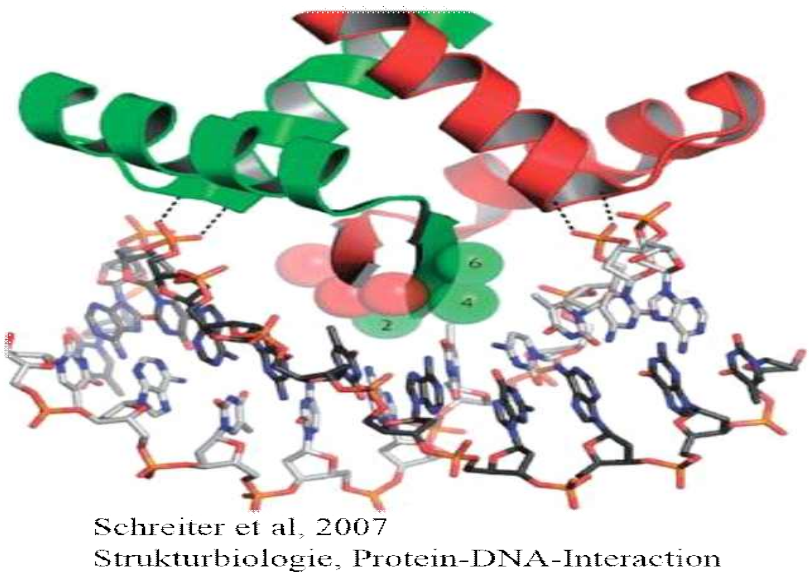

Figure 17. The DNA binding model on RHH motif.

3.4.3 Electrophoretic Mobility Shift Assay (EMSA)

It is an experiment to check DNA binding activity of protein in vitro. DNA was used as a probe and the protein concentration was adjusted to a gradient. And, as you know, figure2 shows The DNA sequence that we deigned. After the reaction was carried out for a certain period of time, the reaction samples were loaded on agarose gel and electrophoresis revealed that Rv0660c binds to the palindromic sequence of the promoter DNA of the MazEF gene as you can see that the bend shifts up (figure 18). This shows that Rv0660c functions as a transcriptional factor.

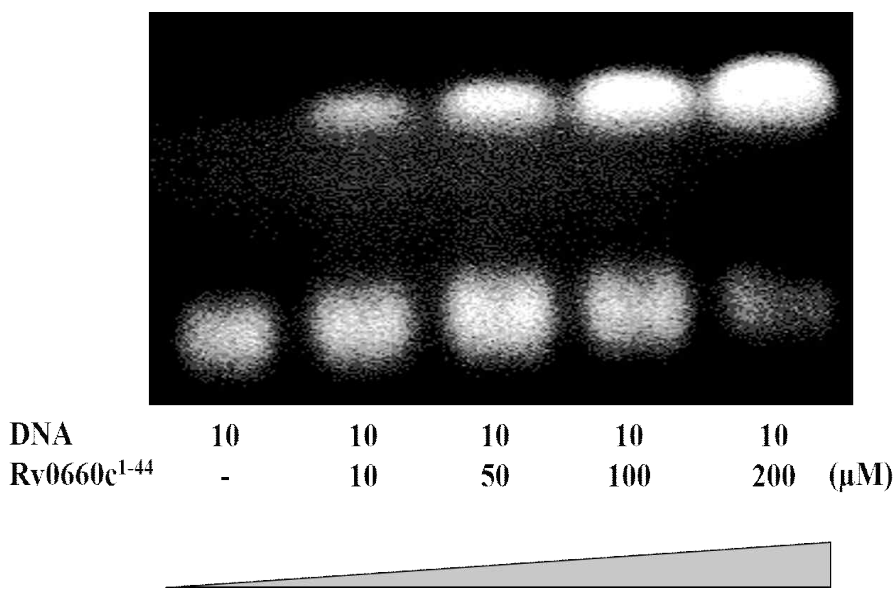

Figure 18. The agarose gel result of Electrophoretic Mobility Shift Assay of $Rv0660c^{1-44}$.

Ⅳ. Discussion

The structure of $Rv0660c^{1-44}$ was obtained by x-ray crystallography, and NMR and EMSA experiments were performed to obtain information on DNA binding. The NMR experiments showed that the secondary structure could be predicted, which is actually the same as the crystal structure. Also, through DNA titration, the backbones with greatly changed chemical shifts could be identified by the site of DNA interaction. And we could confirm the DNA binding activity of Rv0660c in vitro through EMSA.

Therefore, DNA binding information of antitoxin can be used as a part of the antibacterial strategy when TA gene promoter is combined with antitoxin-like compound to act as TA gene repressor and activate toxin to cause cell death.

Once the TA system has been studied as an antibacterial strategy through artificial activation. It is the ability to disrupt the complex through a compound that binds to a toxin or antitoxin, rather than a stress condition, and then to exert its toxin function to cause cell death. Among them, MazEF is the most popular TA. It is the most efficient way to combine with antitoxin to free toxin. This is because the way to disrupt the complex in combination with toxin depends on which part of the toxin is bound, which may affect the activity of the toxin. [Julia J. Williams and Paul J. Hergenrother., 2012]

Therefore, it is important to note the structure of antitoxin and DNA biding information when approaching new drug development by disruption of TA complexes and by inhibiting complex formation.

Ⅴ. References

Issar Smith, *et al*. (2003). *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. ***Clin Microbiol Rev***. 16(3): 463-496.

World Health Organization. (2015). Fact sheets. http://www.who.int/mediacentre/factsheets/fs104/en/. (Accessed 15 May 2015)

Yoshihiro Yamaguchi, Jung-Ho Park, and Masayori Inouye. (2011). Toxin-Antitoxin Systems in Bacteria and Archaea. ***Annual Review of Genetics***. 45: 61-79.

Kenn Gerdes and Etienne Maisonneuve. (2012). Bacterial Persistence and Toxin-Antitoxin Loci. ***Annual Review of Microbiology***. 66: 103-23.

Ambre Sala, Patricia Bordes and Pierre Genevaux. (2014). Multiple Toxin-Antitoxin Systems in *Mycobacterium tuberculosis*. ***Toxins***. 6, 1002-1020.

Minor, W., Cymborowski, M., Otwinowski, Z. (2002). Automatic system for crystallographic data collection and analysis. ***Acta. Phys. Pol.*** A 101:613-619.

P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, *et al*. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. ***Acta Cryst***. D66, 213-221.

P. Emsley and K. Cowtan. (2004). Coot: model-building tools for molecular graphics. ***Acta Cryst***. D60, 2126-2132.

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. ***J Biomol NMR***. 6(3): 277-93.

Johnson BA, Blevins RA. (1994). NMR View: A computer program for the visualization and analysis of NMR data. ***J Biomol NMR***. 4(5): 603-14.

Julia J. Williams and Paul J. Hergenrother. (2012). Artificial activation of toxin-antitoxin systems as an antibacterial strategy. ***Trends in Microbiology***. Vol. 20, No. 6.

Yurong Wen, Ester Behiels and Bart Devreese. (2014). Toxin-Antitoxin systems: their role in persistence, biofilm formation, and pathogenicity. ***Pathogens and Disease***. 70, 240-249.

Laurence Van Melderen. (2010). Toxin-antitoxin systems: why so many, what for?. ***Current Opinion in Microbiology***. 13: 781-785.

Nathalie Goeders and Laurence Van Melderen. (2014). Toxin-Antitoxin Systems as Multilevel Interaction Systems. ***toxins***. 6, 304-324.

Schreiter *et al*. (2007). Protein-DNA-interaction-Regulation of gene expression. Strukturbiologie.

Remy Loris, Irina Marianovsky, Jurij Lah, Toon Laeremans, Hanna Engelberg-Kulka, Gad Glaser., *et al*. (2003). Crystal Structure of the Intrinsically Flexible Addiction Antidote MazE. ***THE JOURNAL OF BIOLOGICAL CHEMISTRY***. Vol. 278, No.30, 28252-28257.

Christian Dienemann, Andreas Bøggild, Kristoffer S. Winter, Kenn gerdes and Ditilev E. Brodersen. (2011). Crystal Structure of the VapBC Toxin-Antitoxin Complex from *Shigella flexneri* Reveals a Hetero-Octameric DNA-Binding Assembly. ***Journal of Molecular Biology***. 414, 713-722.

Robert Schleif. (1988). DNA Binding by Proteins. ***Science***. 241, 4870.

Valentina Zorzini, Lieven Buts, Evelyne Schrank, Yann G.J. Sterckx, Michal Respondek, Hanna Engelberg-Kulka., *et al*. (2015). *Eschericia coli* antitoxin MazE as transcription factor: insights into MazE-DNA binding. ***Nucleic Acids Research***.

Martin Overgaard, Jonas Borch and Kenn Gerdes. (2009). RelB and RelE of *Escherichia coli* Form a Tight Complex That Represses Transcription via the Ribbon-Helix-Helix Motif in RelB. ***J. Mol. Biol***. 394, 183-196.

Andreas Bøggild, Nicholas Sofos, Kasper R. Andersen, Ane Feddersen, Ashley D. Easter, Lori A. Passmore, et al. (2012). The Crystal Structure of the Intact *E. coli* RelBE Toxin-Antitoxin Complex Provides the Structural Basis for Conditional Cooperativity. ***Structure***. 20, 1641-1648.

Gerhard Mittenhuber. (1999). Occurence of MazEF-like Antitoxin/Toxin Systems in Bacteria**. *J. Mol. Microbiol. Biotechnol.*** 1(2): 295-302.

국문초록

결핵균(*Mycobacterium tuberculosis*)은 1882년 Robert Koch에 의해서 처음 발견되었으며, 인체 감염성 균주로서 호흡기를 통해 폐에 질병을 일으킨다. Mycobacterium은 일반적인 세균과는 달리 세포 표면에 두터운 협막을 갖고 있어 그람 염색이 되지 않으나 Ziehl-Neelsen으로 염색시 산, 알코올, 끓이는 것 등에 의해서도 탈색되기 어려우며, 이 성질을 항산성(acid-fast)이라 한다. 그람 염색 여부로는 음성이지만, 이들은 외부 세포막이 없어서 항산성 그람 양성 세균으로 분류된다. 결핵균에 감염된 환자의 치료는 주로 약에 의해 이루어지며 쓰이는 약물로는 이소니아지드, 리팜핀, 피라진아미드 등이 쓰이며 스트렙토마이신 등이 쓰이기도 한다.

결핵균에는 세균 중 가장 많은 Toxin-Antitoxin system 쌍이 존재한다. Toxin-Antitoxin system은 antitoxin 성질에 따라 여섯 가지 종류로 분류되며 본 연구의 타겟 단백질인 Rv0660c는 Type2 Toxin-Antitoxin system의 하나인 MazEF system의 antitoxin인 MazE다.

Toxin-Antitoxin system은 외부 자극에 의해 활성화된 toxin이 균의 성장을 저해하거나 사멸로 이끄는 역할을 한다. Antitoxin은 평상시 toxin과 결합하여 toxin의 활동을 저해하지만 극한상황이 되었을 때 분해된다. 이로 인해 해리된 toxin이 성장 저해, 사멸 등을 유도한다. 이는 또한 잠복기를 유도하여 항생제에 감수성을 떨어뜨리게 만든다.

본 연구는 이 Rv0660c 단백질의 삼차구조를 규명하기 위해 Rv0660c의 N1-44 construct를 과 발현시켰다.

Immobilized Metal Affinity Chromatography(IMAC)를 이용하여 His-tag을 붙인 Rv0660c 단백질을 정제하였고, Thrombin cutting을 거쳐 His-tag을 제거하여 단백질의 순도를 더욱 높여 만들어낸 crystal로 1.69Å의 높은 resolution으로 구조를 구할 수 있었다.

NMR, EMSA를 통하여 Rv0660c의 DNA binding실험을 진행하였다. Rv0660c를 ^{13}C와 ^{15}N의 동위원소로 label하여 heteronuclear multidimensional NMR spectrum을 측정하여, HNCO, HNCA, HNCACO, HNCOCA, HNCOCACB, HNCACB spectrum을 통해 backbone assignment를 하였다. TALOS 프로그램을 이용하여 2차구조를 예상했으며 DNA titration을 통해 DNA와 상호작용하는 부분을 규명하였다. Electrophoretic Mobility Shift Assay(EMSA)를 통해 *in vitro*에서 Rv0660c의 DNA 결합 여부도 확인하였다.

주요어: SBDD(Structure Based Drug Design), Toxin-Antitoxin system, Antitoxin, x-ray crystallography, NMR(Nuclear Magnetic Resonance), EMSA(Electrophoretic Mobility Shift Assay)