Flavin-dependent Thymidylate Synthase Candidate Gene MSC0676 Product in *Mycoplasma mycoides* subsp. *mycoides* Small Colony

Salah Eddin Albasoumi

Uppsala
2010

Swedish University of Agricultural Sciences
Faculty of Veterinary Medicine and Animal Science
Department of Anatomy, Physiology and Biochemistry
Section of Biochemistry

Masters Degree project (30 hp)
Section of Biochemistry

SLU
Swedish University of Agricultural Sciences

Flavin Beroende Kandidat Protein MSC0676 i Mykoplasma mycoides

Salah Eddin Albasoumi

Supervisor : Liya Wang, Department of Anatomy, Physiology and Biochemistry, Section of Biochemistry

Examinator : Staffan Eriksson, Department of Anatomy, Physiology and Biochemistry, Section of Biochemistry

Master thesis in Animal Sciences, Uppsala 2010
Faculty of Veterinary Medicine and Animal Science
Department of Anatomy, Physiology and Biochemistry

Course code: EX0562, Advanced E, 30hp

Key words: Flavin-dependent thymidylate synthase (FDTS), *Mycoplasma mycoides* subs. *mycoides* small colony (MmmSC), Recombinant protein expression; nucleotide phosphotransferase

Online publication of this work: http://epsilon.slu.se

Uppsala 2010

Table of contents

Abbreviations

FDTS; flavin-dependent thymidylate synthas

DHFR; dihydrofolate reductase.

TK; thymidine kinase.

TS; Thymidylate synthase.

ThyX; flavin-dependent thymidylate synthase.

dUMP; 2′-deoxyuridine monophospate.

MTHF; 5,10-methylenetetrahydrofolate.

FAD; flavin adenine dinucleotide. .

NADH; nicotinamide adenine dinucleotide.

AMP; Adenosine monophosphate.

GMP; Guanosine monophosphate.

CMP; Cytidine monophosphate.

UMP; Uridine monophosphate.

IMP; Inosine monophosphate.

dAMP; Deoxyadenosine monophosphate.

dGMP; Deoxyguanosine monophosphate.

dCMP; Deoxycytidine monophosphate.

dUMP; Deoxyuridine monophosphate.

dTMP; Deoxythymidine monophosphate.

dIMP; Deoxyinosine monophosphate.

dThd; Deoxythymidine.

dAdo; Deoxyadenosine.

NADPH; nicotinamide adenine dinucleotide phosphate.

CDP; Cytidine diphosphate.

GDP; Guanosine diphosphate.

UDP; Uridine diphosphate.

TDP; Thymidine diphosphate.

dTTP; 2'-deoxythymidine triphosphate.

dATP; 2'-deoxyadenosine triphosphate.

dCTP; 2'-deoxycytidine triphosphate.

dGTP; 2'-deoxyguanosinetriphosphate.

dTDP; 2'-deoxythymidine diphosphate.

DHFR; dihydrofolade reductase

ADP; Adenosine diphosphate.

dGuo; Deoxyguanosine.

Abstract

Flavin-dependent thymidylate synthase FDTS is essential in the *de novo* thymidylate synthesis for many microorganisms. To characterize FDTS MSC0676 candidate gene products, in *Mycoplasma mycoides* subs. *mycoides* small colony (*MmmSC*), the gene was cloned and studied biochemically. The encoding sequence was fragmented in four parts due to the predicted secondary structure, cloned, and over-expressed, with Trx*Tag thioredoxin protein and the His tag at the N-terminal. The fragments were subcloned into the pET-32a vector which was transformed into *Escherichia coli* Top-10 strain. The E. coli strain for protein expression was BL21_codon plus (DE3) RIPL. Thrombin was used to remove the extra tags, and a nickel affinity column to purify the proteins. SDS-PAGE and western blot were used to monitor purity. Enzymatic activities were studied by using spectrophotometric and radiochemical assays. This study demonstrated high expression and purification yield of three of the four fragments. Characterization of the protein revealed ATPase, GTPase and phosphotransferase activities in case of the two N-terminal fragments. These findings indicate new and important functions for the MSC0676 gene products in energy and metabolism. This gene may be a future antibacterial drug target.

Introduction

Basic Mycoplasma molecular biology

There are 204 investigated mollicutes species, 119 of which are under genus mycoplasma. Mycoplasmas are host and tissue specific species. Mycoplasma species are cell wall-less bacteria (Bizarro and Schuck, 2007). Cloning of mycoplasmas genes is not easy, because mycoplasma unlike almost all other bacteria. Use the stop codon UGA to encode the tryptophan amino acid. For instance, when a mycoplasma gene is cloned in *Escerichia coli*, the translation will stop at the UGA codon, which leads to truncated proteins (Nicholas and Bashiruddin, 1995).

Mycoplasma mycoides subs. *mycoides* small colony (*MmmSC*) belongs to the *Mycoplasma mycoides* group. The *MmmSC*'s plasma membrane consists of galactan. The fibril-structure of *MmmSC* consists of lipoglycans which may have an important role in adhesion (Nicholas and Bashiruddin, 1995). Mycoplazma bacteria are considered to be the smallest self replicating microorganisms. Their genome sizes range between 580 kb for *Mycoplasma genitalium* to 1358 kb for *Mycoplasma penetrance* (Pilo et al., 2007), all with low guanine contents (Bizarro and Schuck, 2007).

Properties of Some Species of Mycoplasma

Phylogeny studies of mycoplasma showed a close relation of these bacteria to the *Bacillus-Lactobacillus-Streptococcus* subgroup, which are considered as gram positive bacteria with low guanine content in their DNA. It is also close related to the *Clostridium innocuum* and *Clostridium ramosum* group (Nicholas and Bashiruddin, 1995).

Mycoplasma mycoides subs. *mycoides* small colony (*MmmSC*) is the causative agent of contagious bovine pleuropneumoniae (CBPP). The disease is considered as the most severe infectious disease in cattle, with a big economical importance. Low mortality with acute inflammatory lesions in lung, high morbidity and low performance are associated with this disease (Totte et al., 2007; Tweyongyere and Nicholas, 1999).

Mycoplasma hyopneumoniae (*Mhp*) is the etiological factor for enzootic pneumonia. It causes severe chronic respiratory infection in pigs (Calus et al., 2007). The disease is highly prevalent in almost all pigs in meat production industry. *Mhp* is also involved in the porcine respiratory disease complex. The increased need for medication and decreased production yield results in high economic losses. *Mhp* adheres to the ciliated epithelium and destroys it and causes infection (Maes et al., 2008). The principal symptom of infection with *Mhp* is coughing (Ameri et al., 2006).

Mycoplasma capriculum (*Mcap*) in goats is the etiological agent of caprine pleuropneumonia. The disease may lead to mortality in up to 70% of a herd (Ezzi et al., 2007). The most important symptom is severe arthritis. Fever is a symptom in young goats (Wesonga et al.,

2004). Naseem and his coworkers (2009) have described that *Mcap* leads to a CBPP-like disease in cattle.

Ureaplasma urealyticum (*Uu*) causes an inflammation of the urogenital tract of humans, abortion, septicemia, meningitis, and pneumonia in offspring. It is considered as an opportunistic pathogen for pregnant women and causes human non-gonococcal urethritis (Xie and Zhang, 2006)

mRNA and Proteins Profiles

Transcriptional profiling studies of *Mhp* during the infection period have been done to investigate potential virulence related genes. This study showed that 79 *Mhp* genes were differently expressed and were up-regulated or down-regulated (Madsen et al., 2007). Transcriptional studies of *Mhp* during heat shock have been done by Madsen et al., (2005), with temperature shift from 37°C to 42°C. This study identified 91 genes that were up or down- regulated in response to the heat shock, more than half of them with unknown functions.

Calus et al., (2007), described mechanisms involved in altering *Mhp* transcription profiles, by doing a comparison of the genome in three different *Mhp* strains. They showed that there are 12 proteins which differ in their number of tandem amino acid repeats, resulting in a variation in structure, physiochemical and antigen characteristics. The changes in the number of tandem amino acid repeats were caused by genetic processes i.e. DNA polymerase slippage during DNA synthesis.

Ferraz and his colleagues (2000) investigated the difference in protein profiles between the *Mycoplasma gallisepticum* (*Mg*) strains (wild-type S6 (208) and a vaccine-type F-K810) by the SDS-PAGE method, and found a specific peptide band for the vaccine-type Mg F-K810 strain which was called p75. From these results they concluded that the *Mg* genome is subjected to changes producing novel antigens which do not exist in the reference strain.

Protein profiles of the house finch (*Carpodacus maxicanus*) cells as host for *Mg* were determined identifying a large number (around 100) of new transcripts which were up or down regulated. These results lead to a new understanding of the molecular co-evolution studies of *Mg* and its host (Wang et al., 2006).

Mycoplasma Pathogenicity

Pilo et al., (2007), reported that the pathological consequences of *MmmSC* are severe inflammations in the lungs of cattle. It is hypothesized that the adhesion of mycoplasma to the host cells creates a cascade of signals, which are transmitted to the host cells and cause inflammation. *MmmSC* adheres tightly to the epithelial cells surface but without any penetration into them.

The virulence of other pathogenic bacteria is due to toxins, cytolysins and invasions. Previously there were no identified genes noted as putative virulence factors in the sequenced mycoplasma species genomes (Pilo et al., 2007). Recently, Kannan and his colleagues identified Community Acquired Respiratory Distress Syndrome toxin responsible toxic gene (MPN_372) of *Mycoplasma pneumonia*, which ensured its survival and to cause disease in the host. The mycoplasma membrane also plays an important role in virulence (Pilo et al., 2007). These characteristics may give the ability of mycoplasma to invade the host cells without stimulating the immunity system (Nicholas and Bashiruddin, 1995).

Prevention and Treatment of Mycoplasma

Attempts to prevent mycoplasmas infection have been described in many studies. Management practices are of great interest and should be employed to reduce the infection symptoms. Production systems that depend on all-in, all-out regimes are important; it reduces the stress of the animal, making gaps between herds, and gives the opportunity to the farmer to clean all the animal facilities. Early weaning less than 3 weeks of piglets are considered as an important practice to reduce the transmission of the microorganism to the offspring, although, it is not allowed in the European Union. Purchasing of animals from a herd with known health status is important, as is adequate space for each animal. Biosecurity considerations like hygiene, insects and rodents control, prevents cross-contamination by reducing the overlap between different environments. Well managed animal's house condition is very important to reduce the stress (Maes et al., 2007).

A preventive measure like vaccination against Mhp (Meyns, et al., 2006) has shown unexpected results. Vaccinated pigs with conventional vaccine do not influence the establishment of lung infection. The most common way to reduce the spread of the CBPP is by vaccination, but the problem was that their data showed many CBPP cases in vaccinated herds (Mbulu et al., 2004). The best way to control the disease is by slaughter of *MmmSC* positive tested animals (Tweyongyere and Nicholas 1999).

Niang et al., (2006), reported that experimentally *MmmSC* infected cattle produced IgA locally, and suggested that the immune response against CBPP should be monitored. Increased levels of CD4 Th1-like T-cells was noted in the serum of CBPP recovered animal over long period (Pilo et al., 2007).

The susceptibility of three strains of *MmmSC* to tilmicosin, enrofloxacin, tylosin and florfenicol were tested *in vitro*. All these antimicrobial agents showed inhibitory effects with superiority for tilmicosin, and it was considered to be mycoplasmacidal. On the other hand, the authors highlighted the needs for *in vivo* and molecular studies to confirm the findings (Tweyongyere and Nicholas, 1999). Xie and Zhang (2006), studied the resistance of *Uu* to antimicrobial agents by conducting an *in vitro* tests of *Uu* isolates from 804 patients. They concluded that the resistance of *Uu* to the antibiotics used in the experiment, quinolones, tetracyclines, or erythromycins increases with time. The sensitivity of *MmmSC* to the antibiotics partially depends on the amount of capsular polysaccharide it produces, normally

MmmSC produces large amount of capsular polysaccharide that makes these organisms less sensitive to antibiotics (Pilo et al., 2007).

Nucleotides Metabolism

By using two different pathways, nucleotides can be synthesized in cells. One of these pathways is called de novo pathway, where NH_3, CO_2, amino acids, and ribose phosphate compounds are combined to form a nucleotide. The other pathway which is called the salvage pathway, use the free bases and nucleosides from degraded nucleic acids and converts them to nucleotides (Stryer et al., 2007).

Deoxyribonucleotide monophosphates (dN), the constructing units of DNA are also used to repair DNA. In the *de novo* pathway, the precursors for dNs are the ribonucleotides, the reaction is catalyzed by ribonucleotide reductase (RNR). In that reaction, the ribonucleoside diphosphate are reduced to deoxyribonucleoside diphosphate (Wehelie, 2006).

In the salvage pathway, the dNs are phosphorylated to deoxyribonucleoside monophosphate (dNMP), the enzymes catalyzing this reaction are the deoxyribonucleoside kinases. After that phosphorylation of dNMPs to diphosphates and triphosphates occur in order to use them in DNA synthesis and repair (Wehelie, 2006).

Two types of deoxynucleoside kinases are found in mycoplasmas. Deoxyadenosine kinase (dAK) is one of them, which phosphorylates deoxyadenosine, deoxyguanosine and deoxycytedine, with superiority to deoxyadenosine. The other one is thymidine kinase (TK) which phosphorylates thymidine, deoxyuridine and nucleoside triphosphates are needed as phosphate donor. Mononucleotide synthesis in mycoplasmas is initiated from nucleobases in two pathways. In a one step irreversible essential reaction with phosphoribosylpyrophosphate (PRPP) mediated by phosphoribosyltrasferase, or in two steps with reversible reactions through the production of nucleosides (Wehelie, 2006).

Flavin-dependent Thymidylate Synthase

Thymidylate synthase (TS) or ThyA is an essential enzyme in folate metabolism by converting 5, 10-methylenetetrahydrofolate to dihydrofolate. TS provide nucleosides for DNA synthesis. Conversion of deoxyuridine monophosphate to deoxythymidine is the role of this enzyme (Shi et al., 2005).

Flavin-dependent thymidylate synthase (FDTS) has been identified i.e. in *Paramecium bursaria* Chlorella Virus-1 (PBCV-1) by A674R open reading frame analysis; it showed around 53% identity of the poly peptides sequence to the ThyX which is recently discovered. TS and FDTS both catalyze the production of dTMP from dUMP. The differences are no similarity in these sequences and structures. Dihydrofolate reductase (DHFR) is an enzyme associated with TS and it produces tetrahydrofolate by reducing dihydrofolate using NADPH (Koehn and Kohen, 2010). One more difference is that TS uses 5, 10-methylenetetrahydrofolate as a source for methylene group and hydrogen but FDTS uses it only as a source of methylene groups. FDTS enhances the production of thymidylate (dTMP),

only when reduced pyridine nucleotides and oxidized FDA are available (Graziani et al., 2004); dUMP, NADPH, FAD, and methylenetetrahydrofolate are required for FDTS (Ulmer et al., 2008). FDTS was not only identified in PBCV-1 but also in many double-stranded DNA viruses and several pathogenic bacteria as reported by Graziani and his colleagues (2004).

ThyA has a homodimer structure; each subunit has one active site (Koehn and Kohen, 2010). Studies have been conducted to identify the active site of the ThyX in *H. pylori* and it was demonstrated that the active site is located at the interphase of the ThyX homotetramer (Graziani et al., 2004). The gene that codes ThyX is the *thyX* gene, and it exists in approximately 30 % of all microorganisms. Microorganisms depend on one of the thymidylate synthase enzyme families, with some exceptions (Koehn and Kohen, 2010). MmmSC, Mhp, Mcap and Uu have only the ThyX enzyme (Wehelie, 2006) and ThyX is a target for chemotherapeutic agents. Finding of a ThyX inhibitor is of great interest. Each substrates of dTMP has its own binding site on the ThyX. There is also particular interest for the partially overlapping 5, 10-methylenetetrahydrofolate binding site either with the binding site of dUMP or with the isoalloxazine ring of FAD. One suggestion is to find a subclass of particular ThyX mechanism-based inhibitors which can stop ThyX activity by inhibiting the oxidation of NAD(P)H (Graziani et al., 2004).

Structural study of the FDTS by Ulmer et al., (2008) in *Mycobacterium tuberculosis* has shown that a serine residue was acting as the nucleophile and histidine as the catalytic residue. Another study with *Helicobacter pylori* FDTS showed that histidine48, arginine74, and serine84 are functionally essential residues, forming a single active site in the ThyX tetramer (Leduc et al., 2004). A large number of microorganisms require FDTS to synthesis thymidylate (deoxythymidine 5'-monophosphate) which is essential for replication of in the cell (Leduc et al., 2004).

The aim of this master degree project is to study the gene of MSC0676 products from *MmmSC* which in earlier studies by Wehelie (2006) have been identified as a FDTS candidate gene by using biochemical and immunological methods.

Materials and Methods

Bacterial Growth and Protein Expression

The *Escherichia coli* strain that was used for cloning of the gene and propagation of the plasmid was Top-10 (Invitrogen), and for protein expression was BL21_codon plus (DE3) RIPL (Stratagene). The plasmid used as cloning and expression vector was the pET-32a vector (Novagen). Medium for bacterial growth was Luria-Bertani broth (LB) with ampicillin (100 mg/ml) for initial bacterial culture and the medium for the protein expression consists of 700 ml superbroth (35g tryptone, 20 g yeast extracts, and 5 g NaCl), 70 ml 10X M9 salts, 0.7 ml riboflavin (10 mM), 0.7 ml $MgSO_4$ (1 M), 0.7 ml ampicillin (100 mg/ml), and 3 ml glycerol (87 %). Inducing agent was 0.4 ml Isopropyl-L-thio-b-D-galactopyranoside (IPTG) 0.4 mM/liter.

The full length of MSC0676 is 872 amino acids which equal 103.65 kDa was fragmented into four fragments to facilitate the protein expression. It wasn't possible to express the protein in E. coli in full length. The fragments of MSC0676 are described in the (Table. 1) and (Fig. 1) bellow.

Table. 1: MSC0676 fragments[1].

Fragment	Weight kDa	AA[1]
F2	33.6	348-630
F3	29.1	631-872
F5	20.5	177-347
F6	21	1-177

[1]: The amino Acid sequence positions

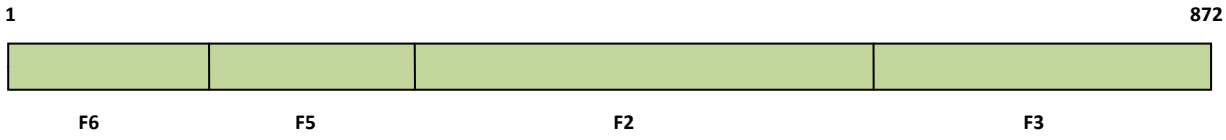


Fig.1. Schematic representation of the different MSC0676 fragments.

These fragments were cloned into the pET-32a vector (Novagen) unpublished data (Wang, 2010). This vector is commonly used to get high level expression of coding sequences inserted with 109 amino acids Trx*Tag^{TM} thioredoxin protein and His tag (Novagen). The extra tags ≈ 16 kDa present an affinity sequence for the protein fragments and to increase solubility (Stryer et al., 2007).

Around 10 ml bacteria stock was taken from -20^{o} C and inoculated to 100 ml LB and incubated overnight at 37^{o} C with 200 rpm/m shaking. Then the overnight culture was transferred to 900 ml protein expressing medium and incubated until OD600 nm reaches around 1.0 at 37^{o} C with 200 rpm/m shaking. Then changed to room temperature and 0.4 ml IPTG 0.4 mM was added as an inducing agent and incubated for 4 h to induce the recombinant protein expression. The bacterial culture was centrifuged at 4500 rpm for 20 min at 4^{o} C to harvest the bacteria. The pellets were resuspended in 20 ml 50 mM Tris/HCl pH 7.5 and 0.4 N NaCl.

Gentle sonication 5 min with pulse for every 5 sec was employed with the bacterial cells which were kept in an ice/water bath. To get the soluble proteins, centrifugation was used at 15000 rpm (Sorvall centrifuge, Avanti^{TM} J.25 Centrifuge, USA), 4^{o} C for 30 min and the supernatant was stored at -20^{o}C. The pellets were resuspended in 20 ml extraction buffer (0.4 M NaCl, 0.5% Triton X-100, and 50 mM Tris-HCl with pH 6.8, 7.0, 7.5, 8.0, 8.5, 9.0, respectively, and incubated overnight with agitation at 4^{o}C. After that the suspensions were centrifuged at 15000 rpm, 4^{o}C for 30 min and the supernatant stored at -20^{o}C.

Protein Purification

Ni Sepharose^{TM} columns from GE Healthcare Bio-Sciences AB were used for Affinity Chromatography. The column was washed with adequate amount of water and equilibrated with 10 ml of binding buffer (50 mM Tris/HCl pH 7.5, 0.4 M NaCl, 0.1% Triton, and 5 mM Imidazole). Then both the supernatants S1 and S2 were mixed and loaded on the column and flow through were collected. The column was then washed with 20 ml binding buffer. After that the column was washed with 20 ml washing buffer (same as binding buffer but with 30 mM Imidazole). The protein was eluted with 5 ml elution buffer (same as binding buffer recipes but with 300 mM Imidazole) and the fractions were analyzed by SDS-PAGE. Thrombin was used to cleave the purified protein with a ratio of 1:50.

SDS-PAGE

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) on a 12% polyacrylamide gel was used to analyse the expression level of the MSC0676 fragments in E-coli as described in ©EnCor Biotechnology Inc. 2010 protocol. The gel was run for 75 min at 150 V and 60 mA, 6 µl protein ladder (Fermentas, USA) was used as reference for protein molecular weight estimation. The gel were stained by Coomassie Brilliant Blue (CB) for 45 min and destained by boiling in water for 15 min.

Western Blot

After running the gel, polyvinylidene fluoride (PVDF) membrane (Millipore Corporation, USA) was cut and pre wetted in 100% methanol for 10 sec and washed with distilled water for 5 min and soaked in transfer buffer for 10 min as described in westernbloting.org protocol. After that wet transferring according to MitoScience protocol to PVDF membrane was employed for 1 h at 200 mA. The membrane was then blocked by using Odyssey Blocking

Buffer (OBB) 1:1 ratio 1X TBS pH 7.6 for 1 h. The membrane was incubated with Anti-His primary antibody 1:5000 in Odyssey Blocking Buffer (OBB) dilution for 2 h at room temp. Then the membrane was washed 4 times for 5 min each at room temp using TBS + 0.1% Tween-20. Fluorescently labeled donkey anti-rabbit (IRDye 800) was used as secondary antibody (1:5000 OBB) and incubated with the membrane for 60 min at room temp. Washing the membrane in the same way as after the incubation with primary antibody was performed. The membrane was rinsed with 1X TBS pH 7.6 and scanned in (Odyssey Scanner 9120, LINCOLN,NE. USA). For more details refer to the LI-CORE Biosciences protocol 2004.

Enzyme Activity Assays

Spectrophotometric Assay Coupled Enzyme System

Using the light absorbency reduction of NADH at 340 nm by as an indication of enzymatic activity for the substrate represents the main idea of the spectrophotometer assay (Bank et al., 2005). The reaction is a process where ATP reacts with the enzyme and produce ADP. ADP reacts with phosphoenolpyruvate and pyruvate kinase and produce pyruvate and ATP. Pyruvate reacts with NADH catalyzed by lactate dehydrogenaze and produce lactate and NAD^+.

The reduction in light absorbency for NADH at 340 nm was monitored in a spectrophotometer (HITACHI U200) for 4000 sec at 25^{o}C. The reactions was carried out by preparing a stock solution (150 µl of 50 mM Tris-HCl pH 7.6, 6 µl of 2 mM $MgCl_2$, 16.7 µl of 1 mM ATP, 75 µl of 0.5 mg/ml BSA, 30 µl of 5 mM DTT, 15 µl of 1 mM PEP, 15 µl of 0.1 mM NADH, 1.5 µl of 2 U/ml LDH, 2.4 µl of 2 U/ml PK, 7.6884 ml distilled water). 150 µl the stock solution was mixed with 30 µl F6 for the ATPase assay. 20 µl of GTP, CTP, dATP, dGTP, NMP, or dNMP were mixed with 120 µl stock solution and 30 µl F6 in separated reactions. ADP was used as positive control and distilled water was used as a negative control. The decrease in OD340 nm over time was recorded. The same protocol was used for the other protein fragments.

Radiochemical Assay

To detect phosphotransferase activity of the protein fragments, ^{32}P-ATP was used as the labeled substrate (phosphate donor) and Thin-layer chromatography (TLC) was used to separate the reaction products. For each reaction 4.5 µl cocktail mixture (50 mM Tris-HCl, 5 mM DTT, 5 mM $MgCl_2$, 0.5 mg/ml BSA, and 0.1 mM ATP) mixed with 0.5 µl ^{32}P-ATP and 10 µl protein solution. NMPs, dNMPs, dNs, or NDPs were added to final concentration of 1 mM. The reaction mixture was incubated in a water bath 37^{o}C for 30 min, after that 1.5 µl reaction mixture was spotted onto a TLC sheet and dried. The products were separated by using 0.2 M NaH_2PO_4 for assays with NMPs and dNs as phosphate acceptor and 0.4 M NaH_2PO_4 for assays with NDPs and dNDPs as phosphate acceptor. Then the products were analyzed by phosphoimaging analysis (FUJIFILM BIO-IMAGING ANALYZER IPR 2500).

Results

Protein Expression

The MSC0676 DNA fragments were cloned into the pET-32a vector and transformed to the BL21_codon plus (DE3) RIPL E-coli strain. Expression of all fragments was induced by adding 0.4 ml IPTG 0.4 mM to the bacterial culture. Bacteria were harvested by centrifugation. Bacteria were lysed by sonication and centrifuged to separate the soluble protein from the pellets. Proteins were extracted using extraction buffer. Expression was seen in the gel after staining with CB and showed different expression levels for all protein fragments in the supernatant and pellets. More expression of the recombinant proteins in the supernatant was seen with F2 with less intense band in the pellets, while in the pellets of F3, F5, and F6 intense bands in the pellets were seen. The results are shown in (Fig. 2).

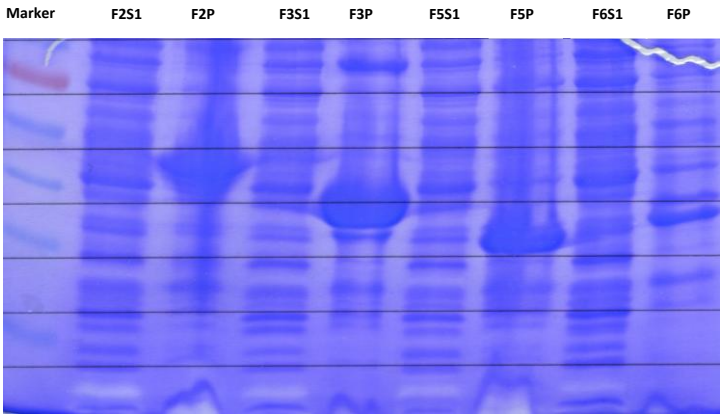


Fig.2. SDS-PAGE analysis of resuspended MSC0676 recombinant proteins; 10 µl protein was stained with CB. F= Fragment; S= Supernatant; P= Pellet.

Higher bands intensity were seen in the pellets with standard pH (7.5) of the extraction buffer in case of F2 and F3 which indicate high protein solubility. Therefore no optimization for extracting condition was needed for these fragments. F5 and F6 showed low solubility at pH 7.5. Attempts to increase the solubility were made by optimizing the extraction conditions.

Optimizing the Extraction Conditions

The pellet bands shown in Fig. 2 were obtained after resuspension of the pellets in extraction buffer containing 50 mM Tris/HCl pH 7.5 for all protein fragments. An experiment was conducted by using extraction buffers of different pH values for F5 and F6. The protein solutions were resuspended overnight at 4^{o}C with agitation.

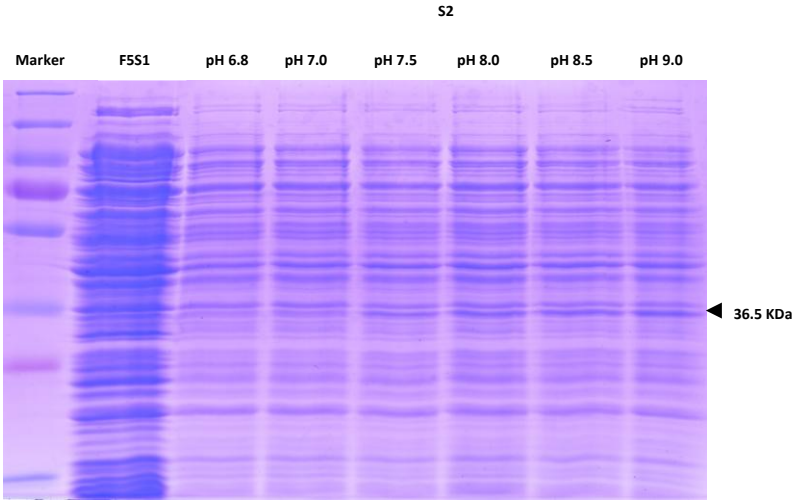


Fig. 3. F5 resuspended with different pH values in the extraction buffer. SDS-PAGE analysis of resuspended MSC0676 F5 recombinant proteins; 20 µl protein was stained with CB. F= Fragment; S= Supernatant.

SDS-PAGE analysis of the centrifuged protein solutions after resuspensio showed higher bands intensity of the 36.5 kDa F5 protein with pH between 8.0 and 9.0. Very low solubility at pH 6.8 was seen, low solubility with pH 7.0, and the solubility started to increase at 7.5 since higher band intensity was observed (Fig. 3).

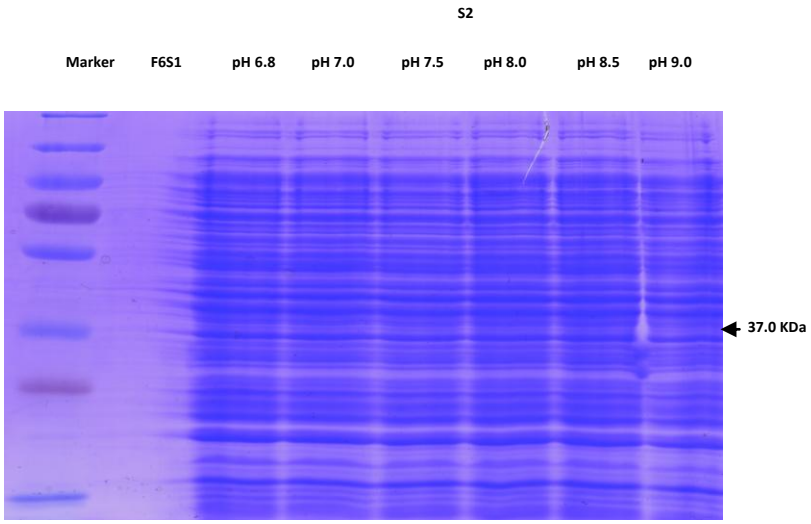


Fig.4. F6 resuspended with different pH values in the extraction buffer. SDS-PAGE analysis of resuspended MSC0676 F6 recombinant proteins; 20 µl protein was stained with CB. F= Fragment; S= Supernatant.

The solubility did not change for F6, even with extraction buffer of different pH values since no difference was found in the band intensity (Fig. 4).

Affinity Chromatography

The protein fragments were purified by Ni Sepharose™ columns as all fragments were N-terminally His tagged. The protein was bound to the Ni columns, while the other proteins were in the flow through. The imidazole substitutes the protein by binding to the immobilized metal ions (Stryer et al., 2007). The columns were washed two times and proteins eluted with elution buffer. All fragments showed good purification as seen in the stained gels (Fig. 5). The bands appeared above the expected molecular weight since they contain His and thioredoxin tags. The lowest amount of purified protein was seen with F6 and the highest with F3.

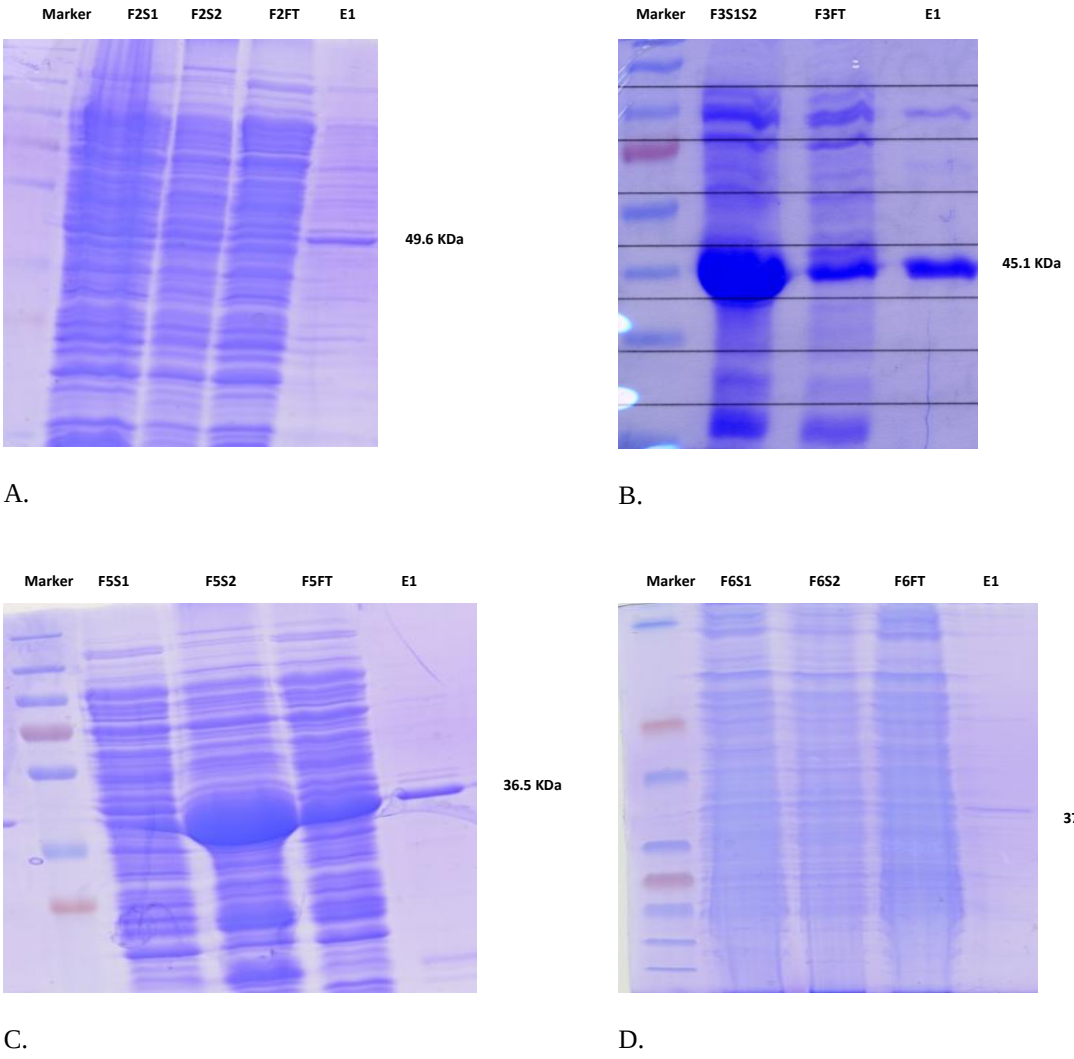


Fig.5. The F2 (A); F3 (B); F5 (C); F6 (D). Protein fragments purified by Ni Sepharose™ cromatography. SDS-PAGE analysis of purified MSC0676 recombinant proteins; 20 µl purified protein was stained with CB. The apparent molecular mass for isolated proteins was 49.6 kDa for F2, 45.1 KDa for F3, 36.5 KDa for F5, and 37.0 KDa. F= Fragment; S= Supernatant; FT= Flow Through; E= Elution.

From the bands intensity in the gels it was obvious that approximately the same amount of purified F2 and F5 protein fragments were obtained. High proportion of the total expressed

protein in E. coli was observed with F3 and F5. Low yield of the total expressed proteins were observed with the F2 fragment and very low with the F6 fragment.

Western Blot

Immunological test were performed with the expressed, purified and thrombin cleaved fragments. PVDF membranes were scanned after transferring and incubation with anti-His primary antibody and fluorescently labeled secondary donkey anti-rabbit antibody (IRDye 800). The results confirmed the results obtained in the CB stained gels and bands were seen with all four fragments (data not shown).

Detection of Enzyme Activity

Two types of assays were used to study the activity of each MSC0676 fragment. ATPase activity was determined with a spectrophotometeric assay. Experiments were conducted using ATP, GTP, CTP, NMPs, dNMPs, dATP, and dGTP as substrates.

The Spectrophotometric method is based on the reduction of NADH to NAD^+ by monitoring the absorbance at 340 nm. After mixing purified MSC0676 F5 protein with NTPs and dNTPs, and loading the mixture in the spectrophotometer, the reaction showed reduction in light absorption which indicates that significant ATPase activity (1.4- 1.6 µmol/min) and no other NTPase or dNTPase activity was detected as shown in (Fig. 6).

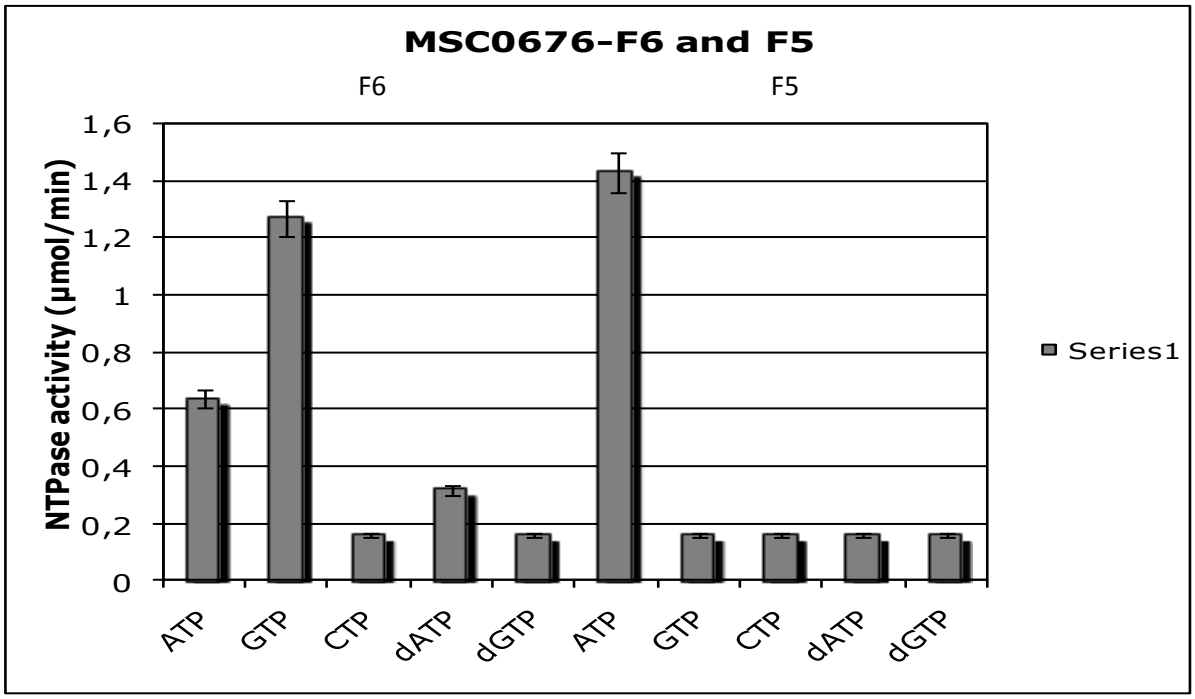


Fig. 6. ATPase assay with MSC0676 F5 and F6 with NTP & dNTP. The reduction in light absorption measured as µmol/min at adding any substrate for 4000 sec time course at 25°C. 16.7 µl ATP and 20 µl for other substrates were used.

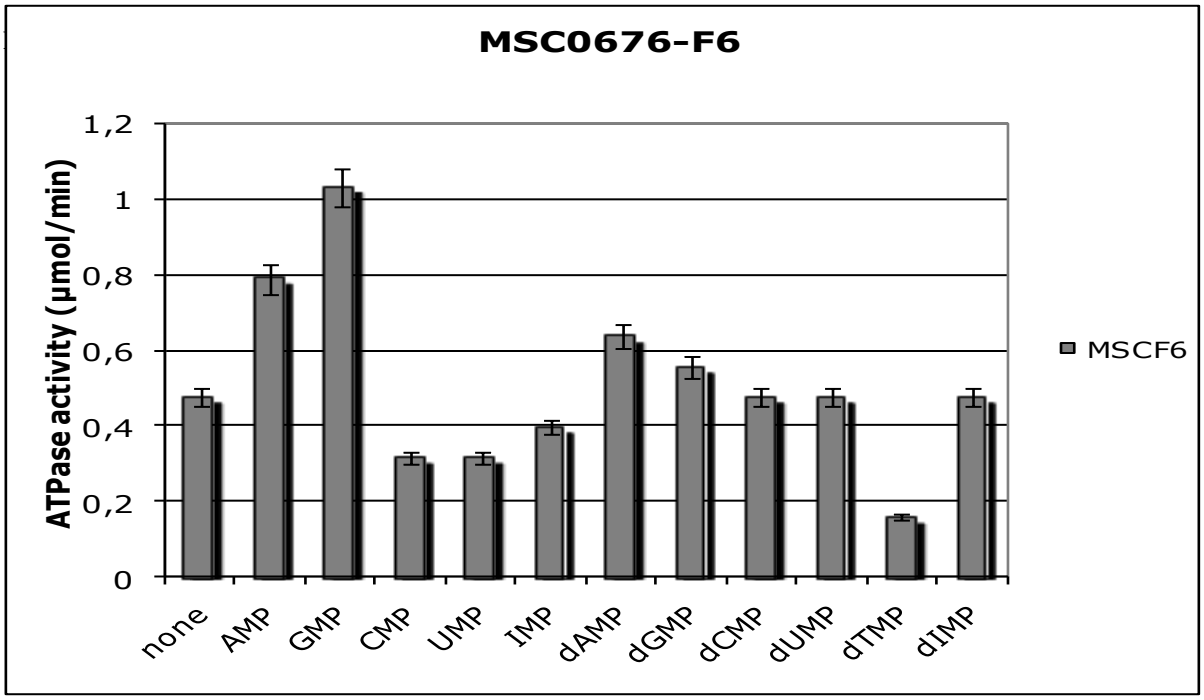


Fig. 7. ATPase assay with MSC0676 F5 and F6 with different NTP & dNTP substrates. The reduction in light absorption measured as µmol/min at adding any substrate for 4000 sec time course at 25°C. 16.7 µl ATP and 20 µl for other substrates were used.

When NTPs and dNTPs were mixed with purified MSC0676 F6 protein, significant GTPase activity (1.0- 1.3 µmol/min) and some ATPase activity (0.4- 0.7 µmol/min) were found. For F5 no activity with NTPase or dNTPs were detected (Fig. 5). For F6 some NMPs like GMP

and AMP stimulated the ATPase activity, while no effects were observed with the other NMPs or dNMPs tested (Fig. 7).

Another experiment was conducted using a radiochemical assay with radio-labeled ^{32}P-ATP as phosphate donor to detect phosphotransferase activities. MSC0676 F5 was tested with AMP, GMP, CMP, UMP, IMP, dAMP, dGMP, dCMP, dUMP, dTMP, dIMP, dThd, dAdo, dGuo, ADP, CDP, GDP, UDP, and TDP. F5 showed phosphotransferase activity with AMP, dAMP, IMP, dCMP, NDP, CMP, and UMP. No phosphotransferase activity was detected with the other substrates. A strong signal with AMP demonstrated that radio labeled phosphate group transfer to AMP and converted it to ADP. Strong signal with dAMP also showed phosphate group transfer to dAMP to produce dADP. Weaker signals were seen with IMP, dCMP, NDP, CMP, and UMP. ATPase activity was also seen by phosphoimaging analysis with all substrates as free ^{32}P-ATP was detected (Fig. 8).

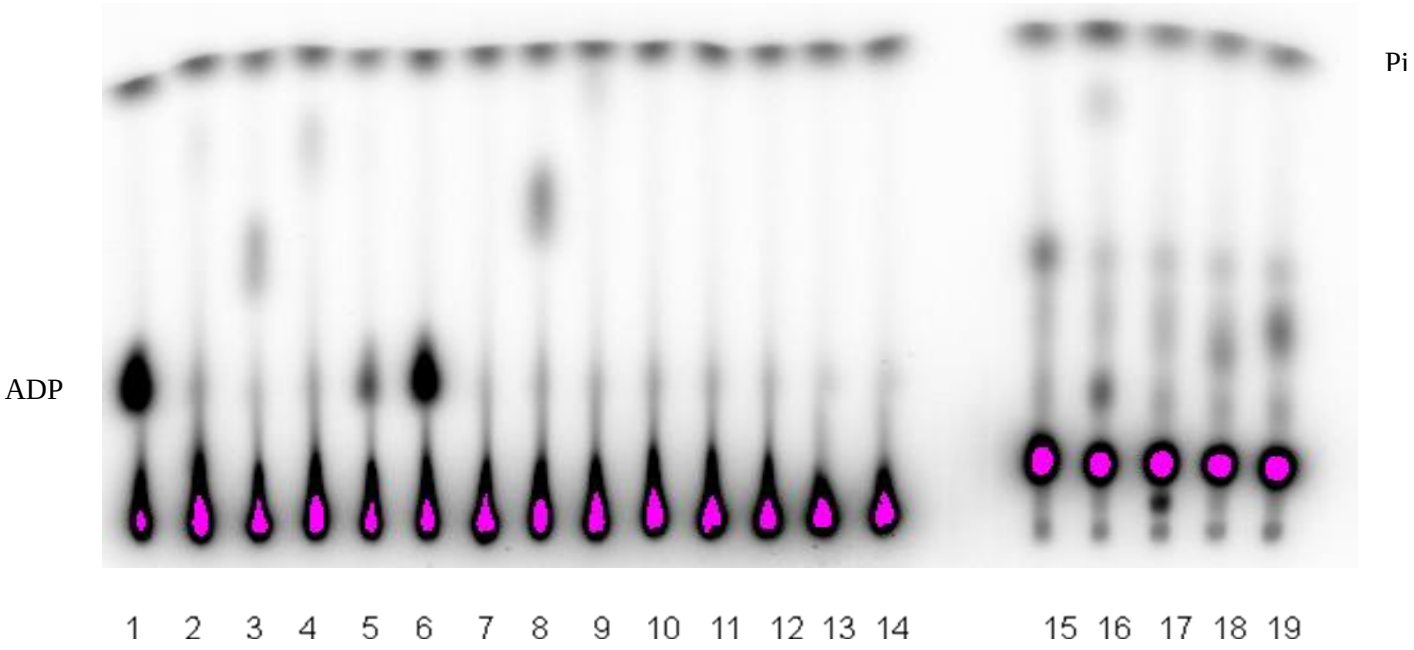


Fig. 8. Phosphotransferase assay with MSC0676- F5. Phosphoimaging analysis shows the effect of different substrate on the reaction. 1. AMP, 2. GMP, 3. CMP, 4. UMP, 5. IMP, 6. dAMP, 7. dGMP, 8. dCMP, 9. dUMP, 10. dTMP, 11. dIMP, 12. dThd, 13. dAdo, 14. dGuo, 15. ADP, 16. CDP, 17. GDP, 18. UDP, 19. TDP.

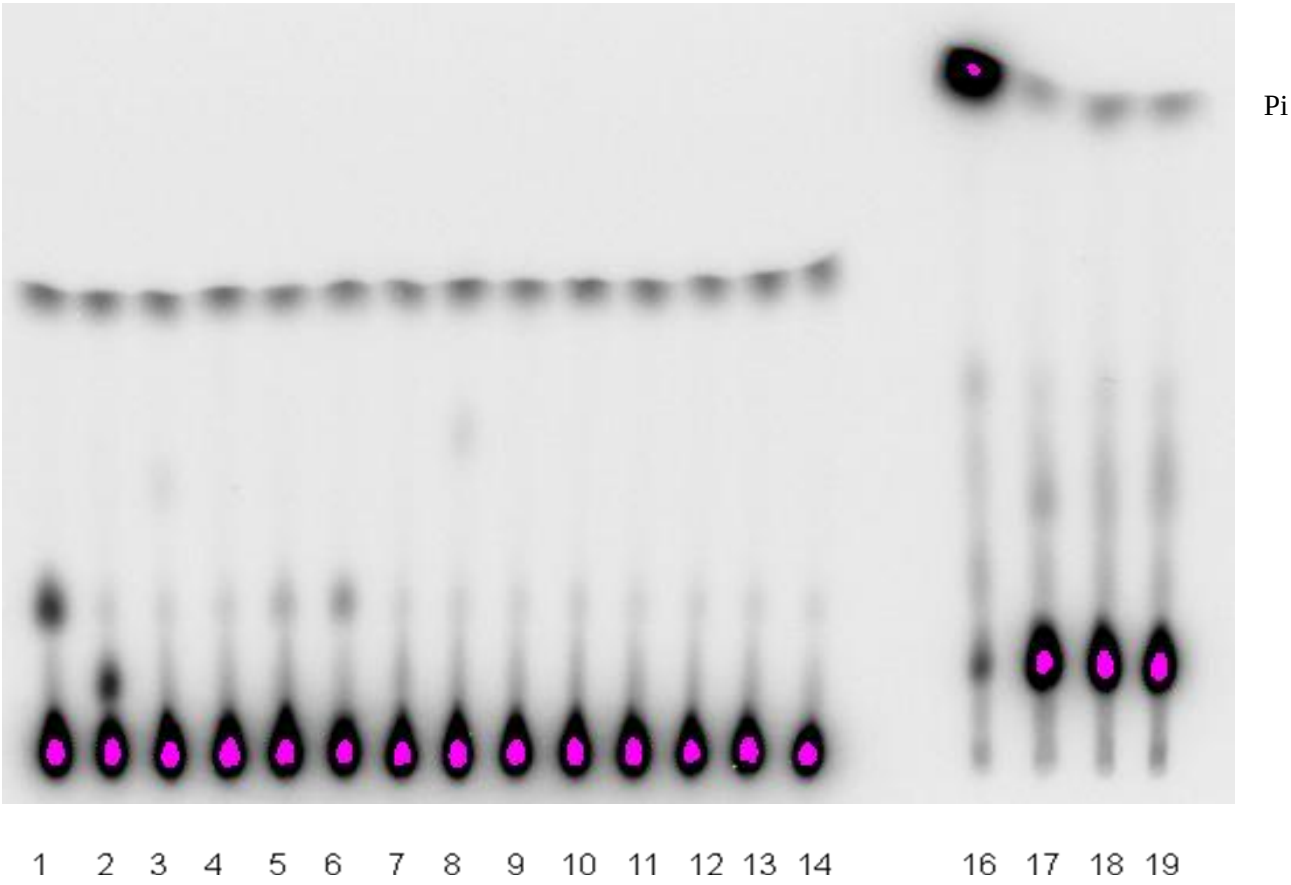


Fig. 9. Phosphotransferase assay with MSC0676- F6. Phosphoimaging analysis shows the effect of different substrate on the reaction. 1. AMP, 2. GMP, 3. CMP, 4. UMP, 5. IMP, 6. dAMP, 7. dGMP, 8. dCMP, 9. dUMP, 10. dTMP, 11. dIMP, 12. dThd, 13. dAdo, 14. dGuo, 15. ADP, 16. CDP, 17. GDP, 18. UDP, 19. TDP.

Phosphotransferase assay with MSC0676 F6 (Fig. 9) was conducted using AMP, GMP, CMP, UMP, IMP, dAMP, dGMP, dCMP, dUMP, dTMP, dIMP, dThd, dAdo, dGuo, ADP, CDP, GDP, UDP, and TDP.

The assay showed that the F6 possesses phosphotransferase activity with AMP, GMP, dAMP, NDP, and IMP in descending order, but no activity with the other substrates. Signals for ADP, GDP, dADP, NTP, and IDP as reaction products for phosphate transfer were seen by phosphoimaging analysis. ATPase activity was observed with adding any of the given substrates with high hydrolysis with CDP. Signals for free phosphate were observed with all substrates.

Discussion

Expression of the Protein Fragments

This project represents the first cloning, expression and analysis of MSC0676 from *MmmSC*. Experiments conducted in our lab with this gene did not succeed with expression of the whole protein in E. coli in considerable amounts. By using bioinformatics approaches it was identified that the protein contains high proportion of highly hydrophobic amino acid residues (Theoretical pI: 8.98) (http://www.expasy.ch/cgi-bin/protparam1?Q6MSU3@noft@) as well as its big size (872 amino acids) which motivated that the protein sequence was fragmented. These were cloned into pET-32a expressing vector, and were transformed into BL21_codon plus (DE3) RIPL E. coli strain, (unpublished, Wang, 2010). MSC0676 was cut into 4 fragments to facilitate the expression. The fragmentation was done depending on the predicted α-helixes in the secondary structure (Appendix 3) (http://bioinf4.cs.ucl.ac.uk:3000/psipred/result/37065) in order not to miss active sites or residues in the protein domains.

The same media and induction conditions were used for expression of the different recombinant protein fragments. Bacteria were harvested by centrifugation and lysed (for details see materials and methods section). The expression levels after extraction of the protein fragments were analyzed by SDS-PAGE and the majority of MSC0676 fragments were in the pellets but F2 was found more in the supernatant. Low expression level for F6 was found in E. coli. Analysis of the ATPase protein sequences using bioinformatics tools (data not shown) suggested that F6 would be toxic to the host cells, and E. coli may inhibit its expression to survive, explaining the low levels of F6 observed.

The success of expressing the protein fragments in E. coli indicate that the fragments hydrophopicity were higher lower than the whole MSC0676 protein. The tags in fussion expression vector, the E. coli strain, selected media, using the time course and the temperature shift protocol all together resulted in good expression of three of four fragments. Intense bands for fragments 2, 3 and 5 were seen in the gels. Glycerol was used in the media in order to get high bacterial density and to reduce the production of by-products at using glucose. More than 1.0 OD_{600} nm was reached for all fragment expressing bacteria before induction. Previous studies showed that use of lactose along with IPTG in the media will increase the expression of the recombinant protein up to 20 fold in comparison to media without lactose (Kim et al., 2007).

Optimization of the Extraction Conditions

Attempts to get more protein in the soluble fractions were conducted by resuspending the pellets in extraction buffer with different pH. Several studies confirmed that change in the pH of extraction buffer often results in increased protein quantity extracted from the sample. More proteins were obtained from plant tissues by using high pH extraction buffers (Woodard et al., 2009). These ideas guided to use different pH extraction buffers. SDS-PAGE analysis for F5 showed that this fraction is more soluble at higher pH. For the other protein fragments, no change in the solubility was observed.

Protein Purification

Metal-affinity chromatography using Ni Sepharose^{TM} columns was applied to purify the His tagged protein fragments. The results showed that Ni Sepharose^{TM} columns for trapping His tagged proteins are an efficient purification method to purify the MSC0676 fragments. The washing buffers that were used in purification procedure resulted in good His tagged protein retention and minimized the binding of other proteins. The proteins were eluted by 300 mM imidazole in the elution buffer. SDS-PAGE analysis showed good purification for two fractions. Clear bands were observed at 45.1 and 36.5 kDa for F3 and F5 respectively. Less intense bands were observed at 49.6 and 37.0 kDa for F2 and F6 respectively. The purification for F6 was low as its expression was also low. Complete elution of the His tagged protein was found by using the mentioned elution buffer in the materials and method section (Al-Ghobashy et al., 2009). The purified proteins were stable when kept at -20^{o}C.

Functional Study of the MSC0676 Fragments

Adenosine triphosphate (ATP) is a primary compound of animal and human cells. The majority of all metabolic processes need ATP. It is a pivotal substance for nucleic acid synthesis (Khlyntseva et al., 2009). Degrading of ATP will results in deleterious effects in tissues.

Unexpected ATPase activity and no other NTPase activity was found with F5. This could lead to detrimental effects in the host cells; it may disturb the energy balance and may retard many metabolic processes (Khlyntseva et al., 2009). From the alignment with other ATPase, it was expected that this fragment might be a ATPase. Purified MSC0676 F6 protein showed GTPase activity and ATPase activity , and no other NTPase or dNTPs activity. GMP and AMP stimulated the ATPase activity, while no effect was seen with other NMPs or dNMPs. The phosphate groups that were produced by the hydrolysis of NTPs were traced by conducting phosphotransferase assays.

Phosphotransferase activity was seen with MSC0676 F5 with AMP, dAMP, IMP, dCMP, NDP, CMP, and UMP. Phosphotransferase activity with MSC0676 F6 with AMP, GMP, dAMP, NDP, and IMP in descending order was detected. Strong ATPase activity was observed in the presence of CDP.

Amino acids 38 to 46 (GPYSSGKTS) (Appendix 1) match the highly conserved nucleotide binding Walker A sequence motife (GxxGxGKT/S) which is observed in all P-loop of the NTPases. NTPases are involved in hydrolysis, it cleaves one phosphate bond from NTPs. ATPases are very important in chemotherapy (Rangrez et al., 2010). For example, using inhibitors for Helicobacter pylori ATPase Cagalpha (VirB11 homolog) were succeed in blocking of CagA transport and cag virulence (Hilleringmann et al., 2006).

The assays revealed hydrolysis and phosphorylation activity for MSC0676 F5 and F6 fragments. These results suggest that MSC0676 is a multi-functional enzyme. MSC0676 involved in many metabolic processes. It has been investigated that *thyX* play an important role in DNA synthesis pathway (Wehelie et al., 2006).

Conclusion and Future perspective

This is the first biochemical characterization of MSC0676. The expression of the protein in fragments was successful in E. coli. Good purification was obtained by using Ni Sepharose™ columns. F6 has GTPase activity and F5 has ATPase activity and both of them have phosphotransferase activity. Future efforts to solve the three dimensional structure of this protein will help explain the phosphotranspherase activities and substrate preferences of the respective proteins and find out the function of this enzyme. We hope to see if it is involved in nucleotides metabolism and if it is good target for chemotherapy.

Acknowledgements

I would like to express my sincere gratitude, considerations and appreciation to everyone who has helped, supported me in any way during these months and made this thesis possible. I would like to thank Liya Wang, my supervisor, for her unbidden supervision throughout the project and his patience in clarifying my doubts. I wish to thank Staffan Eriksson. I am grateful to Hanan Sharif, Elena Sjuvarson, Louise Egeblad and Ren Sun for their co-operation and suggestions. Finally, thanks to Jagarlamudi Kiran Kumar for his help and suggestions during this project work.

References

Al-Ghobashya M. A, Williamsa M. A. K, Brophyb B, Laibleb G, Harding D. R. K. 2009. On-line casein micelle disruption for downstream purification of recombinant human myelin basic protein produced in the milk of transgenic cows. Journal of Chromatography B 877, 1667-1677.

Ameri M, Zhou E, Hsu W. H. 2006. Western blot immunoassay as a confirmatory test for the presence of anti-Mycoplasma hyopneumoniae antibodies in swine serum. J Vet Diagn Invest 18, 198-201.

Bizarro C. V, Schuck D. C. 2007. Purine and pyrimidine nucleotide metabolism in Mollicutes. Genetics and Molecular Biology 30, 1, 190-201.

Calus D, Baele M, Meynsb T, de Kruif A, Butaye P, Decostere A, Haesebrouck F, Maes D. 2007. Protein variability among Mycoplasma hyopneumoniae isolates. Veterinary Microbiology 120, 284-291.

De Bank P. A, Kendall D. A, Alexander S. P. H. 2005. A spectrophotometric assay for fatty acid amide hydrolase suitable for high-throughput screening. Biochemical Pharmacology 69, 1187-1193.

Ezzi A, Pourbakhsh S. A, Moradi bidhendi S. 2007. Study on Mycoplasma pneumonia at the Ziaran abattoir. Archives of Razi Institute 62, 3, 161-164.

Ferraz P. N, Menezes C. C. B, Danelli M. G. M, Lizeu J. O. P, Nascimento E. R, Lucchesi D. M. 2000. Protein Profile Analysis by SDS-PAGE of Mycoplasma Gallisepticum Strains S6(208) and F-K810 Grown in Hayflick's and Frey's Media. Brazilian Journal of Microbiology 31, 113-115.

Hilleringmann M, Pansegrau W, Doyle M, Kaufman S, MacKichan ML, Gianfaldoni C, Ruggiero P, Covacci A. 2006. Inhibitors of Helicobacter pylori ATPase Caga block CagA transport and cag virulence. Microbiology 152, 2919-2930.

Kannan T. R, Musatovova O, Balasubramanian S, Cagle M, Jordan J. L, Krunkosky T. M, Davis A, Hardy A. D, Baseman J. B. 2010. *Mycoplasma pneumoniae* Community Acquired Respiratory Distress Syndrome toxin expression reveals growth phase and infection-dependent regulation. Molecular Microbiology 76, 5, 1127-1141.

Khlyntseva S. V, Bazel Ya. R, Vishnikin A. B, Andruch V. 2009. Methods for the Determination of Adenosine Triphosphate and Other Adenine Nucleotides. Journal of Analytical Chemistry 64, 7, 657-673.

Kim M, Elvin C, Brownlee A, Lyons R. 2007. High yield expression of recombinant pro-resilin: Lactose-induced fermentation in E. coli and facile purification. Protein Expression and Purification 52, 230-236.

Koehn E. M, Kohen A. 2009. Flavin-dependent thymidylate synthase: A novel pathway towards thymine. Archives of Biochemistry and Biophysics 493, 96-102.

Leduc D, Graziani S, Lipowski G, Marchand C, Marchal P. L, Liebl U, Myllykallio H. 2004. Functional evidence for active site location of tetrameric thymidylate synthase X at the interphase of three monomers. PNAS 101, 19, 7252-7257.

Madsen L. M, Nettleton D, Thacker E. L, Edwards R, Minion F. C. 2006. Transcriptional Profiling of *Mycoplasma hyopneumoniae* during Heat Shock Using Microarrays. Infection and Immunity 74, 1, 160-166.

Madsen M. L, Puttamreddy S, Thacker E. L, Carruthers M. D, Minion F. C. 2008. Transcriptome Changes in Mycoplasma hyopneumoniae during Infection. Infection and Immunity 76, 2, 658-663.

Maes D, Segales J, Meyns T, Sibila M, Pieters M, Haesebrouck F. 2008. Control of Mycoplasma hyopneumoniae infections in pigs. Veterinary Microbiology 126, 297-309.

Mbulu R, Tjipura-Zaire G, Lelli R, Frey J, Pilo P, Vilei E. M, Mettler F, Nicholas R. A. J, Huebschle O. J. B. 2004. Contagious bovine pleuropneumonia (CBPP) caused by vaccine strain T1/44 of *Mycoplasma mycoides* subsp. *mycoides* SC. Veterinary Microbiology 98, 229-234.

Meyns T, Dewulf J, de Kruif A, Calus D, Haesebrouck F, Maes D. 2006. Comparison of transmission of Mycoplasma hyopneumoniae in vaccinated and non-vaccinated populations. Vaccine 24, 7081-7086.

Nicholas R. A. J, Bashiruddin J. B. 1995. *Mycoplasma mycoides* subspecies *mycoides* (Small Colony Variant): the Agent of Contagious Bovine Pleuropneumonia and Member of the "*mycoplasrna mycoides* Cluster". J. Comp. Path. 113, 1-27.

Pilo P, Frey J, Vilei E. M. 2007. Molecular mechanisms of pathogenicity of *Mycoplasma mycoides* subsp. *mycoides* SC. The Veterinary Journal 174, 513-521.

Rangrez A. W, Abajy M. Y, Keller W, Shouche Y, Grohmann E. 2010. Biochemical characterization of three putative ATPases from a new type IV secretion system of Aeromonas veronii plasmid pAC3249A. BMC Biochemistry 11, 10.

Shi Q, Zhang Z, Neumann A. S, Li G, Spitz M. R, Wei1 Q. 2005. Case-control analysis of thymidylate synthase polymorphisms and risk of lung cancer. Carcinogenesis 26, 3, 649-656.

Stryer L, Berg J. M, Tymoczko J. L. 2006. Biochemistry. W. H. Freeman and Company, New York pp. 28-32, 70, 709-728.

Totte P, Rodrigues V, Yaya A, Hamadou B, Cisse O, Diallo M, Niang M, Thiaucourt F, Dedieu L. 2008. Analysis of cellular responses to *Mycoplasma mycoides* subsp. *mycoides* small colony biotype associated with control of contagious bovine pleuropneumonia. Vet. Res. 39, 8.

Tweyongyere R, Nicolas R. A. J. 1997. Susceptibility of *Mycoplasma mycoides* subsp. *mycoides* SC antimicrobial agents. Tanzania Veterinary Association Proceedings 20, 192-198.

Ulmer E. U, Boum Y, Thouvene C. D, Myllykallio H, Sibley C. H. 2008. Functional Analysis of the Mycobacterium tuberculosis FAD-Dependent Thymidylate Synthase, ThyX, Reveals New Amino Acid Residues Contributing to an Extended ThyX Motif. Journal of Bacteriology 190, 6, 2056–2064.

Wang Z, Farmer K, Hill G. E, Edwards S. V. 2006. A cDNA macroarray approach to parasite-induced gene expression changes in a songbird host: genetic response of house finches to experimental infection by *Mycoplasma gallisepticum*. Molecular Ecology 15, 1263-1273.

Wehelie, R. (2006) Mycoplasma Pyrimidine Deoxynucleotide Biosynthesis. Thesis. Uppsala: Swedish University of Agricultural Sciences 91-576-7125-7, pp. 12, 15, 23, 25.

Woodard S. L, Wilken L. R, Barros G. O. F, White S. G, Nikolov Z. L. 2009. Evaluation of Monoclonal Antibody and Phenolic Extraction From Transgenic *Lemna* for Purification Process Development. Biotechnology and Bioengineering 104, 3, 562-571.

Xie X, Zhang J. 2006. Trends in the rates of resistance of *Ureaplasma urealyticum* to antibiotics and identification of the mutation site in the quinolone resistance-determining region in Chinese patients. FEMS Microbiol Lett 259, 181-186.

Xiea X, Zhaoa J, Zhenga H. 2004. Expression, purification, and isotope labeling of cannabinoid CB2 receptor fragment, $CB2_{180–23}$. Protein Expression and Purification 38, 61-68.

http://www.westernblotting.org/protocols%20electrophoresis%20transfer.htm

http://www.mitosciences.com/PDF/western.pdf

http://www.encorbio.com/protocols/SDS-PAGE.htm

http://www.proweb.org/kinesin/Methods/ATPase_assay.html

http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0016021

http://www.ebi.ac.uk/interpro/ISearch?query=Q6MSU3

http://www.expasy.ch/cgi-bin/protparam1?Q6MSU3@noft@

Appendices

1. MSC0676 Amino Acid Sequence (http://www.uniprot.org/uniprot/Q6MSU3)

```
        10         20         30         40         50         60
MPIFQDLTPN IKINDPVYDA ALDEAFQKKE IRNIALTGPY SSGKTSTWLS YSRNRKLGKI

        70         80         90        100        110        120
ITVSLGKYNN IKITSNKNKK NYYDNRIERQ IINQISSQIH PLKIPLSKYS FIKNKNLFSI

       130        140        150        160        170        180
IMNIFLSCCF IFSIIGWVHR ILLSDSLVFK ESLNWLLLIL FFVPLVIWIY KFLKNNKLYI

       190        200        210        220        230        240
SRIKFDKMEA NLEEKHLPKT TVFDKDIREI IYLLYSPNTN VVVFEDLDRF HNPKIFIKLK

       250        260        270        280        290        300
EINFILNSYI EAKIFKRKPV KFIYIINNGL FSPHIRTKFF DFIIPIVPVV NYSNSESKFS

       310        320        330        340        350        360
EFIELIGESD SLDKDIIWKI SLYVNDIRLI KNIVNEFFVY NSAIRSENNQ INKNKLFALI

       370        380        390        400        410        420
TLKNLMPKEF DLLQMNKGYV FNVFEKIQKC QKNNSEIIQN KLSNANNELS SLFKNYPEFE

       430        440        450        460        470        480
GNDKLSKTDN ILFFKDIYEL INSDLIKTNF YVIKTIFYDD TTIVEKVQLR ELRDNYPFVN

       490        500        510        520        530        540
HCSNIKKIVV YNDPKIIECY ENIESLEDQL NKKSIKFLLK SLSRDEITDL FDSTSDKQMN

       550        560        570        580        590        600
KRYVSFLKFL FIEGLIDESY LNYISYFYKG SLGINDRIFI KNLLEYVDNH NFDIKLENPK

       610        620        630        640        650        660
LVLKRLDENQ LKSDLSFNFY LFKEFVNSHN SRATMDMFNS VWPSKPQLVA KSLSQFDFDT

       670        680        690        700        710        720
INNFVNIFVR EKDGIVKIRR FLDSIHDNLV VRKYLYSGDI IKDILVSINL KFFKVFRIHF
```

```
       730        740        750        760        770        780
SDFIFNNIPD ILNYKDQHYL RKIIQSLILT GQEERVKYIN NNIDRYSLNI RNVMNMYEFI

       790        800        810        820        830        840
MGEACDFNKL LPNVFEEEKM SKIKEWIEKN FSEFVREYIE LTEIYKYYKN DESILISIIN

       850        860        870
SSEINDDYKR KYIELNKIVL LNLDKITNLK KE
```

2. MSC0676 Amino Acid Sequence Description (http://www.expasy.ch/cgi-bin/protparam1?Q6MSU3@noft@)

Number of amino acids: 872

Molecular weight: 103784.3

Theoretical pI: 8.98

Amino acid composition:

Ala (A) 12 1.4%

Arg (R) 31 3.6%

Asn (N) 90 10.3%

Asp (D) 53 6.1%

Cys (C) 6 0.7%

Gln (Q) 19 2.2%

Glu (E) 55 6.3%

Gly (G) 16 1.8%

His (H) 11 1.3%

Ile (I) 111 12.7%

Leu (L) 89 10.2%

Lys (K) 93 10.7%

Met (M) 12 1.4%

Phe (F) 67 7.7%

Pro (P) 23 2.6%

Ser (S) 67 7.7%

Thr (T) 23 2.6%

Trp (W) 7 0.8%

Tyr (Y) 44 5.0%

Val (V) 43 4.9%

Pyl (O) 0 0.0%

Sec (U) 0 0.0%

(B) 0 0.0%

(Z) 0 0.0%

(X) 0 0.0%

Total number of negatively charged residues (Asp + Glu): 108

Total number of positively charged residues (Arg + Lys): 124

Atomic composition:

Carbon C 4797

Hydrogen H 7468

Nitrogen N 1196

Oxygen O 1332

Sulfur S 18

Formula: $C_{4797}H_{7468}N_{1196}O_{1332}S_{18}$

Total number of atoms: 14811

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 104435

Abs 0.1% (=1 g/l) 1.006, assuming all pairs of Cys residues form cystines

Ext. coefficient 104060

Abs 0.1% (=1 g/l) 1.003, assuming all Cys residues are reduced

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).

>20 hours (yeast, in vivo).

>10 hours (Escherichia coli, in vivo).

Instability index:

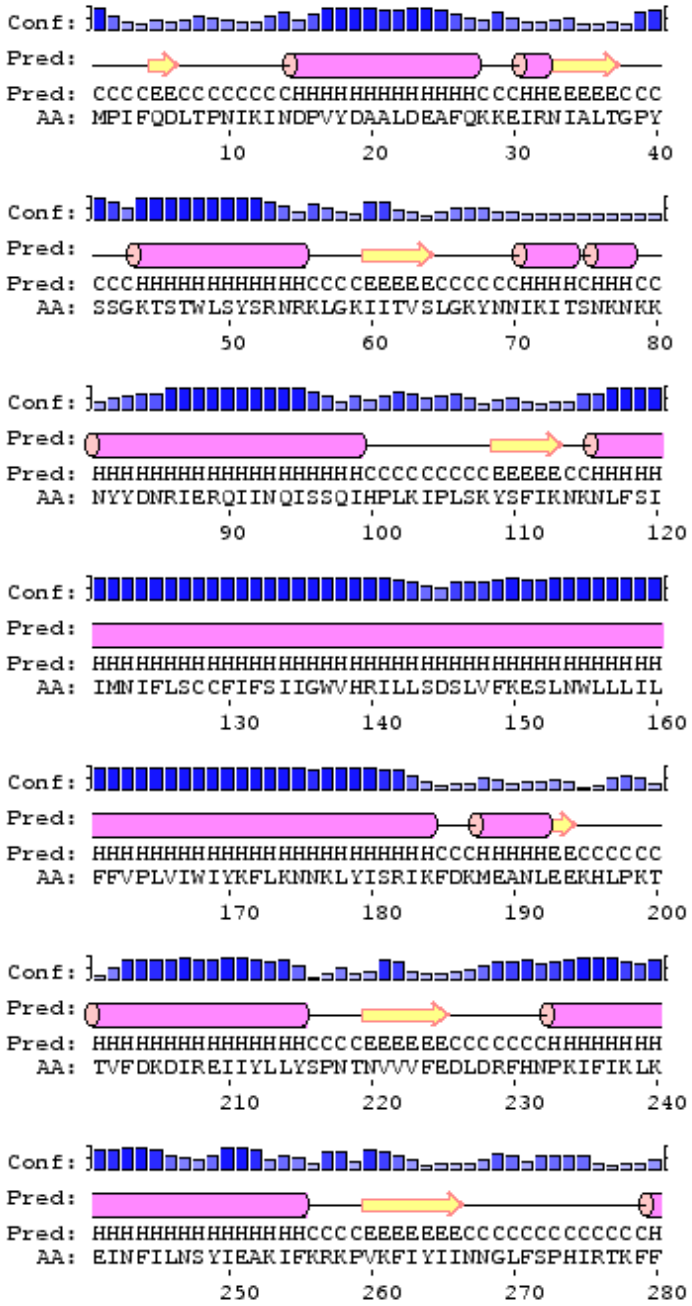
The instability index (II) is computed to be 41.21

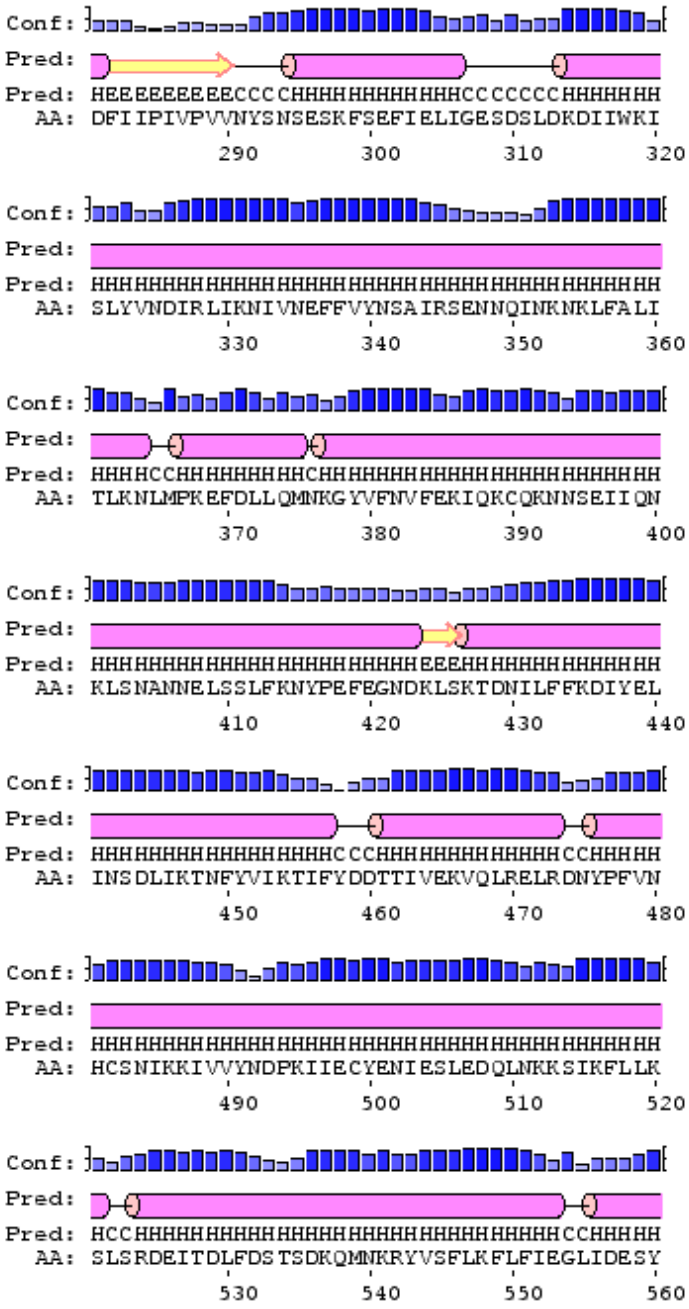
This classifies the protein as unstable.

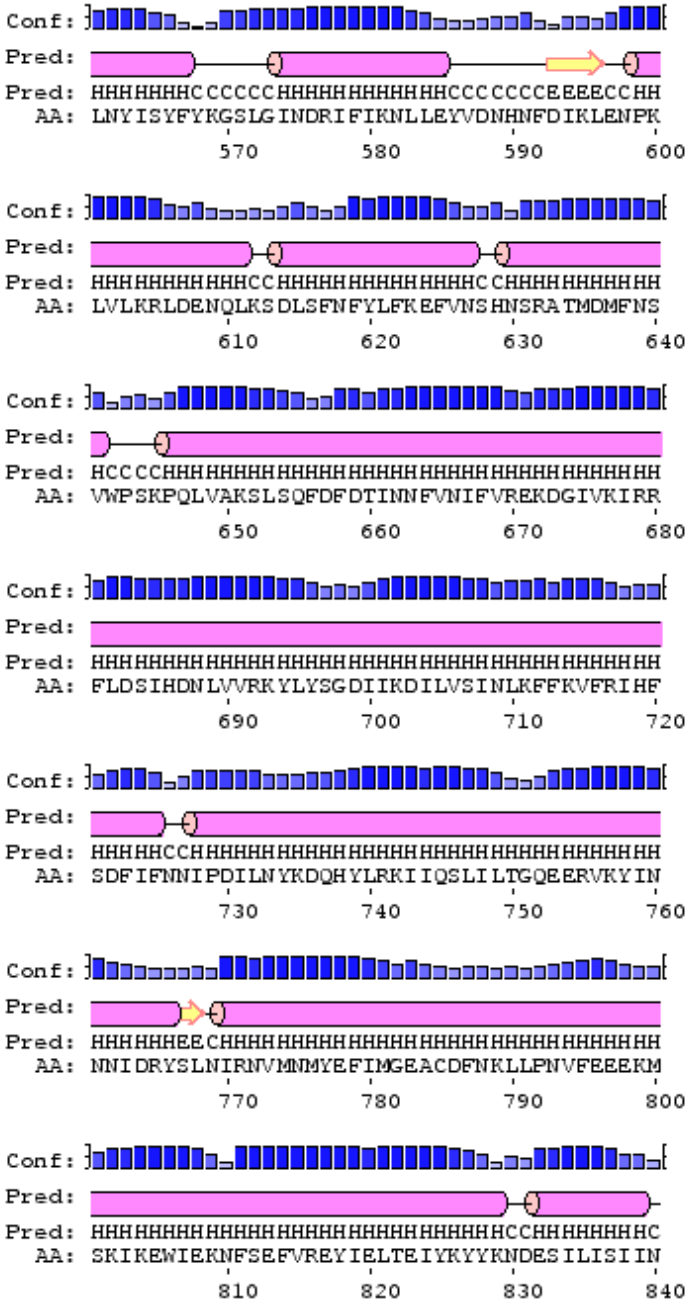
Aliphatic index: 105.13

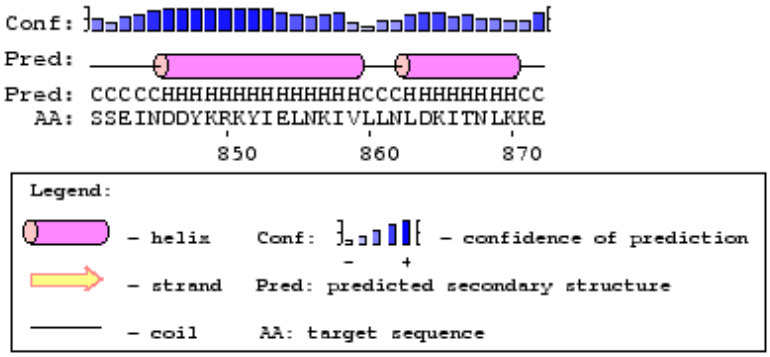
Grand average of hydropathicity (GRAVY): -0.239

3. MSC0676 secondary structure prediction.
(http://bioinf4.cs.ucl.ac.uk:3000/psipred/result/37065).




Conf:
Pred:
Pred: HEEEEEEEEECCCCCHHHHHHHHHHHCCCCCCCHHHHHHH
AA: DFIIPIVPVVNYSNSESKFSEFIELIGESDSLDKDIIWKI
290 300 310 320
Conf:
Pred:
Pred: HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
AA: SLYVNDIRLIKNIVNEFFVYNSAIRSENNQINKNKLFALI
330 340 350 360
Conf:
Pred:
Pred: HHHHCCHHHHHHHHHCHHHHHHHHHHHHHHHHHHHHHH
AA: TLKNLMPKEFDLLQMNKGYVFNVFEKIQKCQKNNSEIIQN
370 380 390 400
Conf:
Pred:
Pred: HHHHHHHHHHHHHHHHHHHHHHEEEHHHHHHHHHHHHH
AA: KLSNANNELSSLFKNYPEFEGNDKLSKTDNILFFKDIYEL
410 420 430 440
Conf:
Pred:
Pred: HHHHHHHHHHHHHHHHHCCCHHHHHHHHHHCCHHHHH
AA: INSDLIKTNFYVIKTIFYDDTTIVEKVQLRELRDNYPFVN
450 460 470 480
Conf:
Pred:
Pred: HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
AA: HCSNIKKIVVYNDPKIIECYENIESLEDQLNKKSIKFLLK
490 500 510 520
Conf:
Pred:
Pred: HCCHHHHHHHHHHHHHHHHHHHHHHHHHHCCHHHHH
AA: SLSRDEITDLFDSTSDKQMNKRYVSFLKFLFIEGLIDESY
530 540 550 560


Conf:
Pred:
Pred: HHHHHHHCCCCCCHHHHHHHHHHHHCCCCCCCEEEECCHH
AA: LNYISYFYKGSLGINDRIFIKNLLEYVDNHNFDIKLENPK
570 580 590 600
Conf:
Pred:
Pred: HHHHHHHHHHCCHHHHHHHHHHHHCCHHHHHHHHHHH
AA: LVLKRLDENQLKSDLSFNFYLFKEFVNSHNSRATMDMFNS
610 620 630 640
Conf:
Pred:
Pred: HCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
AA: VWPSKPQLVAKSLSQFDFDTINNFVNIFVREKDGIVKIRR
650 660 670 680
Conf:
Pred:
Pred: HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
AA: FLDSIHDNLVVRKYLYSGDIIKDILVSINLKFFKVFRIHF
690 700 710 720
Conf:
Pred:
Pred: HHHHHCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
AA: SDFIFNNIPDILNYKDQHYLRKIIQSLILTGQEERVKYIN
730 740 750 760
Conf:
Pred:
Pred: HHHHHHEECHHHHHHHHHHHHHHHHHHHHHHHHHHHH
AA: NNIDRYSLNIRNVMNMYEFIMGEACDFNKLLPNVFEEEKM
770 780 790 800
Conf:
Pred:
Pred: HHHHHHHHHHHHHHHHHHHHHHHHHHHCCHHHHHHHC
AA: SKIKEWIEKNFSEFVREYIELTEIYKYYKNDESILISIIN
810 820 830 840


Conf:
Pred:
Pred: CCCCCHHHHHHHHHHHHHCCCHHHHHHHHCC
AA: SSEINDDYKRKYIELNKIVLLNLDKITNLKKE
850
860
870
Legend:
- helix
- strand
- coil
Conf:
- confidence of prediction
Pred: predicted secondary structure
AA: target sequence