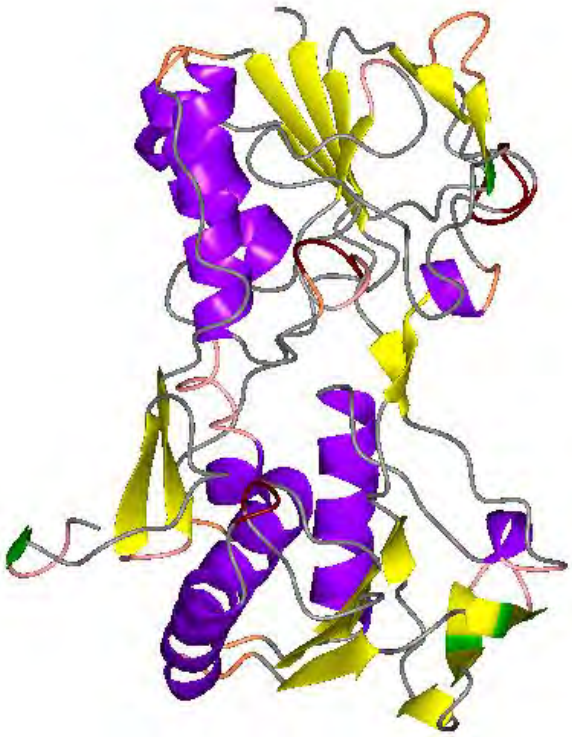
UPPSALA
UNIVERSITET

Structural and Functional Studies of NDH2 from *M. Smegmatis*



Lu Lu

Degree project in biology, Master of science (2 years), 2010
Examensarbete i biologi 30 hp till masterexamen, 2010
Biology Education Centre and Structural biology, dept. of Cell and Molecular Biology, Uppsala University
Supervisor: Torsten Unge

Contents

Abbreviations

bp	Base pair
DMSO	Dimethyl sulfoxide
dNTP	Deoxyribonucleoside
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetraacetic acid
EtBr	Ethidium bromide
IMAC	Immobilized metal ion affinity chromatography
LA	Luria-Bertaini agar
NDH-2	NADH dehydrogenase II
OD	Optical density
PEG	Polyethylene glycol
SDS-PAGE	Sodium dodecyl sulphate – polyacrylamide gel electrophoresis
FAD	Flavin adenine dinucleotide
Tm	Melting temperature
EM	Enzyme mixture
SM	Substrate mixture

Summary

M. smegmatis is considered as a non-pathogenic micro-organism. Proteins in the *Mycobacterium* genus, including pathogens such as those causing tuberculosis and leprosy, are conserved. Therefore *M. smegmatis* has been used in many lab experiments. NADH dehydrogenase II (NDH-2) plays an important role in the respiratory chain in Mycobacteria, it has been considered as a potential drug target, through which, the diseases caused by the pathogens could be cured. Structural and functional studies would help understanding the mechanism of the enzyme and developing inhibitors for medical uses. This paper describes experiments of the protein on NDH-2, from the design of a modified gene to expression and purification, with the final aim of crystallization, functional studies and drug design.

1. Introduction

1.1 Evolutionary relationship in Mycobacteria.

The species in Mycobacteria, including *M. tuberculosis* and *M. leprae*, are evolutionarily close (Gutierrez et al., 2005). *M. tuberculosis* and *M. leprae* are two major pathogens causing tuberculosis and leprosy, respectively. Around two billion people are infected with *M. tuberculosis*, and 1/1000 of the infected die each year, according to a report from WHO (2009). Considering the pathogenic properties of species in the Mycobacterium genus and the small sequence differences between among different species, the non- pathogenic *M. smegmatis* has been selected as a substitute for research purpose in the lab. Moreover, proteins from *M. smegmatis* tend to be more soluble than proteins from other species, and the solubility is fundamental for crystallization.

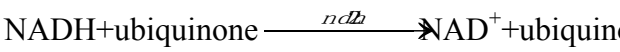
1.2 Selection of Potential Drug targets.

Proteins that are crucial for survival of the pathogen *M. tuberculosis*, have been selected as potential drug targets. The selection of potential targets was guided by the scientific article by a list of essential proteins (Sassetti et al., 2003) The potential drug targets were selected based on several criteria, for instance that the enzyme should play an important role in any phase of growth, infection, or proliferation, and that the enzymes from the pathogen and humans differ in structure so that the potential inhibitors would not affect the corresponding processes in humans.

1.2.1 NADH dehydrogenases

NADH dehydrogenases are generally classified into 3 groups, proton-translocating NADH–quinone oxidoreductase (NDH-1); NADH–quinone oxidoreductases that do not translocate the proton due to the lack of an energy-coupling site (NDH-2); and sodium-translocating NADH–quinone oxidoreductase (Na^+-NDH) (Yagi et al., 2001). NADH dehydrogenase II has been found in several organisms, such as, *Mycobacterium parascroful- aceum* (EMBL entry: ADNV01000294), *Rhodococcus equi* (EMBL entry: ADNW01000002), and *M. tuberculosis* (Cole et al., 1998). Homologs of bacterial enzymes have been found to be expressed in mitochondria of

eukaryotes (Yagi et al., 2001). NDH-2 usually cooperates with flavin adenine dinucleotide (FAD), a redox cofactor, and functions as a single subunit in the respiratory chain (Dong et al., 2009). NDH-2 is a kind of flavoproteins catalyzing the dehydrogenation of NADH, transferring electrons from NADH to ubiquinone (Eschemann et al., 2005), basically, it could function as the equation below:



NADH dehydrogenase II is crucial for optimal cellular growth with several non-fermentable carbon sources (Luttik et al., 1998), it is involved in reductive metabolic pathways, such as glycolysis. NADH dehydrogenase II from *M. smegmatis str. MC2 155* has 457 amino acids and it is 95% identical to NDH-2 from *M. tuberculosis*. Hence the studies of smegmatis NDH-2 would provide guidance for the one from *M. tuberculosis*.

2. Aim of thesis

The main aim of this thesis is to design and express soluble NADH dehydrogenase II constructs from *M. smegmatis*, to purify and crystallize the enzyme, and finally to solve the 3D structure by crystallography. Meanwhile potential inhibitors are tested in an activity assay for further drug design.

3. Materials and Methods

3.1 Bio-informatic analysis

Sequence information of gene and protein was collected from the sequence at NCBI, http://www.ncbi.nlm.nih.gov and also from http://www.expasy.org. Subsequent information based on the sequence, such as proportion of different types of amino acids, and extinction coefficient were obtained from ProtParam tool http://www.expasy. org/tools/protparam.html. pBLAST (protein Basic Local Alignment Search Tool, http://www.ncbi.nlm. nih.gov/BLAST) was employed to search for sequence similarity and to analyze conserved domains. Sequence alignment was also performed in Clustalw (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Known structures of NDH-2 were searched for in PDB database (http://www.pdb.org). The structure of the enzyme from the M. smegmatis was predicted by Protein Homology/analogY Recognition Engine (PHYRE, http://www.sbg.bio.ic.ac.uk/~phyre/) to give an overview of the structure and to allow the design of modifications from the theoretical model. Trans-membrane regions were analyzed by TMHMM2.0 (Center for Biological Sequence Analysis, Technical University of Denmark http://www.cbs.dtu.dk/services/TMHMM-2.0/).

3.2 Isolation of gene and PCR.

A plasmid with the full length gene of NADH dehydrogenase II was already available in the lab. Three pairs of primers were designed manually based on the analysis above and ordered from Invitrogen™, to construct three different expression vectors. Primers were stored -20°C in the stock solution (10mM Tris-HCl buffer, pH 8.0), with

the concentration of 0.1mM.

The isolation of the target gene was analyzed by agarose gel electrophoresis (1% agarose, EtBr) in 1X Tris-EDTA buffer pH 8 (50mM Tris-acetate and 1mM EDTA).

PCR were programmed as shown in Table 1.

Table 1. PCR-programs for amplification and $(His)_6$ addition

Program		Cycles
Pre-denaturation	94°C 2min	
denaturation	94°C 1min	
Annealing[1]	61°C 1min	40X
elongation	72°C 2min	
Final extension	72°C 2min	
hold	4°C	

The preparation of PCR mixture was done as shown in Table 2.

Table 2. PCR mixture

Chemical	Volume (μl)
ddH_2O	36
10X pfu buffer	5
DMSO	2.5
dNTP	1
Forward primer	2
Reverse primer	2
template	1
pfu enzyme	0.5
	50

3.3 Ligation and transformation.

Amplified PCR products were extracted and purified by cutting out the band from the gel and centrifuging at 11300g for 5 min, mixed with 17.8 μl water, 2 μl 10X Taq buffer and 0.2 μl Tag enzyme (Invitrogen, LA, USA) and 0.5 μl dNTP to add adenine at 3', and incubated at 72°C for 10 min. 1 μl of the mixture above was incubated at room temperature for 20min, with 0.5 μl salt solution (pEXP5-TOPO kit, Invitrogen, LA, USA), 1 μl water and 0.5 μl pEXP5-CT vector. Following the ligation, the ligated plasmids were transformed to TOP10 cells (Invitrogen, LA, USA). Competent cells were thawed on ice for 3 min, mixed with plasmids, incubated on ice for 5 min, heat shocked at 42°C for 30 seconds, incubated immediately on ice for 2 min, and finally spread onto the LB plate prepared with 50 μg/ml ampicillin (LA plate). The plate was incubated at 37°C overnight.

3.4 Plasmid extraction and transformation to expression host.

Clones were picked up from the plate with ampicillin (50 μg/ml), incubated overnight.

Plasmids were extracted using the miniprep kit, following the protocol (Qiagen™). PCR was programmed as shown in Tables 3, 4, and 5.

Table 3. Analytical PCR-program

Program		Cycles
Pre-denaturation	94°C 2min	
denaturation	94°C 1min	
annealing	55°C 1min	25X
elongation	72°C 2min	
Final extension	72°C 2min	
hold	4°C	

Table 4. Analytical PCR mixture

Chemical	Volume (µl)
ddH2O	18.3
10X Taq buffer	2.5
DMSO (3%-5%)	1
dNTP	0.5
T7 promoter Forward primer	1
T7 terminus Reverse primer	1
Template	0.5
Taq enzyme	0.2
	50

Table 5. Analytical PCR mixture

Chemical	volume(µl)
ddH2O	18.3
10X Taq buffer	2.5
DMSO (3%-5%)	1
dNTP	0.5
T7 promoter forward primer	1
Ndhsmeg6	1
Template	0.5
Taq enzyme	0.2
	50

T7 promoter forward primer and T7 terminus reverse primer are a pair of universal primers annealing to corresponding complementary segment from the vector, which were applied in analytical PCR to test the presentence of the gene in the plasmid. Another pair of primers, T7 promoter forward primer and ndhsmeg6 was chosen for verification of both the presence and the orientation of the gene. The products of PCR were analyzed by agarose gel electrophoresis in the same way as in section 3.2. The

elongation time was extended to 2 minutes so that the PCR reaction would have sufficient time to finish every circle to the very end. Positive clones were picked up and incubated in 3 ml LB medium with ampicillin at 37℃ . For expression 1 μl of extracted plasmids were transformed into *E. coli* BL21-AI cells as described above. Clones were analyzed with test expression.

3.5 Cell lines (E. coli)

Only two cell strains were used in this thesis, Top10 and BL21-AI. Top10 cell line was chosen to amplify the cloning vector; it lacks the T7 promoter so that recombinant protein would theoretically not be expressed. BL21-AI cell line is the expression host in most laboratories for the study of recombinant proteins. It has high tolerance to toxic proteins and high regulation of the expression by arabinose-araBAD promoter interaction.

3.6 Test Expression

Clones on the plate were cultured in 4 ml LA medium, to reach an OD_{600} between 0.6 and 1. The culture was cooled down, then induced with arabinose (2mg/ml) at three temperature series, 16°C, room temperature and 37°C. Cell culture, 100 μl, was saved as backup for further experiments. The expression was tested by SDS-PAGE with PhastGel™. A 100 μl cell culture was collected and spun down, 20 μl X sample buffer (pH 8.0, 1% SDS, 2% beta-mercaptoethanol, 4% glycerol, 5 mM EDTA, 20 mM Tris-HCl) was added and thoroughly mixed with cell pellets. The mixture was incubated at 95°C for 10 min to lyse the cells and denature the proteome, and vortexed 20 sec to break down the DNA, finally centrifuged for 1 min to get homogeneous proteome samples.

3.7 Large-scale expression, cell lysis and solubility test

Backup cells collected at the previous step was cultured as an inoculum in 100 ml LA medium and grown at 37°C on shaker. Inoculum was incubated in 1 l LA medium for large scale till the absorbance at 600nm of the cell culture reached 0.6. The cell culture was cooled down at room temperature and 10 μM FAD added. Expression was induced at room temperature by 2 mg/L arabinose. Cells were collected by centrifugation at 16100g for 20 min, washed in 1X SSP buffer, centrifuged again at 7000g for 30 min. The ingredients of LB medium were 10 g of tryptone, 10 g of yeast extract, and 5 g of sodium chloride in 1 l distilled water.

Cells were lysed in 12 ml 50 mM phosphate buffer with 300 mM sodium chloride, 10 mM imidazole, 10% of glycerol, and 0.5% of Triton X-100, with the pH value adjusted to 8.0. Five mg of Lysozyme, 2.5 mg RNAse A, 5 mg DNase I, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 7 mM beta -mercaptoethanol were added in the lysis buffer. These substances have different functions in the lysis process: lysozyme helps lysing the cells, breaking down the membrane, RNase and DNase digest the RNA and DNA respectively, PMSF prevents the digestion of proteins from proteases, and mercaptoethanol is used as the reducing agent to break down the disulfide bonds. Cells were disrupted by cell disruptor, followed by centrifugation at 14000g for 10 min to separate into supernatant and cell pellets. The result was

analyzed by SDS-PAGE.

3.8 IMAC and size exclusion chromatography (SEC)

The supernatant after lysis was loaded to a 0.5 ml Nickel column, pre-equilibrated with 10 ml lysis buffer, and mixed thoroughly for 20 min in cold room. The column was left to sediment, and washed by 20 ml 50 mM phosphate buffer with 300 mM NaCl, 20 mM imidazole, then eluted with 3 ml the same washing buffer containing 250 mM imidazole. 10 μM FAD was added in every buffer. The different fractions were analyzed by SDS-PAGE.

Buffer exchange was performed in a DG-10 column, for the substitution of elution buffer with 20 mM Bis-Tris Propane with 200 mM NaCl, 10 mM FAD, 10 mM EDTA and PMSF (buffer BTSSE).

AKTA primer (Pharmacia Biotech) protein purification system was used to further purify the protein by size. The Superdex^{TM}75 column was selected, due to the suitable separation capability, and pre-equilibrated with the buffer BTSSE. Different fractions were collected automatically and analyzed by SDS-PAGE. Additionally the OD280 was measured as the indicator of the protein concentration. Based on the result of SDS-PAGE, different elution fractions were selectively pooled and concentrated for further studies.

3.9 Stability test

Purified NADH dehydrogenase II was incubated overnight under four different conditions, at room temperature, at 42°C, on ice at 4°C, frozen and thawed, respectively. After treatments, SDS-PAGE was performed to analyze the stability of the protein.

3.10 Activity assay

The dehydrogenase activity of NDH-2 was measured and calculated based on the chemical reaction:

$$\text{NADH+menaonine} \xrightarrow{ndh} \text{NAD}^{+}\text{+menaonine-H}$$

The measurement was based on the absorbance peak of NADH at 340nm, which is proportional to the concentration of the NADH in the solution. As the NADH is consumed during the reaction, the absorbance decreases, so that the catalytical dynamics parameters could be calculated, such as K_m, K_{cat}, and V_{max}. Table 10 shows the reaction conditions, EM represents enzyme mixture while SM represents substrate mixture.

The reaction mixture was prepared following the protocol as above, and incubated for 10 min at room temperature to stabilize the system before enzyme was added into the system. The absorbance at 340 nm was measured in 96-well Microplate (Greiner bio-one^{TM}) by 2104 multilabel reader (PerkinElmer^{TM}). The reaction was monitored

for 1 h to get the change of instantaneous absorbance values.

Potential inhibitors from AstraZeneca Bangalore™ were added in the assay and incubated with mixture for 10 min to stabilize before the enzyme was added.

Table 6. Activity Assay of NDH-2 (each reaction)

EM			SM	
Chemical (final C)	V (µl)		V (µl)	Total (µl)
HEPES 50mM	2		1.75	
NaCl 100mM	0.8		0.7	
Brij(%)	0.32		0.28	
NADH 0.3 mM	0.45	Menadione 50µM	0.0375	
MQ water	36.4		31.79	
total	40		35	75

3.11 Crystallization

Purified NADH dehydrogenase II was concentrated to at least 10mg/ml in vivaspin with the cut-off of 10000 Da (Vivascience™). 10 mM DTT and 1 mM substrate NADH were added into enzyme solution before the crystallization screens were setup with the assistance of screening robot with 0.6 µl protein and 0.6 µl reservoir buffer in the sitting drop and 75 µl in every well. Buffer screen was setup with 24 different buffer conditions. Two commercial screens, Morpheus and JSCG+, were also setup on the Crystal QuickPlates (Greiner bio-one™).

4. Results

4.1 Bioinformatical analysis

The gene MSMEG_3621 contains 1374 nucleotides representing 457 amino acids (see Appendix 1) and a stop codon. It has only one cysteine in the sequence, which indicates that it is more likely to fold correctly, considering that no disulfide bond would be formed in the monomer. The extinction coefficient is low due to the small proportion of Trp, Phe, and Tyr, the first of which is mostly contributing to absorbance at 280 nm. However, coloured FAD that was added in every step has an observable absorbance at 280 nm, suggesting that the extinction coefficient could not be used for measurement of protein concentrations.

NADH dehydrogenase II was predicted by TMHMM to have a trans-membrane helix approximately from 382 to 404 amino acids (Fig. 1). Engineered protein expressed in prokaryote cells, specifically E. coli, exposes in the aqueous environment, high entropy would cause problems during the folding process. Hence constructs were designed with a truncation of trans-membrane domain.

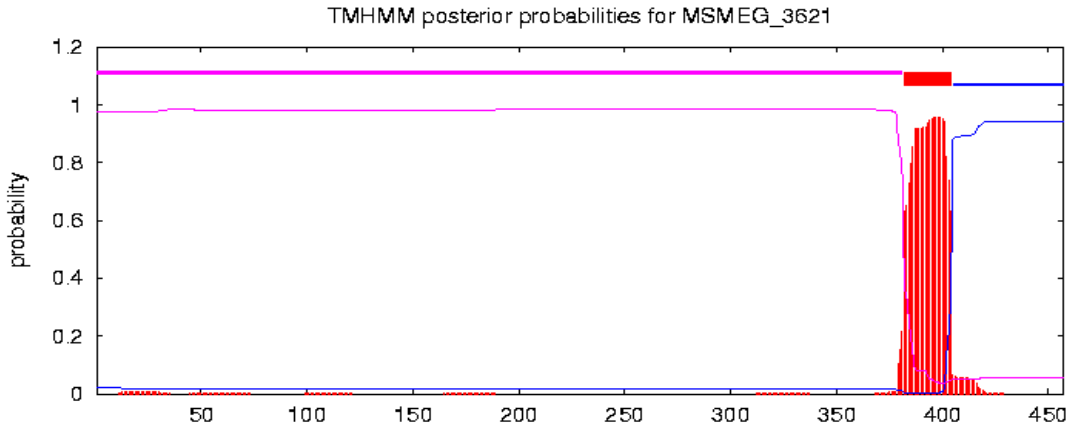


Fig. 1 TMHMM prediction

To get a preliminary image and to some extent useful information of the structure of NADH dehydrogenase II, Phyre prediction was employed on the Phyre server. Full length of NDH-2 amino acids sequence was input to the server for modeling. The predicted structure is shown in Fig. 2.

The best predicted structure with an e-value of 8.6e-39 and estimated precision value of 92%, starts from amino acid 13 to 380, while it was based on the sequence alignment to NADH oxidase from *Pyrococcus furiosus* (pdb entry: 1xhc), chain A. The prediction has suggested that two proteins could both have a NADH binding domain (NADB), which also has been proven by the conserved domain analysis (data not shown). The E-values stand for the average number of times in multiple testing that the statistic test is as extreme as the value is observed, to indicate the significance of the deduced test results, the precision of predictions is inversely correlated to e-value. The structure has two main domains linked by one long loop and two beta strands, one of which shows a beta-beta-alpha sandwich motif, while the other one has a beta-beta-alpha-beta sandwich structure.

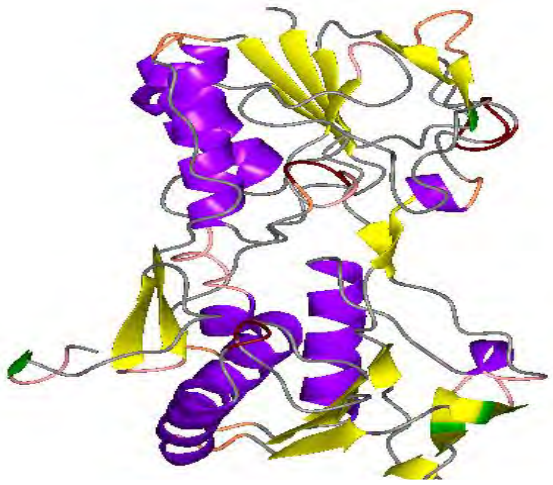


Fig.2. Phyre prediction of full length of NDH-2

The amino acids in the interface between protein and membrane are positively charged to interact with phosphate on the membrane (Fig. 3). The location of the membrane was evident through the position of the C-terminus present in the prediction (residue 380). The residues 224Lys, 228Lys, and 378Lys are stretching out

to the membrane.

Three constructs were finally designed to produce soluble NADH dehydrogenase II from *M. smegmatis*, construct 1 with a C-terminal truncation and a C-terminal his-tag, construct 2 with a C-terminal truncation and an N-terminal his-tag, and construct 3 with the same truncation and split his-tags at both termini.

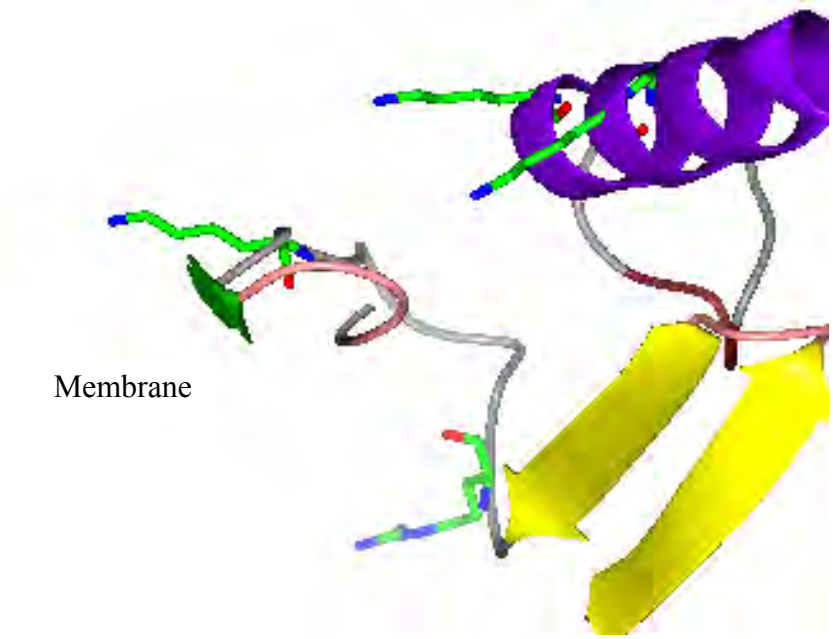


Fig. 3. C-terminal positive-charged amino acids.

4.2 Isolation of gene and mutation PCR

Genes were amplified from the plasmid with full length of NDH-2, as shown in Fig.4. Three constructs were successfully cloned and truncated at C-terminus, and his-tags were added onto the fragments. Verification of the cloned gene fragments was done by the agarose gel electrophoresis.

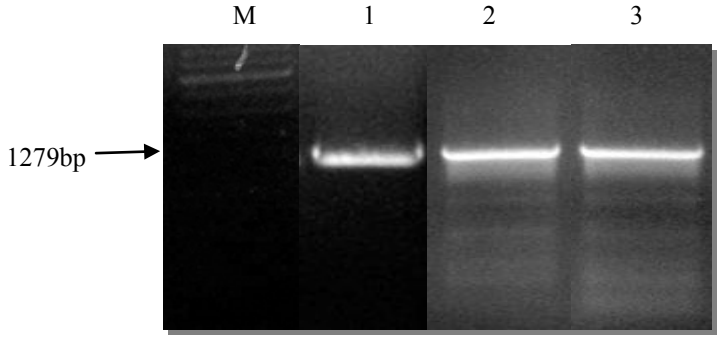


Fig.4. PCR truncation products. Lane M, 200bp marker from O'rangeRuler™, Lane2, 1176bp isolated gene with Cterninal his-tag (Construct 1); Lane3, 1179bp isolated gene with Nterminal his-tag (Construct 2); Lane4, 1179bp isolated gene with NC terminal his-tag (Construct 3).

Construct 1 was theoretically 3 bp shorter than the other two constructs because a smart codon GCT corresponding to alanine was added in the last two constructs, which could prolong the half-life of the proteins in the expression hosts and therefore, stabilize the expression of the recombinant proteins in *E. coli*. However, the difference of 3 bp could not be observed on the gel.

4.3 Ligation and transformation

The isolated PCR products were extracted from the gel, ligated to the TOPO-CT vector, and transformed to TOP10 cells. The ampicillin-resistant gene included in the vector would help selecting for transformed cells. Due to the properties of TOPO ligation, the gene could be bi-directionally ligated to the vector, so the orientation of the gene had to be verified by analytical PCR, with one of the primers on the vector. Plasmid DNA was extracted from the cells of positive clones, and the concentrations of the plasmid from two constructs were measured. Yields of plasmid were 90 ng/µl, 203 ng/µl respectively. The gene of construct 3 has not been proven to ligate to vector yet. The result of analytical PCR is shown below (Fig.5).

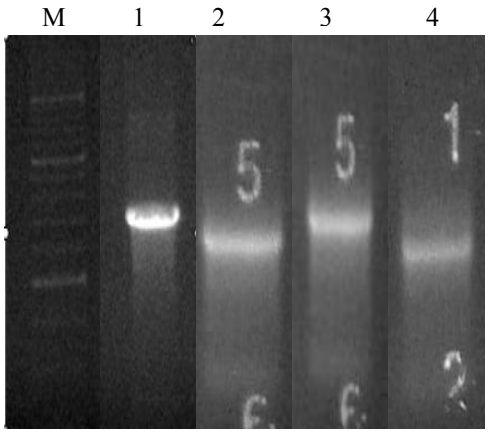


Fig. 5. Products of analytical PCR. Lane M, 200bp marker from O'rangeRuler™, lane 1, construct 1 amplified with T7 promoter forward primer and terminal reverse primer; lane 2, construct 1 with ndh_smeg_for1 and terminal reverse primer; Lane 3 construct 2 with the same pair of primers as lane1; lane 4, construct 2 with ndh_smeg_for2 and T7 terminal reverse primer.

The amplified gene using T7 promoter forward primer and terminal reverse primer contains two segments from the vector, 144 bp downstream and 78 bp upstream. The gene amplified with a primer from vector and the other one from gene is approximately 100 bp shorter, which could be observed on the gel. Consideration should be made when selecting the primers for analytical PCR, primers containing his-tag codons could unspecifically bind to the complementary part in the vector. Hence, to reduce the false positive chance of analytical PCR, different primers combinations were applied.

The single bright bands indicated that ligation and transformation were successful for the first two constructs but not construct 3. So far, none of the clones isolated from the LA-plates for this construct were positive according to the analytical PCR. Plasmids of positive clones were sent for sequencing. The analysis showed that construct 2 contained a point mutation from C to T, from arginine to cysteine. Experiments were carried on despite of the possible effect of the mutation. Plasmids were transformed to

BL21-AI cells.

4.4 Test Expression

After induction, intensive expression was observed at three different temperature series (Figs. 6 and 7), SDS-PAGE showed a protein band at the correct position around 45kD according to the low molecular weight marker. The size is slightly larger than its theoretical value, but the migration of protein can be affected by several

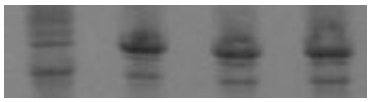


Fig6 . Expression of construct 1. Lane 1, not induced, lane 2, at 16°C, lane 3 22°C , and lane 4, 37°C.

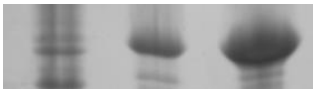


Fig7. Expression of construct 2. Lane1, not induced, lane 2, at 22°C, Lane3, 37°C.

Induction at lower temperature has been proven to be an effective way to solubilize protein and prevent formation of inclusion bodies. One explanation is that proteins express at lower level at lower temperature so that they do not aggregate. Another explanation is that it may offer much more time for proteins to fold correctly.

Another consideration was the degradation occurring during fermentation, because of too long time of induction. In conclusion, both constructs were induced to express at 22°C for 4 hours. The back-up cells were inoculated to prepare for large-scale fermentation.

4.5 Solubility test and purification

4.5.1 Construct 1 2 grams cells were collected from 1 l of culture, and lysed with cell disruptor thoroughly. The purification IMAC and SEC were carried on immediately after lysis. To avoid the risk of proteolytic degradation, the inhibitors PMSF and EDTA were added to the lysis buffer. Every collected fraction was analyzed by SDS-PAGE. (Fig. 8)

M I P S F1 W E1 F2 E2 19 23 25

Fig.8. M represents low molecular weight marker, I induced cell pellets, P precipitates after lysis, S supernatant after lysis, F1 flowthrough of the IMAC, W wash fraction of IMAC, E1 elution fraction of IMAC, F2 flowthrough of DG10 buffer exchange column, E2 elution fraction of DG10, 19, 23, 25, relative fractions of SEC.

Fig.8 shows that a lot of protein formed inclusion bodies and stayed in cell pellets, and that another overexpressed protein closely beneath was also present in the cell pellets before lysis, indicating that degradation happened during fermentation. The bright side of the experiment was that it was in the cell pellet fraction if it degraded.

NDH-2 was purified by IMAC. Over-loading with protein could increase the purity of target protein since the his-tagged protein has the highest affinity to the matrix.

NDH-2 was in the flow-through fraction, proving that the column was overloaded.

The small band beneath the target band represents cleaved NADH dehydrogenase II. Most the material was interestingly in the wash fraction. One possible explanation could be that the histidine-tag has been cleaved off either entirely or partly.

Fig. 8 shows the pattern of the SEC, the second peak represents NDH-2, according to the SDS-PAGE, but given the profile of SEC itself, the construct 1 was purified. Some degradation occurred during SEC.

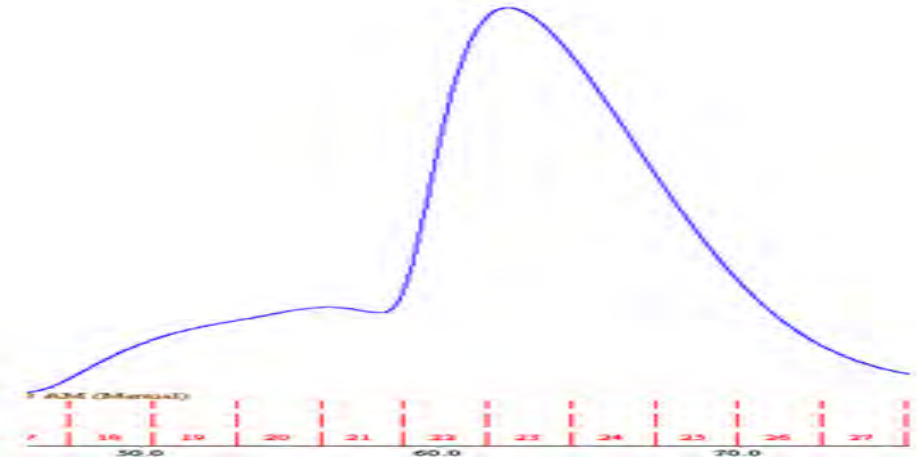


Fig.9 Graph of SEC the numbers in red colour represents different fractions.

The concentration of proteins was measured in every step, so that the loss and recovery could be followed (Table 7).

For 1 l culture, around 8.1 mg of pure protein was obtained, some materials was lost in every step of purification.

Table 7. Concentration measurement (Bradford) (1 l culture)

process	C(mg/ml)	V(ml)	m(mg)	OD_{595}	D-factor[1]
After IMAC	4.3	3	12.9	0.233	20
After SEC	10.7	0.75	8.1	0.968	10

D-factor means dilution factor

4.5.2 construct 2 This construct showed different properties during purification (Fig. 10).

Most of the overexpressed construct 2 appeared in the inclusion-body fractions. The soluble fraction isolated by IMAC was still inhomogeneous.

Furthermore, SEC was employed to purify the proteins by size. The pattern of SEC is shown in Fig. 11. Therefore, due to the inhomogeneity of construct 2 materials after purification, it was not our main target to continue working on. Because of the point mutation, a mutation PCR would have to be applied to mutate back from the cysteine to arginine.

Nevertheless, different fractions from SEC were loaded in the activity assay to test

whether it could transfer the hydrogen from NADH to menadione.

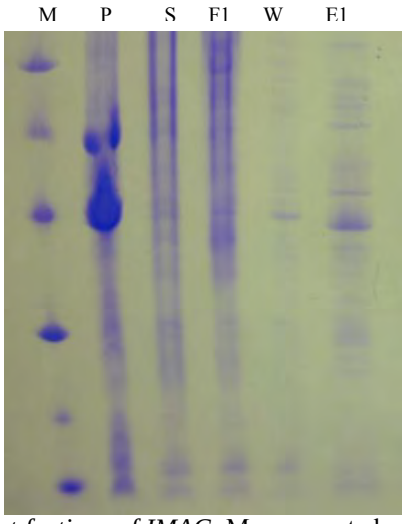


Fig.10. SDS-PAGE of different factions of IMAC. M represents low molecular weight marker, P precipitates after lysis, S supernatant after lysis, F1 flowthrough of the IMAC, W wash fraction of IMAC, E1 elution fraction of IMAC.

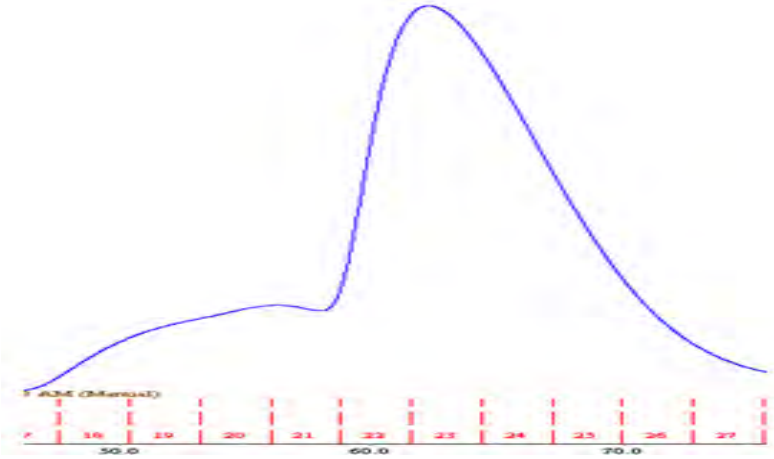


Fig.11 Graph of SEC. The numbers in red colour represents different fractions.

4.6 Stability test

4.6.1 Construct 1. As the degradation was observed in purification process, stability test was run to figure out the optimal condition for the storage of NADH dehydrogenase II (Fig.12).

There was a faint band under the target band in the elution fraction after IMAC, and afterwards construct 1 degraded to different extents under every condition. Construct 1 stored at 42°C seems less degraded. Some bands were present above the target band, which suggests that irregular aggregation occurred. According to the gel, The best way to store construct 1 is to freeze and thaw it.

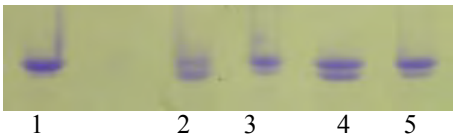


Fig. 12. Stability test of construct 1 by SDS-PAGE analysis. Lane 1, elution fraction directly after IMAC; lane 2, at room temperature; lane 3, at 42°C; lane4, on ice in cold room; lane5, frozen and thawed. Sample lanes 2-5 were incubated overnight.

4.7 Activity assay

Both constructs were tested with activity assay. Enzymatic activity could be considered as an indication of a correct conformation.

4.7.1 Construct 1. Construct 1 showed no or weak activity although it had a yellowish color, suggesting that it binds to the co-factor.

4.7.2 Construct 2. Construct 2 has high activity compared to the previous construct; the NADH in the assay has been consumed after the enzyme was added. (Fig. 13)

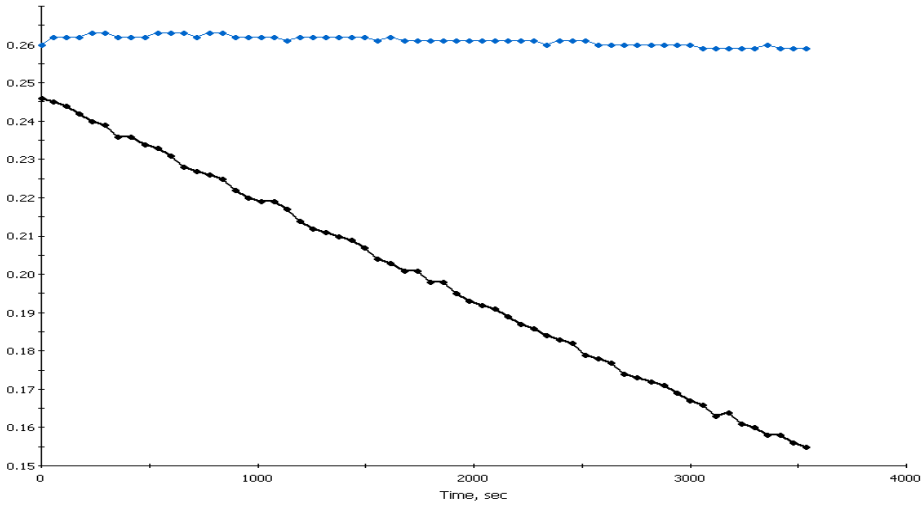


Fig.13. Activity assay of construct 2. Blue curve represents the blank without enzyme. The black curve shows that the absorbance dropped sharply during 1 h reaction.

The activity measurement indicates that the enzyme shows hydrogen transferase activity from NADH to menadione, thus there is a consumption of NADH. Absorbance at 340nm is theoretically proportional to the concentration of NADH. The assay was done with NADH dehydrogenase II material from the SEC elution fraction 20. According to the SDS-PAGE analysis this fraction contained very little material.

Different potential inhibitors have been tested in the assay, which could also further prove that the NDH-2 material is correctly folded. The inhibitors tested were confirmed to inhibit the activity of DXR, which catalyzes a reaction transferring a hydrogen from MEP to NADPH. The inhibitors could mimic the NADH or NADPH and block the NADPH binding site by competitive inhibition. However, none of them could inhibit the activity of construct 2.

4.8 Screening of crystallization

Buffer screen provided useful information of which buffer is most suitable for NDH-2, since it is well designed to contain pH value series and different buffer contents. Construct 2 is sensitive to pH value, based on the observation that it precipitated a lot at lower pH value from 3.5 – 5, but it remains in clear drops at pH value 6 – 7. So far no crystals were observed.

5 Discussion

The Phyre prediction is based on sequence alignment to a known structure, and the given e-value and the estimated precision represent the predicted accuracy of each prediction. Since the environmental elements are not considered in any prediction methods, the conformation of last 60 amino acids which have interactions with membrane *in vivo*, are too complex to predict, so they were not included. The reason of absence of first 12 amino acids in the structure, could be that the terminus is disordered; constructs where the termini are truncated are common in crystallography.

The trans-membrane helix has some particular features regarding to the distribution of amino acids, due to the docking position in the di-phospholipid membrane. It usually has non-polar residues coiling through the membrane, which could be problematic during folding and crystallizing, if they were exposed in the polar solvent. One option to solve the instability are truncation of the trans-membrane.

A his-tag was added to provide convenience in IMAC purification, based on the high affinity of histidines to Ni-matrix. The localization of $(His)_6$-tag in engineered proteins was not as critical as the modification of constructs, for instance, mutation, truncation, or removal of certain amino acids, which could potentially increase the entropy of protein. Addition of a his-tag normally does not affect the folding, activity or stability. However, it has been reported that it could prolong the half-life of engineered protein in an expression host (Alexander, 1997).

Due to the fact that NADH dehydrogenase II is a membrane protein, partially interacting with the membrane, attention should be paid on decreasing the negative effect of the positively charged residues in the interface between protein and membrane during folding and crystal packing. Addition of high salt is one of the solutions to solve the problem, since the anions would neutralize the surface charges. Point mutation to replace the positively charged amino acids with alanine or glycine could also be introduced.

Precipitation of the construct 1 was observed during purification, especially in the buffer screen. The protein was partially degraded, according to the result of SDS-PAGE. Optimization of different buffer conditions has been tried; bis-tris-propane was found to be a suitable buffer for NDH-2 to solve the problem of precipitation. However, the degradation still occurred under every condition, and the presence of a faint band beneath the target band on the gel indicated that NDH-2 was not stable. It could be explained by the fact that the flexible ends of the engineered protein cause stability problems during folding in the expression hosts. Therefore, different constructs need to be tested to find a stabilized NDH-2.

No crystals have been observed yet, but potential candidates were spotted. These hits showed some regular aggregation pattern, suggesting an initiation of crystal formation, following a disturbance to the continuous crystal packing. The degradation of the C-terminal his-tag could be responsible for the precipitation in the buffer screen and also the disturbance during packing, since the inhomogeneous materials would not aggregate regularly and continuously to form crystals.

This research will be carried on with different constructs if we fail to figure out in which buffer current constructs remain soluble and intact.

To conclude, NDH-2 from *M. Smegmatis* was successfully cloned and expressed. The material was soluble. The construct 1 material is, however, unstable under the conditions tried so far. A dramatic improvement of the stability was obtained by the addition of 10 μM FAD to the buffer. But still the material gets proteolyzed at the C-terminus with concomitant loss of FAD. Apparently additional stabilizing components have to be added. These components could be negatively charged molecules mimicking the membrane phospholipids. Additional modifications of the C-terminal sequence might also be necessary. It is not unexpected that the majority of the construct 2 material was insoluble. The mutation of an Arg to a Cys removes a positive charge and furthermore introduces a SH group with the potential to form S-S bonds. Since no activity was observed with the construct 1 material the conclusion must be that the binding site for the NADH is not correctly arranged. While the construct 2 catalyzes the consumption of NADH, which might indicate that the his-tag at the beginning to the peptide confirms the correct folding of NDH-2, or at least the NADH binding site. Since NADH is a substrate, one must expect it to bind loosely. A stable protein could, however, be obtained with one of the NADH-mimicking inhibitor compounds.

6 Acknowledgement

First of all, I would like to thank my supervisor Torsten Unge, this work could not be done without his guidance and patience, I myself was deeply impressed by his optimism and inspiration, I would never have the same intensive passion as he does, on the structures. And special thanks for the time he spent on reviewing my thesis.

Also thanks to Alwyn Jones, for his tremendous advices and sometimes understandable humor. Thanks to Annette Roos, Terese Bergfors, Adrian Suarez, Kristina Backbro, for the assistance and guidance in the lab.

Thanks to my girl for our endless discussion and being there for me all the time.

References

Alexander, V. (1997). The N-end rule pathway of protein degradation. Genes to Cells 2 : 13-28.

Dong, C.K., Patel, V., Yang, J.C., Dvorin, J.D., Duraisingh, M.T., Clardy, J., Wirth, D.F. (2009). Type II NADH dehydrogenase of the respiratory chain of Plasmodium falciparum and its inhibitors. Bioorg. Medicin. Chem. Lett. 19, pp. 972–975.

Cole S. T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S. V., Eiglmeier K., Gas S., Barry C. E. 3rd, Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998). Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 393**,** 537-44.

Eschemann, A., Galkin, A., Oettmeier, W., Brandt, U., Kerscher, S. (2005). HDQ (1-hydroxy-2- dodecyl-4(1H) quinolone), a high affinity inhibitor for mitochondrial alternative NADH dehydrogenase: evidence for a ping-pong mechanism . J. Biol. Chem 280:3138-42.

Gutierrez, M. C., Brisse, S., Brosch, R., Fabre, M., Omais, B., Marmiesse, M., Supply, P., and Vincent, V. (2005). Ancient origin and gene mosaicism of the progenitor of Mycobacterium tuberculosis. PloS. Pathog. 1**,** e5.

Global tuberculosis control: epidemiology, strategy, financing (WHO report 2009)

Luttik, M. A., Overkamp, K. M., Kotter, P., de Vries, S., van Dijken, J. P., and Pronk, J. T. (1998). The Saccharomyces cerevisiae NDE1 and NDE2 genes encode separate mitochondrial NADH dehydrogenases catalyzing the oxidation of cytosolic NADH. J. Biol. Chem. 273**,** 24529-34.

Nallamsetty, S., and Waugh, D. S. (2006). Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. Protein. Expr. Purif. 45**,** 175-82.

Sassetti, C. M., Boyd, D. H. and Rubin, E. J. (2003), Genes required for mycobacterial growth defined by high density mutagenesis. Molecular Microbiology, 48: 77–84.

Yagi, T., Seo, B. B., Di Bernardo, S., Nakamaru-Ogiso, E., Kao, M. C., and Matsuno-Yagi, A. (2001). NADH dehydrogenases: from basic science to biomedicine. J. Bioenerg. Biomembr. 33**,** 233-42.

Appendix 1 Sequence information

>MSMEG_3621

ATGAGCCATCCCGGAGCTACGGCGTCGGATCGGCATAAAGTCGTCATCATCGGTTCGGGT

TTCGGTGGTCTCACCGCTGCCAAGACCCTCAAGCGCGCTGACGTCGACGTCAAGCTAATC

GCCCGTACCACGCACCACCTCTTCCAGCCGCTGCTCTACCAGGTGGCGACCGGCATCATC

TCCGAGGGCGAGATCGCCCCGGCCACTCGAGTGATCCTCCGCAAGCAGAAGAACGCCCAG

GTCCTTCTCGGCGATGTGACGCACATCGATCTGGAGAACAAGACCGTGGATTCGGTGCTG

CTCGGTCACACCTACTCGACGCCCTACGACAGCCTCATCATCGCCGCGGGCGCGGGTCAG

TCCTACTTCGGCAACGACCACTTCGCCGAGTTCGCACCCGGCATGAAGTCGATCGACGAT

GCGCTGGAGCTGCGCGGTCGCATCCTCGGCGCGTTCGAACAGGCCGAGCGCTCCAGCGAC

CCGGTGCGCCGCGCGAAGTTGCTGACGTTCACCGTCGTCGGCGCGGGCCCGACCGGCGTC

GAGATGGCCGGACAGATCGCCGAATTGGCCGACCAGACTTTGCGGGGCAGCTTCCGCCAC

ATCGATCCCACCGAGGCCCGGGTGATCCTGCTCGACGCCGCACCGGCCGTGCTACCGCCC

ATGGGCGAGAAGCTCGGCAAGAAGGCGCGGGCCCGTCTGGAGAAGATGGGCGTCGAGGTC

CAGCTGGGTGCGATGGTCACCGACGTCGACCGCAACGGCATCACCGTCAAGGATTCCGAC

GGGACCATCCGTCGCATCGAGTCGGCGTGCAAGGTGTGGTCGGCCGGTGTGTCGGCCAGC

CCTCTCGGCAAGGATCTCGCCGAGCAGTCGGGTGTCGAACTCGACCGCGCGGGCCGGGTC

AAGGTACAGCCCGACCTGACGCTGCCCGGTCACCCGAACGTGTTCGTCGTGGGCGACATG

GCGGCCGTCGAGGGCGTGCCCGGTGTGGCGCAGGGCGCCATCCAGGGTGGCCGCTACGCC

GCGAAGATCATCAAGCGTGAGGTCAGTGGCACCAGCCCGAAGATCCGCACGCCGTTCGAG

TACTTCGACAAGGGCTCGATGGCGACGGTGTCGCGGTTCTCCGCGGTGGCCAAGGTGGGT

CCCGTCGAGTTCGCGGGCTTCTTCGCCTGGTTGTGCTGGCTCGTGCTGCACCTGGTGTAC

CTGGTCGGGTTCAAGACGAAGATCGTCACACTGCTGTCGTGGGGCGTGACGTTCCTGAGC

ACCAAGCGTGGTCAGCTCACCATCACCGAGCAGCAGGCCTATGCGCGAACCCGCATCGAG

GAGCTCGAGGAGATCGCGGCGGCGGTGCAGGACACCGAGAAAGCCGCGTCCTAG

>MSMEG_3621

MSHPGATASDRHKVVIIGSGFGGLTAAKTLKRADVDVKLIARTTHHLFQPLLYQVATGII

SEGEIAPATRVILRKQKNAQVLLGDVTHIDLENKTVDSVLLGHTYSTPYDSLIIAAGAGQ

SYFGNDHFAEFAPGMKSIDDALELRGRILGAFEQAERSSDPVRRAKLLTFTVVGAGPTGV

EMAGQIAELADQTLRGSFRHIDPTEARVILLDAAPAVLPPMGEKLGKKARARLEKMGVEV

QLGAMVTDVDRNGITVKDSDGTIRRIESACKVWSAGVSASPLGKDLAEQSGVELDRAGRV

KVQPDLTLPGHPNVFVVGDMAAVEGVPGVAQGAIQGGRYAAKIIKREVSGTSPKIRTPFE

YFDKGSMATVSRFSAVAKVGPVEFAGFFAWLCWLVLHLVYLVGFKTKIVTLLSWGVTFLS

TKRGQLTITEQQAYARTRIEELEEIAAAVQDTEKAAS