CLONING, EXPRESSION AND PURIFICATION OF RECOMBINANT ELONGATION FACTOR Tu FROM HYPERTHERMOPHILIC BACTERIA *GEOBACILLUS ANATOLICUS*

by

Barış Akalın

B.S. in Biology, Istanbul University, 2003

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Molecular Biology and Genetics
Boğaziçi University
2006

Dedicated to humanity

ACKNOWLEDGEMENTS

This work is supported by a grant from Boğaziçi University Research Fond (BAP project no: 03S102).

I want to thank my thesis advisor Prof. Neş'e Bilgin wholeheartedly. She made this work possible and I think that she represents the future of humanity in terms of human mind.

I'm grateful to Prof. Zehra Sayers and Assoc. Prof. Esra Battaloğlu. They accepted to become my thesis jury and spent their valuable time to evaluate my work.

I want to thank my lab mate and my thesis twin Toros Şahin. We complemented each other and whenever I needed a hand, he was always ready to help. As the founder of "Hattian" philosophy of life, his most insightful quote will always echo in our ears: "As good as it precipitates" or shortly "haaaaaaatti"…

I want to thank my lab mates Tunç Öztunç and Mehmet Doğan. They always shared their experiences with me without any hesitation. I'm really grateful to them and don't forget about Ferrari staff.

Thanks to my lab mate Melike Çağlayan… As the most diligent scientist in our lab, she was always a role model for me… I will never forget what you have done for me.

I want to thank my family and my eternal friend Onur Yazgan for their sincere support and encouragement.

Special thanks to my ex-loves Bilge Şipal and Nıvart Taşçı. You know what I mean.

I want to thank Nadire Duru "the blessed of Dionysos". She was always there whenever I needed her. Is this an eternal love? We still don't know what Suzanne is going to do… :)

ABSTRACT

CLONING, EXPRESSION AND PURIFICATION OF RECOMBINANT ELONGATION FACTOR Tu FROM HYPERTHERMOPHILIC BACTERIA *GEOBACILLUS ANATOLICUS*

In this study, elongation factor Tu (EF-Tu) from *Geobacillus anatolicus* was cloned, expressed and purifed as a His-tagged recombinant protein in *Escherichia coli*. At first, the nucleic acid sequence of *Geobacillus anatolicus tuf* gene was determined by using sequence data obtained from the bacteria phylogenetically related to *Geobacillus anatolicus.* The *tuf* gene and its neighbouring genes of the *tuf* gene sequences were used to design appropriate oligonucleotide primers and these primers were used to amplify the chromosomal region carrying the *tuf* gene of *Geobacillus anatolicus*. The sequence information from this fragment was then used to design primers to sequence and clone the complete *tuf* gene into an appropriate expression vector. This vector was used to express and purify the protein in high quantities in *Escherichia coli*. Six additional histidine residues were inserted into the recombinant protein of *Geobacillus anatolicus* at its C-terminal region to purify the protein using Ni-affinity chromotography to homogeneity. It was found that recombinant *Geobacillus anatolicus* EF-Tu is fully competent in forming a binary complex with *Geobacillus anatolicus* EF-Ts. It was determined that it has a significantly higher mobility rate in non-denaturing PAGE as compared to *Escherichia coli* EF-Tu, not explainable with molecular mass difference between these proteins. The mobility of the *Geobacillus anatolicus* EF-Tu.EF-Ts complex does not differ from the mobility of the *Geobacillus anatolicus* EF-Ts, indicating a large conformational change of EF-Tu upon binding to EF-Ts. The recombinant *Geobacillus anatolicus* EF-Tu can form a ternary complex with GTP and Phe-$tRNA^{Phe}$. *Geobacillus anatolicus* retained its GDP binding activity fully after incubation at 60°C for 10 min, and it still has approximately %20 of its activity after incubating for 10 min at 80°C.

ÖZET

HİPERTERMOFİLİK *GEOBACILLUS ANATOLICUS* BAKTERİLERİNDEN REKOMBİNANT ELONGASYON FAKTÖRÜ Tu'NUN KLONLANMASI, EKSPRESYONU VE SAFLAŞTIRILMASI

Bu çalışmada, *Geobacillus anatolicus* elongasyon faktörü Tu klonlanmış, *Escherichia coli*'de ekspresyonu yapılmış ve His-kuyruklu bir rekombinant protein olarak saflaştırılmıştır. İlk olarak, *Geobacillus anatolicus tuf* geninin dizisi belirlenmiştir. *Geobacillus anatolicus tuf* geninin dizisinin belirlenmesi için, *Geobacillus anatolicus*'a filogenetik olarak yakın bakterilerin *tuf* geni ve *tuf* geninin etrafındaki genler için belirlenmiş dizi bilgileri kullanılarak uygun oligonukleotit primerler tasarlanmıştır ve bu oligonükleotit primerler kullanılarak *tuf* genini taşıyan kromozomal bölge çoğaltılmıştır. Elde edilen dizi bilgisi daha sonra *tuf* geninin uygun bir ekspresyon vektörüne klonlanması için kullanılmıştır. Rekombinant proteinin daha sonra Ni-afinite kromatografisi ile saflaştırılabilmesi için vektör üzerinde klonlanan proteinin C-terminal bölgesine 6 His aminoasidi eklenmiştir. Bu vektör aracılığı ile *Geobacillus anatolicus* EF-Tu'sunun *Escherichia coli* içinde yüksek miktarda ekspresyonu sağlanmış ve Ni-afinite kromatografisi kullanılarak saflaştırılmıştır . *Geobacillus anatolicus*'tan elde edilen EF-Tu'nun *Geobacillus anatolicus* EF-Ts proteini ile etkileşime girebildiği gösterilmiştir. *Geobacillus anatolicus* EF-Tu'su, *Escherichia coli* EF-Tu'suna göre denatüre etmeyen poliakrilamid elekroforezinde anlamlı derecede hızlı bir ilerleme hızına sahiptir. *Geobacillus anatolicus* EF-Tu.EF-Ts bileşiği ile *Geobacillus anatolicus* EF-Ts'sinin denatüre etmeyen poliakrilamid elektroforezinde aynı ilerleme hızına sahip olması, bileşiğin oluşması sırasında EF-Tu proteininin önemli yapısal değişime uğradığına işaret etmektedir. Rekombinant *Geobacillus anatolicus* EF-Tu'sunun GTP ve Phe-tRNA^{Phe} ile üçlü bileşik oluşturabildiği de yine denatüre etmeyen poliakrilamid elektroforezi ile gösterilmiştir. *Geobacillus anatolicus* EF-Tu'su 60°C'de 10 dakika inkübe edildiğinde GDP bağlama aktivitesini korumaktadır. Bu değer 80°C'de yaklaşık olarak %20'ye inmektedir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv
ABSTRACT v
ÖZET vi
LIST OF FIGURES xi
LIST OF TABLES xiv
LIST OF ABBREVIATIONS xv
1. INTRODUCTION 1
1.1. Protein Synthesis 1
1.1.1. Initiation 4
1.1.2. Elongation 5
1.1.3. Termination and Recycling 8
1.2. Elongation Factor Tu (EF-Tu) 9
1.3. Location of EF-Tu Gene on the Bacterial Chromosome 11
1.4. *Geobacillus anatolicus* 11
1.5. Hyperthermophiles 13
1.6. Thermostability of Proteins in Hyperthermophiles 14
1.7. Biotechnological Implications for Hyperthermophilic Enzymes 17
2. PURPOSE 18
3. MATERIALS 19
3.1. Bacterial Strains 19
3.1.1. *Geobacillus anatolicus* 19
3.1.2. *Escherichia coli* 19
3.2. Oligonucleotide Primers 19
3.3. Enzymes Used Throughout This Study 20
3.4. Chemicals 21
3.5. Buffers and Solutions 23
3.5.1. Bacterial Growth 23
3.5.1.1. Stock Solutions for Growth Medium for *Geobacillus anatolicus* 23
3.5.1.2. Growth Medium for *Escherichia coli* JM109 (DE3)

Strain 24
3.5.2. Stock Solutions for Genomic DNA Isolation 24
3.5.3. Stock Solutions for Polymerase Chain Reaction (PCR) 25
3.5.4. Stock Solutions for Agarose Gel Electrophoresis 25
3.5.5. Stock Solutions for Enzyme Digestions 25
3.5.6. Stock Solutions for Competent Cell Preparation 26
3.5.7. Stock Solutions for Transformation 26
3.5.8. Stock Solutions for Plasmid DNA Isolation 27
3.5.9. Stock Solutions for SDS-Polyacrylamide Gel Electrophoresis ... 27
3.5.10. Stock Solutions for Non-denaturing Polyacrylamide Gel Electrophoresis for Ternary Complex Analysis 29
3.5.11. Buffer for Cell Lysis and Homogenization…………………… 30
3.5.12. Stock Solutions for Non-denaturing Polyacrylamide Gel Electrophoresis for EF-Tu.EF-Ts Binding……………………. 30
3.5.13. Buffers for Ni-NTA Affinity Chromotography 30
3.5.14. Polymix Buffer for Dialysis, Storage and Assay 31
3.6. Expression Vector 32
3.7. PCR Product Purification 33
3.8. Equipments 33
4. METHODS 35
4.1. Extraction of the Genomic DNA from *Geobacillus anatolicus* 35
4.2. Analysis of Extracted DNA by Agorose Gel Electropheresis.................... 36
4.3. Analysis of *Geobacillus anatolicus* Elongation Factor Gene (*tuf*)………… 36
4.3.1. Degenerate Primer Design for Unknown *G. anatolicus* *tuf* Gene…………………………………………………………... 36
4.3.2. Primer Design for Partially Known *G. anatolicus tuf* Gene……... 37
4.3.3. Degenerate Primer Design for Adjacent Genes *fus* and *rplC*……. 37
4.3.4. PCR Amplifications of *tuf* and Adjacent Genes *fus* and *rplC*…… 37
4.3.5. Purification of PCR Products……………………………………. 38
4.3.6. DNA sequencing………………………………………………… 39
4.4. Cloning of *G. anatolicus tuf* Gene.. 39
4.4.1. Restriction Enzyme Digestions of Plasmid pCR® T7/NT-TOPO® and PCR Fragment Obtained

Using TuF2 and TuR2 Primers…………………………………… 39
4.4.2. Purification of the Digested PCR Fragment and Plasmid pT7D3…………………………………………………… 40
4.4.3. Ligation of the Digested PCR fragment and Plasmid pT7D3.......... 40
4.5. Transformation of the *Escherichia coli* JM109 (DE3) Cells 41
4.5.1. Preparation of *E. coli* JM109 (DE3) Competent Cells 41
4.5.2. Transformation 41
4.6. Verification of Existence of *tuf* Gene in the Plasmid pT7D3Tu................... 42
4.6.1. Plasmid Purification 42
4.6.2. Confirmation of *tuf* Fragment in the Plasmid pT7D3Tu by Using PCR.. 42
4.6.3. PCR Amplification of the Plasmid pT7D3Tu for DNA Sequencing.. 43
4.6.4. Storage of JM109(DE3) Cells Containing pT7D3Tu..................... 43
4.7. Expression of *G. Anatolicus tuf* Gene in JM109(DE3) Cells........................ 43
4.8. Isolation of *G. anatolicus* EF-Tu.. 44
4.9. Ni-Affinity Column Chromatography 45
4.9.1. Ni-Affinity Column Regeneration.. 45
4.9.2. Purification Step... 45
4.10. Dialysis and Storage of *G. anatolicus* EF-Tu protein................................ 45
4.11. Determination of Protein Concentration 46
4.12 Mobility-Shift Assay on Non-Denaturing PAGE for EF-Tu·EF-Ts complex formation.. 46
4.13. Mobility-Shift Assay on Non-Denaturing PAGE for Ternary Complex Formation between EF-Tu·GTP·Phe-$tRNA^{Phe}$........................... 47
4.14. Thermostability Test.. 47
5. RESULTS 49
5.1. Determination of *Geobacillus anatolicus* Elongation Factor Tu Gene (*tuf*) Sequence.. 49
5.1.1. Extraction of DNA from *Geobacillus anatolicus*........................... 49
5.1.2. Primer design……………………………………………………... 49
5.1.3. Partial PCR Amplification of *tuf* Gene…………………………... 73
5.1.4. Confirmation of Partial *tuf* Sequence…………………………...... 74

5.1.5. Amplification of the Fragment Containing Complete *tuf* Sequence.. 74
5.1.6. Complete *G.anatolicus tuf* Sequence............................... 75
5.2. Cloning of *tuf* Gene of *G. anatolicus*.. 81
5.2.1. Cloning Method.. 81
5.2.2. Constructing Primers for Cloning of *tsf* Gene........................... 82
5.2.3. Construction of pT7D3Tu.. 83
5.2.4. Confirmation of the presence of the *tuf* Gene in the pT7D3Tu vector... 84
5.3. Expression of *G. Anatolicus tuf* Gene in JM109(DE3) Cells.................... 85
5.4. Ni-Affinity Column Chromatography... 86
5.5. Interactions of *G. anatolicus* EF-Tu with EF-Ts............................. 88
5.6. Ternary Complex Formation between G. *anatolicus* EF-Tu, GTP and Phe-tRNA^{Phe}.. 89
5.7. Thermostability Test.. 90
6. DISCUSSION 92
6.1. Determination of *Geobacillus anatolicus tuf* Gene Sequence............... 92
6.2. Number of *tuf* genes.. 98
6.3. Cloning of the *tuf* Gene.. 99
6.4. Expression of the Recombinant Elongation Factor Tu............................. 99
6.5. Purification of the Recombinant His-tagged Elongation Factor Tu.......... 99
6.6. Mobility-Shift Assay for EF-Tu.EF-Ts complex formation..................... 99
6.7. Ternary Complex Formation between G. *anatolicus* EF-Tu, GTP and Phe- tRNA^{Phe}.. 100
6.8. Thermostability... 100
6.9. Conclusion and Future Perspective... 101
7. REFERENCES 102

LIST OF FIGURES

Figure 1.1. *E.coli* ribosome visualized by cryo-EM…………………………………. 2

Figure 1.2. IF-1, IF-2 and IF-3 are shown with their corresponding locations on the ribosome 30S subunit……………………………………………………. 5

Figure 1.3. The functional cycle of EF-Tu…………………………………………… 7

Figure 1.4. Overview of elongation cycle……………………………………………. 7

Figure 1.5. Overview of termination and recycling…………………………………. 9

Figure 1.6. Structures of bacterial elongation factor EF-Tu…………………………. 10

Figure 1.7. Location of *tuf* gene on the bacterial ribosome………………………….. 11

Figure 1.8. Phylogenetic relationship between *Geobacillus* species based on 16S rRNA sequences…………………………………………………………. 12

Figure 3.1. Map of pCR® T7/NT-TOPO® vector.. 32

Figure 5.1. Genomic DNA of *Geobacillus anatolicus*... 49

Figure 5.2. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.1... 50

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2... 60

Figure 5.4. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.3..……………………………………………………………… 71

Figure 5.5. Partial PCR amplification of *G.anatolicus tuf* gene…………………… 74

Figure 5.6. PCR amplification using FusF and RplC primers……………………… 75

Figure 5.7. Confirmation of the presence of the *tuf* gene in the fragment amplified by using FusF and RplC primers…………………………………………... 75

Figure 5.8. Complete *G. anatolicus tuf* sequence... 76

Figure 5.9. Complete *G. anatolicus* EF-Tu amino acid sequence............................ 78

Figure 5.10. Sequence aligment of thermophilic EF-Tu's and G.anatolicus EF-Tu amino acid sequences.. 79

Figure 5.11. pCR®T7/NT-TOPO® cloning sites... 82

Figure 5.12. TuF2, TuR2.. 82

Figure 5.13. Result of the PCR amplification using TuF2 and TuR2 primers……… 83

Figure 5.14. Result of plasmid isolation from transformants……………………….. 84

Figure 5.15. PCR results on the pT7D3Tu plasmid…………………………………. 85

Figure 5.16. Results of expression experiment……………………………………… 86

Figure 5.17. SDS analysis of Ni-Affinity choromotography………………………... 87

Figure 5.18. *G. Anatolicus* EF-Tu…………………………………………………… 88

Figure 5.19. Mobility-shift assay for complex formation between EF-Tu and EF-Ts. 89

Figure 5.20. Ternary complex formation experiment………………………………… 90

Figure 5.21. Thermostability of *G. anatolicus* EF-Tu as judged by GDP binding….. 91

Figure 6.1. Nucleotide sequence comparison between G. kaustophilus and G. anatolicus.. 93

Figure 6.2. Amino acid sequence comparison between G. kaustophilus and G. anatolicus. 96

Figure 6.3. Amino acid sequence comparison and EF-Tu interactions sites.............. 97

LIST OF TABLES

Table 1.1. Gene IDs of the 16S sequences used for construction of the phylogenetic tree in Figure 1.8…………………………………………… 13

Table 1.2. Growth temperatures of several *Geobacillus* species…………………… 14

Table 1.3. Stability of some enzymes at 100°C... 15

Table 1.4. The relative amino acid compositions of proteins from mesophiles and thermophiles.. 16

Table 1.5. Strategies of thermal stabilization of particular thermophilic enzymes... 16

Table 3.1. Genotype of the *Escherichia coli* strain used throughout this study…… 19

Table 3.2. Oligonucleotide primers used for the amplification of *tuf* gene and for DNA sequencing of *tuf* gene.. 20

Table 3.3. Oligonucleotide primers used for cloning of *tuf* gene............................ 20

Table 3.4. Oligonucleotide primers used for plasmid DNA sequencing.................. 20

Table 5.1. Bacterial *tuf* genes used for degenerate primer design............................ 50

Table 5.2. Bacterial *fus* genes used for degenerate primer design............................ 59

Table 5.3. Bacterial *rplC* genes used for degenerate primer design.......................... 70

Table 5.4. NATSTATS result *G. anatolicus tuf* sequence.. 78

LIST OF ABBREVIATIONS

aa	Aminoacid
A	Adenine
ADP	Adenosine 5'-diphosphate
AMP	Adenosine 5'-monophosphate
APS	Ammonium persulphate
Arg	Arginine
Asp	Aspartic acid
ATP	Adenosine 5'-triphosphate
bp	Base pair
BSA	Bovine serum albumin
C	Cytosine
Ca	Calcium
Cys	Cysteine
C-terminal	Carboxyl terminal
D	Aspartic acid
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotides
dsDNA	Double stranded deoxyribonucleic acid
DTE	1,4-Dithioerythritol
E	Glutamic acid
EDTA	Ethylenediaminetetraacetate
EF-G	Elongation factor G
EF-Ts	Elongation factor Ts

EF-Tu	Elongation factor Tu
EtBr	Ethidium bromide
F	Phenylalanine
G	Guanine
GDP	Guanosine diphosphate
Gla	γ-carboxy glutamic acid
Gln	Glutamine
Glu	Glutamate
GTP	Guanosine triphosphate
h	Hour
H	Histidine
His	Histidine
I	Isoleucine
Ile	Isoleucine
IPTG	Iopropyl-1-thio-β-D-galactoside
K	Lysine
kb	Kilo base
kDa	Kilo Dalton
L	Leucine
LB	Luria-Bertani broth
Leu	Leucine
Lys	Lysine
Met	Methionine
Mg	Miligram

Mg	Magnesium
$MgCl_2$	Magneium chloride
min	Minute
mg	Milligram
mM	Millimolar
mRNA	Messenger RNA
N	Asparagine
ng	Nano gram
Ni	Nickel
NTA	Nitrilotriacetic acid
N-terminal	Amino terminal
Nres	Nucleotide residues
PCR	Polymerase chain reaction
PEG	Polyethyleneglycol
pmol	picomole
PMSF	Phenylmethyl ulfonylfluoride
Q	Histidine
R	Arginine
RNA	Ribonucleic acid
rpsB	ribosomal protein S2 gene
rRNA	Ribosomal RNA
S	Serine
sec	Seconds
Ser	Serine
SDS	Sodiumdodecylsulphate

SDS-PAGE	SDS-polyacrylamide gel electropheresis
smbA	uridylate kinase gene
T	Thymine
Taq	*Thermus aquaticus*
TE	Tris-EDTA
TEMED	N,N’,N’,N’-Tetramethylethylenediamine
Thr	Threonine
Tm	Melting temperature
Tris	Tris(hydroxylmethyl)aminomethane
tRNA	Transfer RNA
Trp	Tryptophan
tsf	elongation factor Ts gene
Tyr	Tyrosine
u	Unit
UV	Ultra Violet
V	Valine
Val	Valine
W	Tryptophan
Y	Tyrosine

1. INTRODUCTION

1.1. Protein Synthesis

Life is dependent on protein synthesis. By using this process, living cells produce structural and functional proteins and reproduce themselves. Given this dependency, it can be said that protein synthesis is a central biological process in all living cells. Protein synthesis takes place on ribosomes where the genetic information transcribed into mRNA is translated into protein. (Clark *et al.*, 1999).

Translational machinery is composed of ribosomes, mRNAs, tRNAs and various protein factors known as translation factors. Peptidyl-transferase activity is intrinsic to the ribosomes and resides in the 23 S ribosomal RNA (rRNA) in ribosomes (Marintchev and Wagner, 2005).

In prokaryotes, the ribosome is made up of a large subunit, referred to as 50S, and a small subunit, referred to as 30S. The 30S subunit binds mRNA and controls whether the codon–anticodon interaction between tRNA and mRNA is correct or not. The 50S subunit performs peptide bond formation and interacts with the acceptor end of the aminoacyl-tRNAs. The two subunits therefore, based on rRNA and ribosomal protein content, perform different tasks and contact with the L-shaped tRNA at the opposite ends. On the ribosome there are three different tRNA-binding sites named A-, P- and E- sites (Figure 1.1) (Nilsson and Nissen, 2005).

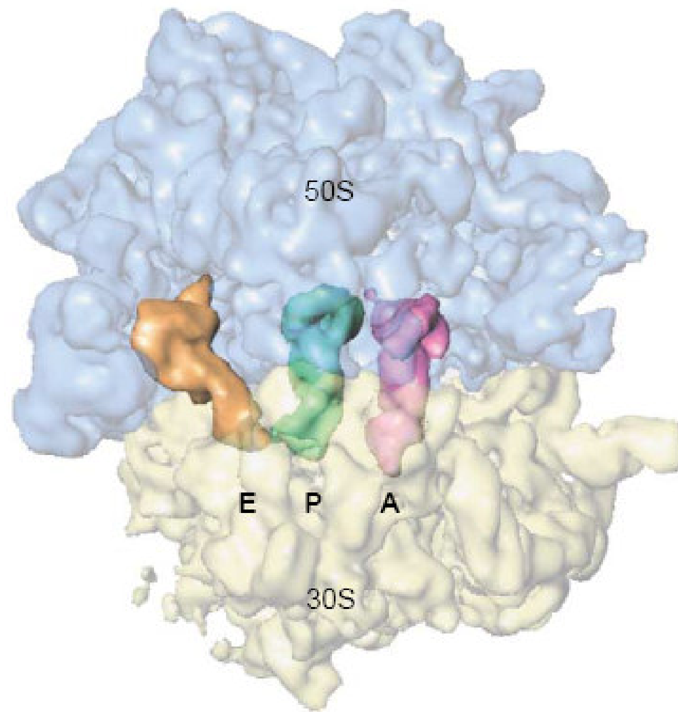

Figure 1.1. *E.coli* ribosome visualized by cryo-EM. There are three tRNAs on the A-, P- and E-sites (Nilsson and Nissen, 2005) (PDB: 1QZB)

In the aminoacyl (A) site, aminoacyl-tRNA (aa-tRNA) is accommodated; at the A site of the small subunit - the decoding center -, it was checked whether interaction of the anticodon of tRNA with the codon on mRNA is correct or not and after that tRNA is rejected or accepted.

In the peptidyl (P) site, peptidyl-tRNA is accommodated through codon-anticodon interactions on the small subunit. Another functional center, the peptidyltransferase center is present as a pocket-like structure on the large subunit. Following peptidyl transfer the elongated peptidyl-tRNA in the A-site is translocated to the P-site while deacylated tRNA in the P-site moves to the E-site.

In the exit (E) site, deacylated tRNA exits ribosome (Frank, 2003).

There are four major steps in the protein synthesis process: initiation, elongation, termination and recycling.

Translation elongation factors can be considered molecular motors in the protein synthesis on the ribosome. In the presence of them, amino acids are added to the nascent polypeptide chain and the polypeptide chain elongates. Because translation is a crucial process in terms of liveliness, general biochemical structure of the translation is well preserved in all biological kingdoms (Andersen *et al*., 2003).

Among bacterial translational elongation factors, elongation factor thermo unstable (EF-Tu) has special significance. In translation elongation cycle, its major role is the binding and transporting of the appropriate codon-specified aminoacyl-tRNA to the aminoacyl site of the ribosome. In performing its function, it forms a ternary complex with GTP and aminoacyl-tRNA and it interacts with ribosomes. A second elongation factor, EF-Ts is responsible for GDP/GTP exchange on EF-Tu (Richarme, 1998).

In *E.coli*, EF-Tu is a monomeric protein of 43.1 kD (Young and Bernlohr, 1991). It is one of the most abundant proteins in the cell such that approximately 10% of the total protein content is EF-Tu. It was determined that there is about one EF-Tu molecule for every tRNA molecule under different growth conditions (Gouy and Grantham, 1980).

In addition of its crucial role in polypeptide elongation, EF-Tu can have other functions. It was determined that EF-Tu is an essential host-donated subunit of the Qβ replicative complex, the transcriptional apparatus and membranes; it has a role in the degradation of N-terminally blocked proteins by 26S proteasome; in the presence of stress factors, EF-Tu can display chaperone-like properties in protein folding and renaturation and it has protein disulfide isomerase activity (Richarme, 1998).

As a crucial part of an ancient and common biological process and as a protein highly conserved across the three major kingdoms of life, EF-Tu can be used as a tool for construction of a universal tree in phylogeny studies (Baldauf *et al*., 1996).

Because it has a crucial role in translation elongation cycle and it may have diverse possible functions, to gain insight into fundamental principals of protein structure and function, EF-Tu is being extensively studied in a wide range of species.

1.1.1. Initiation

Initiation is a process where start codon on the mRNA is correctly located and placed on the P-site of the ribosomes with the help of initiation factors and initiator tRNA. Prokaryotic protein synthesis is the most studied protein synthesis system. The initiation phase of prokaryotic protein synthesis is controlled by initiation factors (IFs), IF-1, IF-2 and IF-3. Among them, IF-2 forms a ternary complex with GTP and initiator $tRNA^{fMet}$ (methionyl-tRNA). All three factors and the additional components form an initiation complex; this complex is composed of initiation factors, initiator $tRNA^{fMet}$, the ribosomal subunits and mRNA. Eukaryotic initiation is much more complex and involves a large number of eukaryotic initiation factors (eIFs) (Clark *et al.*, 1999) (Figure 1.2).

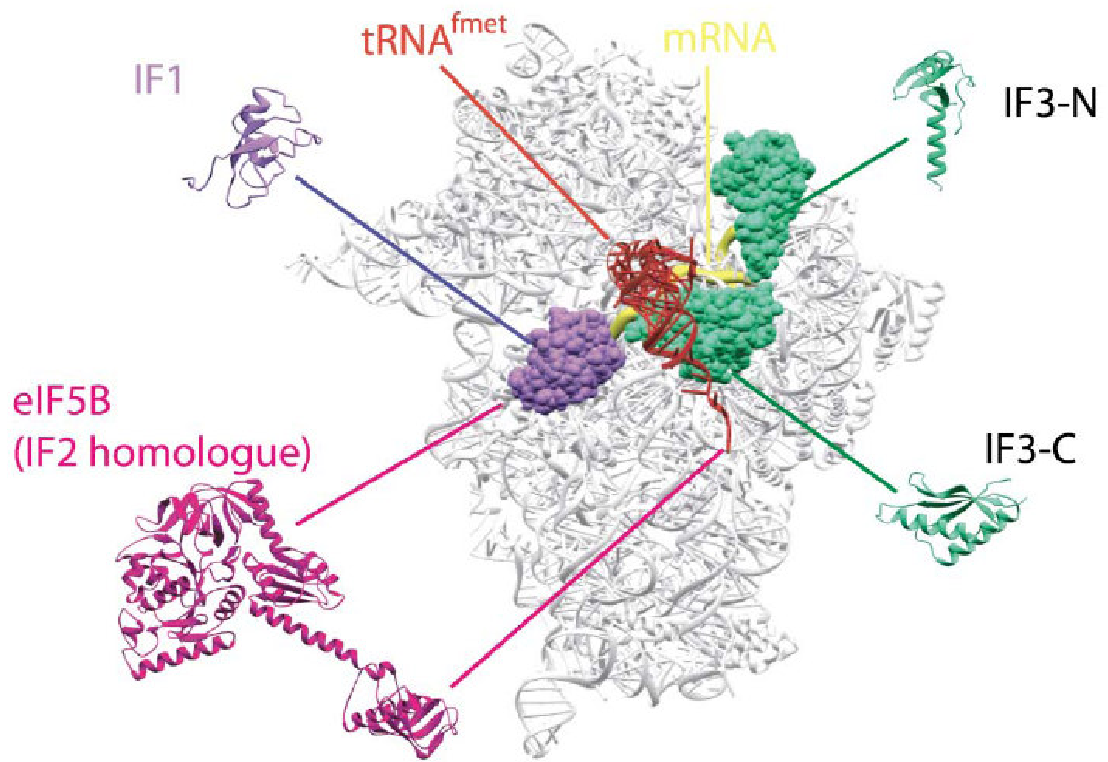

Figure 1.2. IF-1, IF-2 and IF-3 are shown with their corresponding locations on the ribosome 30S subunit (Ramakrishnan, 2002) (PDB: 1HR0 for 30S-IF1)

1.1.2. Elongation

The most intensively studied phase of the translation process has been the elongation phase and it has been examined both functionally and structurally over the last 30 years. In the elongation phase ribosome selects the correct aminoacyl-tRNAs through codon-anticodon interactions following the sequence on the mRNA and catalyzes peptide bond formation. There are three elongation factors (EFs) that speed up this process (Krab and Parmeggiani, 1998).

Elongation factor EF-Tu forms a ternary complex with GTP and aminoacylated tRNAs (aa-tRNAs) in procaryotes. Formation of the ternary complex protects the amino acid ester bond against hydrolysis. EF-Tu carries the aa-tRNA to the ribosomal A-site at which the decoding of mRNA by codon-anticodon interaction is performed. If correct

codon-anticodon interaction is present, ribosome stimulates GTP hydrolysis on EF-Tu and EF-Tu:GDP is released (Clark *et al*., 1999).

The nucleotide exchange factor EF-Ts converts EF-Tu:GDP into active EF-Tu:GTP (Figure 1.3).

The accuracy of the aa-tRNA in the A-site is further checked in a proofreading step (Hopfield, 1974). Proofreading continues until the acceptor end of the aa-tRNA fully enter the A-site of the 50S subunit during which incorrectly accepted aa-tRNA's can be released (Valle *et al.*, 2003). When docking of the aa-tRNA is complete, the aa-tRNA contacts with the peptidyl-tRNA in the ribosomal P-site. Then peptide bond formation occurs and the nascent polypeptide chain is elongated by one amino acid (Clark *et al.*, 1999).

The active site for peptide bond formation, the peptidyl transferase (PT) center, is located on the large ribosomal subunit. By using high-resolution crystal structures of the large ribosomal subunits, it has been determined that the PT center, as localized by an intermediate-state analog, is composed of RNA (23S rRNA). This suggests that the PT reaction is catalyzed by RNA (Nissen *et al.*, 2000).

Translocation occurs at different times in the 30S and 50S subunits. This situation is called hybrid-states model. For example, after peptide bond formation, peptidyl-tRNA is at a hybrid site, in which the anticodon end is at the A-site on the 30S subunit, whereas the acceptor end is at the P-site on the 50S subunit (Moazed and Noller, 1989).

Elongation factor, EF-G, forms a complex with GTP and is responsible for the translocation of tRNAs on the ribosome (Clark *et al.*, 1999) (Figure 1.3.).

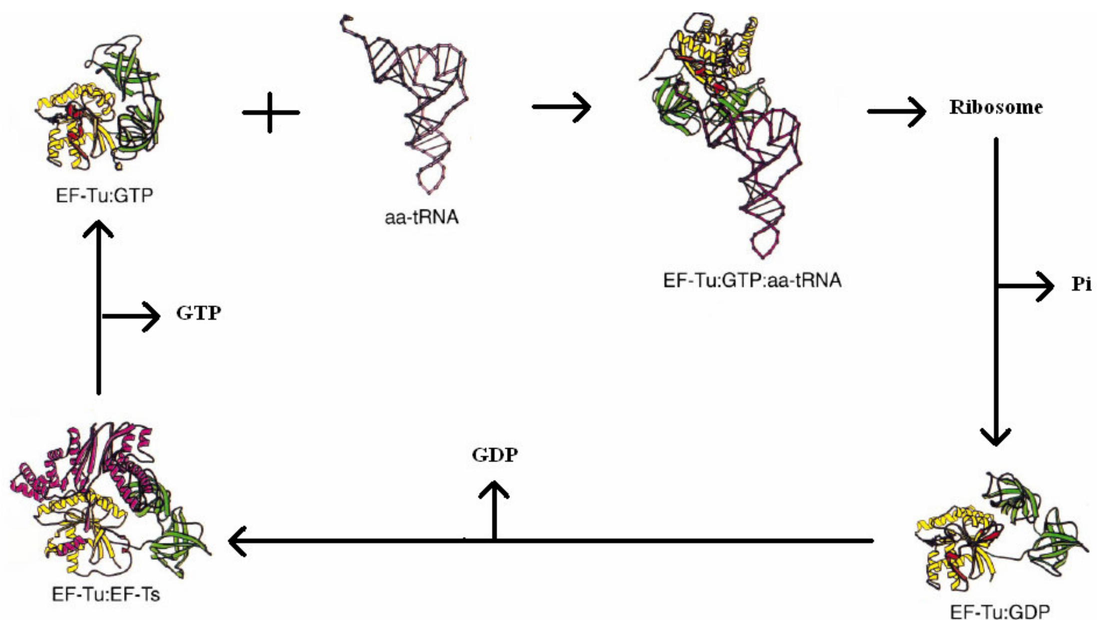

Figure 1.3. The functional cycle of EF-Tu (adapted from Clark *et al.*, 1999)

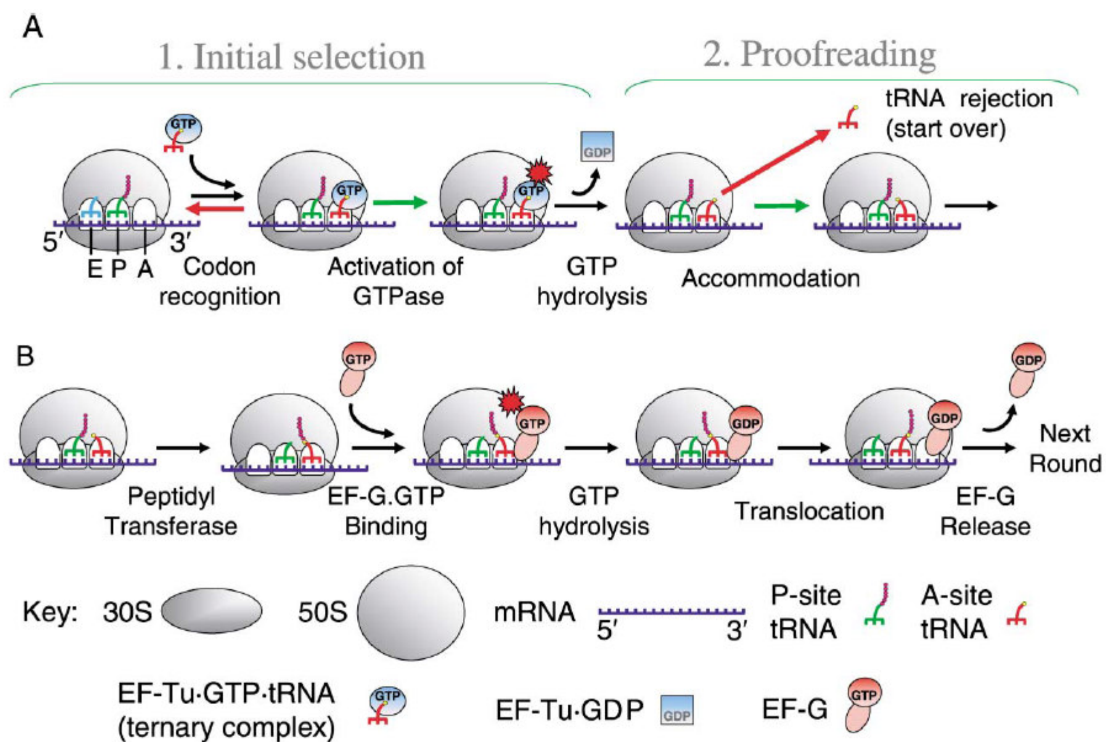

Figure 1.4. Overview of elongation cycle (Ramakrishnan, 2002)

1.1.3. Termination and Recycling

Termination phase of the protein synthesis begins when any one of the stop codons (UAA, UAG, UGA) on the mRNA enters to the A site on the ribosome. In bacteria, recognition of the stop codon is performed by two “class I” release factors, RF1 and RF2. Both factors recognize UAA codon; RF1 additionally recognize UAG and RF2 additionally recognize UGA codons (Kisselev and Buckingham, 2000). A “class II” release factor, RF3, which is a GTPase protein, binds to the complex of RF1/2 with the ribosome (Ramakrishnan, 2002).

When the appropriate stop codon is present in the A site of the ribosome and the RF1/2 complex binds to the ribosome, the hydrolysis reaction occurs on the tRNA and the peptide chain is released from the tRNA in the P site. Dissociation of RF1 and RF2 is facilitated by RF3. GTP doesn’t bind to the RF3 on the ribosome until the hydrolysis of peptidyl tRNA by RF1/2 is completed. After binding of GTP to the RF3, RF3 switches into a new conformation; this conformation has a high affinity for ribosomes and facilitates the dissociation of RF1/2 from the ribosome. The hydrolysis of GTP is required for subsequent dissociation of RF3 (Ramakrishnan, 2002).

After the polypeptide chain is released, mRNA and the deacylated tRNA are still on the ribosome. For the ribosome to begin a new cycle of protein synthesis, this complex should be disassembled. Ribosome recycling factor (RRF) along with EF-G is required for this process. By using GTP hydrolysis, RRF and EF-G lead to the dissociation of ribosomes into its subunits (Gao *et al.*, 2005) (Figure 1.5).

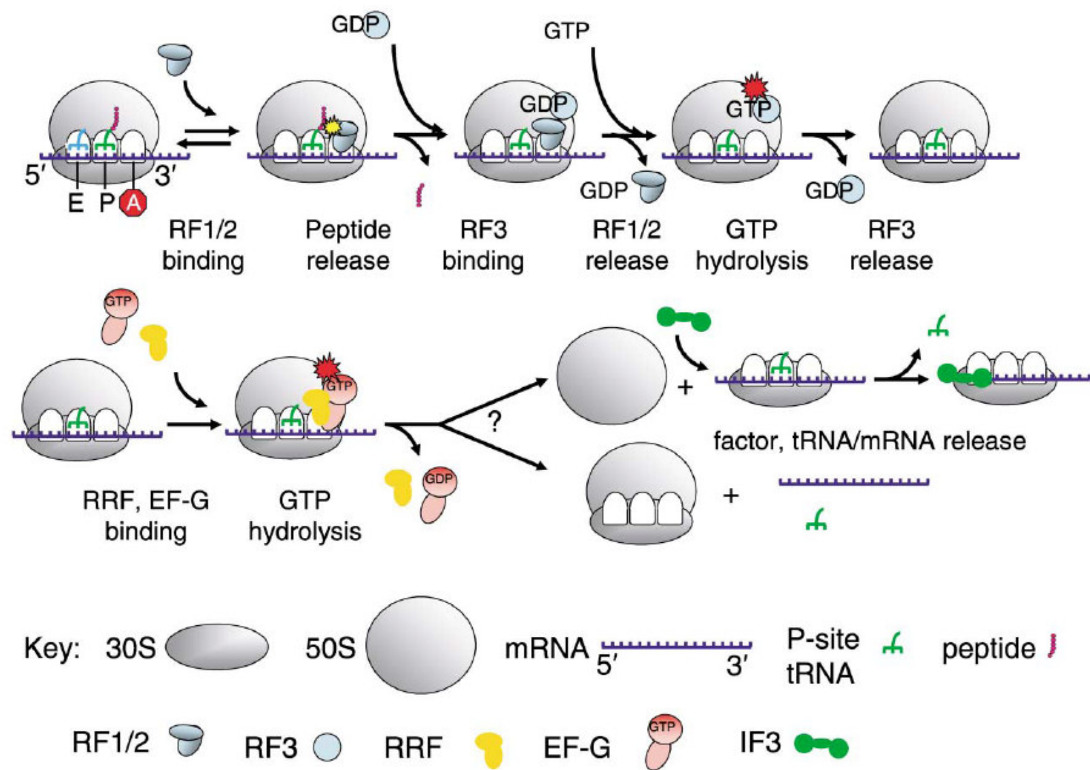

Figure 1.5. Overview of termination and recycling (Ramakrishnan, 2002)

1.2. Elongation Factor Tu (EF-Tu)

The elongation factor that is responsible for carrying aa-tRNA's to ribosomes is called EF-Tu. In the recently changed nomanclature it is called EF-1A. EF-Tu is composed of three structural domains. Domain 1 (referred to as G-domain) is about 200 amino acid long and responsible for binding of GDP or GTP; this domain is a typical nucleotide-binding domain; it has a central β-sheet surrounded by α-helices. It is structurally highly conserved as in the other G-proteins (Kjeldgaard *et al.*, 1996). Domains 2 and 3 are about 100 residues long. Each one of them is β-barrels and based on structure studies, it was determined that these domains are kept together by strong interdomain interactions as one structural unit. EF-Tu has an extraordinary ability; its two structural units (domain 1 and domains 2 and 3 together) may have large variations in their relative orientations. The structures of EF-Tu:GDP and EF-Tu:GTP are very different from each other (Kjeldgaard *et al.*, 1996). This difference was provided by structural changes in two regions on the G-domain of EF-Tu called switch I and switch II. Switch I changes from a β-hairpin in EF-

Tu:GDP to a short α-helix in EF-Tu:GTP. In switch II, the β-helix moves on the sequence by approximately four residues and the axes of the helix are rotated by about 45°. Because this helix (called helix B) is part of the interface between domains 1 and 3, after its activation, EF-Tu can show large conformational changes. The change in helix B correlates with the introduction of a γ-phosphate in the nucleotide-binding site; this phosphate induces an 180° peptide flip at a conserved Gly just prior to the α-helix (Clark *et al.*, 1999).

Crystal structures of EF-Tu-GDPNP (where GDPNP is a non-hydrolyzable analog of GTP) have been determined to high resolution from both *Thermus thermophilus* (Berchtold *et al.*, 1993) and *Thermus aquaticus* (Kjeldgaard *et al.*, 1993). Similarly, crystal structures of EF-Tu-GDP have been determined to high resolution from *Escherichia coli* (Abel *et al.*, 1996) and *T. aquaticus* (Polekhina *et al.*, 1996) (Figure 1.6).

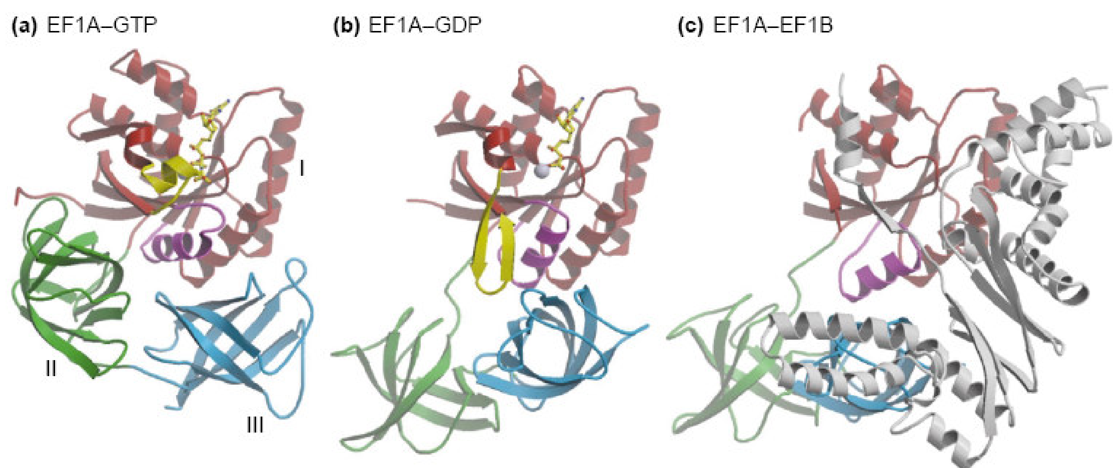

Figure 1.6. Structures of bacterial elongation factor EF-Tu. (a) EF-Tu–GDPNP from *Thermus aquaticus* (PDB: 1TTT), (b) EF-Tu–GDP from *Escherichia coli* (PDB: 1EFC), and (c) the EF-Tu–EF-Ts complex from *E. coli* (PDB: 1EFU). Domain I of EF-Tu is red, domain II is green and domain III is blue. EF-Ts is gray. The switch 1 region of EF-Tu is yellow and switch 2 is magenta. Nucleotides are shown in ball-and-stick models. Mg^{2+} ions are show as gray balls (Andersen *et al.*, 2003)

1.3. Location of EF-Tu Gene on the Bacterial Chromosome

In many organisms, the EF-Tu gene (*tuf*) is the most distal part of the *streptomycin* (*str*) operon which is composed of the following four successive genes: 5'-*rps*L(S12)-*rps*G(S7)-*fus*(EF-G)-*tuf*(EF-Tu)-3'. In *B. stearothermophilus* it was determined that *tuf* gene is in the same place on the *str* operon and is between *fus* gene and *rpl*C gene of the S10 operon which follows the *str* operon (Figure 1.7) (Krasny *et al.*, 1998).

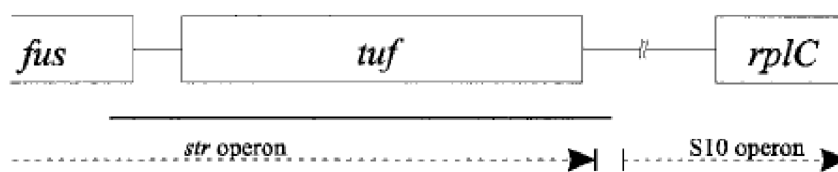

Figure 1.7. Location of *tuf* gene on the bacterial chromosome (Krasny *et al.*, 1998)

1.4. *Geobacillus anatolicus*

Bacillaceae is a large and diverse family of bacteria including various aerobic and facultatively anaerobic, rod-shaped, gram-positive to gram variable, endospore-forming species. This family includes various thermophilic, psychrophilic, acidophilic, alkalophilic, freshwater and halophilic bacteria. *Geobacillus* is recently identified within this family (Nazina *et al.*, 2001).

23 *Geobacillus* species have so far been defined and registered into the Bacterial GenBank database. In this study, as a source of EF-Tu, a hyperthermophilic *Geobacillus* species *Geobacillus anatolicus* was used. *Geobacillus anatolicus* was isolated from a thermal vent in Balıkesir, at 98°C, and based on 16S rRNA sequence comparisons, identified as a new hyperthermophilic bacteria in 2001 (Uysal *et al.*, 2001). No sequence information has been available for the translational components of the *Geobacillus* species when this work had started. Only recently the genome sequence of *Geobacillus kaustophilus* has been completed (Takami *et al.*, 2004).

A phylogenetic relationship between all known *Geobacillus* species including *G. anatolicus* in this genus was drawn by using ClustalW multiple aligment technique based on the 16S rRNA sequence data available for these species in the GeneBank (Figure 1.8.). In Table 1.1., gene IDs of the 16S sequences used for construction of the phylogenetic tree were shown.

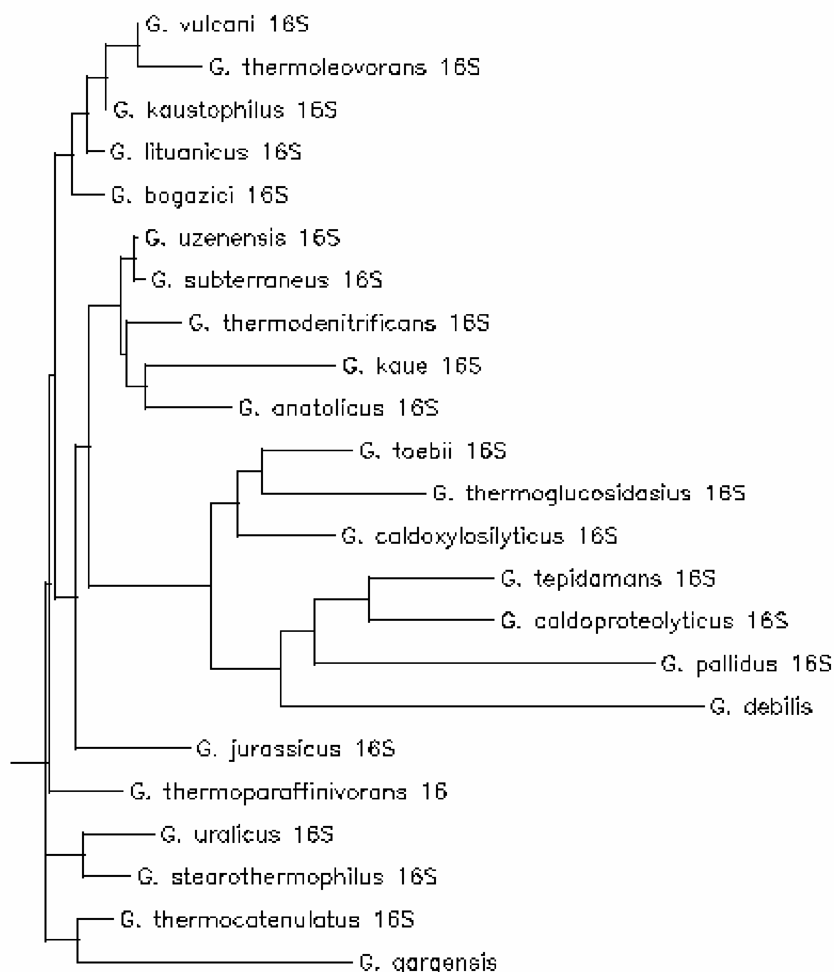

Figure 1.8. Phylogenetic relationship between *Geobacillus* species based on 16S rRNA sequences

Table 1.1. Gene IDs of the 16S sequences used for construction of the phylogenetic tree in Figure 1.8

Organism name	**GeneBank ID**	**Organism name**	**GeneBank ID**
G. vulcani	11877135	*G. caldoxylosilyticus*	54261298
G. thermoleovorans	6406842	*G. tepidamans*	45272587
G. kaustophilus	49482179	*G. caldoproteolyticus*	37538487
G. lituanicus	49482190	*G. pallidus*	57753943
G. bogazici	32815087	*G. debilis*	31414367
G. uzenensis	49482204	*G. jurassicus*	32442349
G. subterraneus	49482203	*G. thermoparaffinivorans*	83637833
G. thermodenitrificans	54261341	*G. uralicus*	22651495
G. kaue	16118568	*G. stearothermophilus*	82655168
G. anatolicus	16118566	*G. thermocatenulatus*	49482180
G. toebii	77819901	*G. gargensis*	28395524
G. thermoglucosidasius	49688769		

1.5. Hyperthermophiles

Hyperthermophilic organisms represent the organisms at the upper temperature border of life. Hyperthermophiles belong to phylogenetically distant groups and may represent rather ancient adaptations to heat. As a rule, they grow fastest (optimally) between 80°C and 110°C. Due to their heat tolerence there is an increasing interest for applications in biotechnological processes (Huber and Stetter, 1998). Growth temperatures of several *Geobacillus* species were shown in Table 1.2.

Table 1.2. Growth temperatures of several *Geobacillus* species (adapted from Zeigler, 2005).

Organism name	**Growth temperature (°C)**
Geobacillus stearothermophilus	37ºC - 65ºC
Geobacillus kaustophilus	37ºC - 68ºC
Geobacillus subterraneus	60ºC - 75ºC
Geoacillus uzenensis	60ºC - 70ºC
Geobacillus thermodenitrificans	45ºC - 70ºC
Geobacillus thermoglucosidasius	37ºC - 68ºC
Geobacillus thermoleovorans	42ºC - 75ºC
Geobacillus anatolicus	65ºC - 98ºC

1.6. Thermostability of Proteins in Hyperthermophiles

Normally, proteins from mesophilic organisms quickly become inactivated above the growth temperatures of these organisms (10-45°C). Accordingly, proteins produced by thermophilic microorganisms are more resistant to thermal and chemical denaturation than their mesophilic counterparts (Scandurra *et al.*, 1998). Proteins from hyperthermophilic organisms show extraordinary stability even at 100°C (Table 1.3.).

Table 1.3. Stability of some enzymes at 100°C (adapted from Daniel and Cowan, 2000 and Huber and Stetter, 1998)

Enzyme	**Source Organism**	**Growth temperature**	**$T_{1/2}$ at 100°C**
Cellobiohydrolase	*Thermotoga*	90 °C	> 200 min
β-Glucosidase	*Thermotoga*	90 °C	90 min
Xylanase	*Thermotoga*	90 °C	20 min
Xylosidase	*Thermotoga*	90 °C	150 min
Esterase	*Sulfolobus*	85 °C	60 min
Hydrogenase	*Pyrococcus*	96 °C	120 min
Amylase	*Pyrococcus*	96 °C	360 min
DNA-dependent RNA polymerase	*Thermoproteus*	97 °C	>120 min

In the course of evolution, organisms have succeeded to survive under extreme conditions. Generally it can be said that there are three strategies for adaptation: escaping or compensating the stress or enhancing the stability of cellular structure and content. In the case of temperature, there is no alternative to mutative adaptation for survival. Proteins can have extremely high stabilities by using relatively small number of weak intermolecular interactions. In this respect, proteins from hyperthermophiles don't differ strongly from their mesophilic counterparts. Although there are no universal rules for thermostability, based on previous studies, several principles could be determined (Jaenicke and Böhm, 1998).

Charge clusters, networks of hydrogen bonds, optimization of packing density and hydrophobic interactions are examples of strategies used by thermophilic organism for protein stability (Jaenicke and Böhm, 1998).

For example, in thermophiles, while number of charged residues such as Glu, Asp, Lys and Arg increases, number of polar residues such as Ser, Thr, Asn and Gln decreases (Fukuchi and Nishikawa, 2001). This observation is called *CvP* bias and was very clearly shown in a study by Sanderova *et al.* in which *E. coli* and *B. stearothermophilus* elongation

factor Tu's were compared (Sanderova *et al.*, 2004). Similarly, in Table 1.4. this bias is presented explicitly.

Table 1.4. The relative amino acid compositions of proteins from mesophiles and thermophiles (adapted from Jaenicke and Böhm, 1998)

Amino acid	**Mesophiles**	**Thermophiles**
Charged residues (DEKRH)	24.11 %	29.84 %
Polar/uncharged residues (GSTNQYC)	31.15 %	26.79 %
Hydrophobic residues (LMIVWPAF)	44.74 %	43.36 %

As mentioned above, there are several methods for the establishing of thermostability. In Table 1.5., for particular proteins some of them were shown.

Table 1.5. Strategies of thermal stabilization of particular thermophilic enzymes (partially adapted from Jaenicke and Böhm, 1998)

Protein (Organism)	**Major cause(s) of thermostability**
CS (*P. furiosus*)	Increased compactness, enhanced subunit interactions, increased number of intersubunit ion pairs, shortening of loops
GAPDH (*T. maritima*)	Large number of additional salt bridges
IPMDH (*T. thermophilus*)	Increased number of ion pairs and hydrogen bonds, extended hydrophobic subunit interactions and improved packing of the hydrophobic core, shortened chain termini
LDH (*T. maritima*)	Increased number of ion pairs, decreased hydrophobic surface area, increased helicity, less cavity volume
Ferredoxin (*S. elongatus*)	Extension of hydrophobic core, a unique hydrophobic patch on the surface of a β sheet, two unique ion-pair networks

1.7. Biotechnological Implications for Hyperthermophilic Enzymes

As a consequence of the growth temperatures of hyperthermophiles, their enzymes are highly temperature stable, making them attractive for new biotechnological applications. There is an unanticipated diversity of species of hyperthermophiles within high temperature environments. Based on their unique properties, hyperthermophiles proteins may play an important role in technical processes (Huber and Stetter, 1998).

Until now, various thermophilic proteins have been commercialized. In molecular biotechnology, *T. aquaticus (Taq)* DNA polymerase (essantial for PCR technology), *P. furiosus (Pfu)* DNA polymerase and DNA ligase and *Thermus* Rt41A serin protease PRETAQ are only a few examples which are widely in use. In industry, especially in strach processing, hyperthermophilic β-amylases, α-glucosidases and pullulanases have great potential. Using hyperthermophilic enzymes in cellulose degredation and ethanol production may support important advancements in these areas. In addition to the direct industrial applications, complete understanding of thermostability in hyperthermophiles is going to make it possible to design enzymes which have specific properties via protein engineering techniques (Vieille and Zeikus, 2001).

2. PURPOSE

Up to now, the question "how hyperthermophilic bacteria perform translation process properly at such high temperatures at which Watson-Crick base-pairing is very weakened?" has not been fully answered. The ultimate aim of this study is to understand translation process at high temperatures. One of the essential enzymes of protein synthesis is EF-Tu. Therefore, the structure and function of EF-Tu which will be purified from *Geobacillus anatolicus*, a newly described hyperthermophilic bacteria isolated from a hot-spring at 98℃ in Turkey, will be studied. For this purpose, the *tuf* gene coding EF-Tu will be cloned, expressed and purified in order to study it functionally in its interactions with EF-Ts, aminoacyl-tRNAs and ribosomes. As a future prospect, the data which will be obtained from this study is going to be used to construct an *in vitro* thermophilic protein synthesis machinary for further analysis of protein synthesis at high temperatures.

3. MATERIALS

3.1. Bacterial Strains

3.1.1. *Geobacillus anatolicus*

Geobacillus anatolicus was used as the source for Elongation factor Tu gene (*tuf*) in molecular cloning studies. *Geobacillus anatolicus* was previously collected from a hydrothermal vent at a water temperature of 98ºC in Hisaralan, Balıkesir, Turkey and it was identified as a new species in our laboratory (GenBank Accession No: AF411064) (Uysal *et al*., 2001).

3.1.2. *Escherichia coli*

Escherichia coli strain JM109(DE3) was used as the plasmid host for cloning and for expression of the *tuf* gene from *Geobacillus anatolicus*. The genotype of the JM109(DE3) is given in Table 3.1.

Table 3.1. Genotype of the *Escherichia coli* strain used throughout this study

JM1O9(DE3)	*end*A1, *rec*A1, *gyr*A96, *thi*, *hsd*R17 (rk^-,mk^+), *rel*A1, *sup*E44, *mcr*AΔ(*lac*-*pro*AB), [F′, *tra*D36, *pro*AB, *laq*[^{q}ZΔM15], λDE3

3.2. Oligonucleotide Primers

Oligonucleotide primers used throughout this study were shown in Table 3.2., 3.3. and 3.4.

Table 3.2. Oligonucleotide primers used for the amplification of *tuf* gene and for DNA sequencing of *tuf* gene
(V: any base except T, R: A or G, N: a base, M: C or A, H: any base except G, Y: C or T, W: T or A, D: any base except C)

TuF413	5'- GCC VCA RAC NMG NGA RCA -3'
TuR1174	5'- TCN CCN GGC ATN ACC AT - 3'
TuF888	5'- AGA AAT GTT CCG TAA GCT GC - 3'
TuR766	5'- ATC GAG AAA ACG TCC TCA AC - 3'
FusF	5'- CHG ARA TGT TYG GWT AYG CDA C - 3'
RplC	5'- CCR ATY TTW CKW CCT AAG ATT CCT TT - 3'

Table 3.3. Oligonucleotide primers used for cloning of *tuf* gene

TuF2	5'- CGG GGG GCA TAT GGC TAA AGC GAA ATT TGA GCG TAC GAA A - 3'
TuR2	5'- CAA AAA ATT CGA ATT AAT GAT GAT GAT GAT GAT GCT CAA TGA TTT CCG ATA CGG AAC CAG C – 3'

Table 3.4. Oligonucleotide primers used for plasmid DNA sequencing

T7 Forward	5'-TAA TAC GAC TCA CTA TAG GG
pRSET Reverse	5'-TAG TTA TTG CTC AGC GGT GG

3.3. Enzymes Used Throughout This Study

Taq DNA polymerase (EC# 2.7.7.7) : Promega, USA

Nde I (EC# 3.1.21.4) : Promega, USA

Csp45 I (EC# 3.1.21.4) : Promega, USA

T4 DNA Ligase(EC # 6.5.1.1) : Promega, USA

Lysozyme (EC # 3.2.1.17)	:	Appligene, USA
RNase-A (EC # 3.1.27.5)	:	Promega, USA
Proteinase K (EC # 3.4.21.64)	:	Promega, USA

3.4. Chemicals

Agar	:	Oxoid, United Kingdom
Agarose	:	Sigma, USA
Ampicillin	:	Mustafa Nevzat, Turkey
Ammoniumpersulfate (APS)	:	Saveen, Sweden
Bromophenol Blue	:	Sigma, USA
Chloroform	:	Ambresco, USA
Coomassie Brilliant Blue R-250	:	Sigma, USA
Crystal Violet	:	Merck, Germany
D-Glucose	:	Riedel de-Häen, Germany
DNA molecular size markers	:	100 bp (Promega, USA) 1 kb (Promega, USA)
Dithioerythritol (DTE)	:	Fluka BioChemika, USA
Ethidium Bromide	:	Sigma, USA

Glacial Acetic Acid	:	Merck, Germany
Glass beads (0.1 mm diameter)	:	Biospec Products INC., USA
Glycerol	:	Merck, Germany
Guanosine 5'-diphosphate (GDP)	:	Sigma, USA
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	:	Sigma, USA
Imidazole	:	Sigma, USA
Isopropyl-ß-D-thiogalactopyranoside (IPTG)	:	Saveen, Sweden
β-Mercaptoethanol	:	Merck, Germany
Nitrilotriacetic acid (NTA)	:	Fluka, Switzerland
N,N,N’,N’-Tetramethyl-ethylenediamine (TEMED)	:	Sigma, USA
Phenylmethylsulfonylfluoride (PMSF)	:	Sigma, USA
Phenol	:	Riedel de-Häen, Germany
Phenol Red	:	Sigma, USA
Polyethylene glycol (PEG 6000)	:	Merck, Germany
Protein Molecular Weight Markers	:	PageRuler Prestained Protein Ladder (Fermentas, USA)

		PageRuler Protein Ladder (Fermentas, USA)
Sodium Dodecyl Sulfate (SDS)	:	Kebolab, USA
Sucrose	:	Merck, Germany
Tryptone	:	Oxoid, United Kingdom
Yeast Extract	:	Oxoid, United Kingdom

All other laboratory salts are analytical grade and obtained from Merck and Fluka, Germany.

3.5. Buffers and Solutions

3.5.1. Bacterial Growth

3.5.1.1. Stock Solutions for Growth Medium for *Geobacillus anatolicus*

5X Solution A	:	20 g/l yeast extract 40 g/l tryptone 15 g/l NaCl
10X Solution B	:	1 g/l NTA 0.6 g/l $CaSO_4.2H_2O$ 1 g/l $MgSO_4.7H_2O$ 1 g/l KNO_3 6.9 g/l $NaNO_3$ 1 g/l Na_2HPO_4
100X Solution C	:	28 mg/l $FeCl_3$

100X Solution D	:	220 mg/l $MnSO_4.H_2O$
		50 mg/l $ZnSO_4.7H_2O$
		1.6 mg/l $CuSO_4$
		2.5 mg/l $Na_2MoO_4.2H_2O$
		4.6 mg/l $CoCl_2.6H_2O$
		0.5 ml/l concentrated H_2SO_4

Final liquid culture medium (1X) contained appropriate dilutions of the concentrated stock solutions of A, B, C, and D. After autoclaving, D-glucose is added to a final concentration of 2%.

3.5.1.2. Growth Medium for *Escherichia coli* JM109 (DE3) Strain

1X LB-ampicillin Medium (pH 7.0)	:	5 g/l NaCI
		10 g/l Tryptone
		5 g/l Yeast Extract
		150 μg/ml ampicillin
LB–Ampicillin Agar	:	5 g/l NaCI
		10 g/l Tryptone
		5 g/l Yeast Extract
		14 g/l agar
		150 μg/ml ampicillin

3.5.2. Stock Solutions for Genomic DNA Isolation

Homogenization Buffer	:	50 mM Tris-HCl (pH 8.0)
		20 mM EDTA
20 % SDS	:	200 g/l SDS in H_2O
TE Buffer	:	10 mM Tris-HCl (pH 8.0)
		1 mM EDTA

3.5.3. Stock Solutions for Polymerase Chain Reaction (PCR)

10 X PCR Buffer	:	10 mM Tris-HCl (pH 9) 50 mM KCl 0.1 % Triton-X 100 (Promega, USA)
$MgCl_2$	:	25 mM $MgCl_2$ (Promega, USA)
Deoxyribonucleotide mixture	:	100 mM of each dATP, dGTP, dCTP, dTTP (Promega, USA) in H_2O

3.5.4. Stock Solutions for Agarose Gel Electrophoresis

25 X TAE Buffer	:	121 g/l Tris (Base) 28.55 ml/l Glacial acetic acid 25 mM EDTA (pH 8.0)
1 or 2 % Agarose Gel	:	1 or 2 % (w/v) Agarose in 1 X TAE Buffer, containing 0.5 µg/ml Ethidium Bromide
6 X Loading Buffer	:	10mM Tris-HCl (pH 7.6) 0.03% bromophenol blue 0.03% xylene cyanol FF 60% glycerol 60mM EDTA

3.5.5. Stock Solutions for Enzyme Digestions

10X concantrated buffers for enzymes below were purchased together with the enzymes.

10X Buffer D for Nde I digestion	:	60 mM Tris-HCl (pH 7.9)
		1.5 M NaCl
		60 mM $MgCl_2$
		10 mM Dithiotheitol (DTT)
10X Buffer B for Csp45 I digestion	:	60 mM Tris-HCl (pH 7.5)
		500 mM NaCl
		60 mM $MgCl_2$
		100 mM DTT
10X T4 DNA Ligase Buffer	:	300 mM Tris-HCl (pH 7.8)
		100 mM $MgCl_2$
		100 mM DTT
		10 mM ATP

3.5.6. Stock Solutions for Competent Cell Preparation

TSS Buffer	:	10 % PEG 6000
		5 % DMSO
		20 mM MgSO4
		5 g/l NaCl
		10 g/l Tryptone
		5 g/l Yeast extract

3.5.7. Stock Solutions for Transformation

SOC Medium	:	2 % Tryptone
		0.5 % Yeast Extract
		10 mM NaCl
		2.5 mM KCl
		10 mM $MgCl_2$
		10 mM $MgSO_4$

20 mM D-Glucose

3.5.8. Stock Solutions for Plasmid DNA Isolation

TEL Solution	:	10 mM Tris-HCl (pH 8) 1 mM EDTA 5 mg/ml lysozyme
NAS Solution	:	0.2 M NaOH 1 % SDS
K-acetate Solution (pH 4.8)	:	60 ml/l 5M K-acetate 28.5 ml/l Glacial acetic acid
Resuspension Buffer	:	10 mM Tris-HCl (pH 8.0) 1mM EDTA 0.3 M NaCl

3.5.9. Stock Solutions for SDS-Polyacrylamide Gel Electrophoresis

Protein Lysis Solution	:	167 mM Tris-HCl (pH 6.8) 0.33 M SDS 10 % (w/v) sucrose 25 µl/ml β-mercaptoethanol 0.01 % (w/v) bromophenol blue
Solution I	:	20 % acrylamide 1 % N,N'-methylene-bisacrylamide (37.5/1)
Solution II A	:	3 M Tris (pH 8.9)
Solution II B	:	60 g/l Tris (base)

		68.6 g/l $NaH_2PO_4.2H_2O$ (pH 7.8)
Gel Buffer (1:1)	:	200 ml Solution II A 8 ml 20 % SDS 1 ml TEMED 600 ml H_2O
Spacer	:	41.5 ml Solution I 31.25 ml Solution II B 1.25 ml 20 % SDS 0.25 ml TEMED 207 ml H_2O
5X SDS Sample Buffer	:	5 ml 20 % SDS 4 ml Solution II B 1 ml β-Mercaptoethanol 10 ml glycerol
10 % Ammoniumpersulfate	:	0.1 g/ml APS in H2O
40% Acrylamide Stock (37.5:1)	:	37.5 % Acrylamide 1 % N,N'- methylenebisacrylamide (purchased as a ready-made solution from Biorad, USA)
7X Solution 2C	:	24 g/l Tris (base) 115.2 g/l Glycine
SDS-PAGE Electrophoresis running buffer	:	167 ml/l 7X Solution 2C 5.83 ml/l 20 % SDS
Staining Solution	:	40 % Methanol

		10 % Glacial Acetic acid 1 g/l Coomassie Brilliant Blue (R-250)
Destaining Solution	:	30 % Isopropanol 10 % Glacial acetic acid

3.5.10. Stock Solutions for Non-denaturing Polyacrylamide Gel Electrophoresis for Ternary Complex Analysis

10 X PAGE Buffer for Ternary Complex Analysis	:	100 mM MOPS 50 mM Mg Acetate 650 mM NH_4 Acetate 10 mM EDTA pH = 6.65 (adjusted with KOH)
Gel Buffer for Ternary Complex Analysis	:	5 ml acryl:bis 40 % (19:1) 4 ml 10 x PAGE buffer 400 µl APS 10 % 4 µl 100 mM GTP 200 µl 50 mM DTE
Running Buffer for Ternary Complex Analysis	:	100 ml 10 x PAGE Buffer 100 µl 100 mM GTP 154 mg DTE Add up to 1 lt with dH_2O (prepared ice-cold)

3.5.11. Buffer for Cell Lysis and Homogenization

10X Homogenization buffer	:	500 mM Hepes-KOH (pH 7.5)
		1.5 M KCl
		100 mM $MgCl_2$

For 1X homogenization buffer the following components are added freshly:

10 % Glycerol
7 mM β-Mercaptoethanol
10 μM GDP
0.1 mM PMSF

3.5.12. Stock Solutions for Non-denaturing Polyacrylamide Gel Electrophoresis for EF-Tu.EF-Ts Binding

10 X Gel Buffer for EF-Tu.EF-Ts Binding	:	80 mM Tricine (pH = 8.2)
Gel Buffer for EF-Tu.EF-Ts Binding	:	3 ml 10 x Gel Buffer
		7.5 ml acryl:bis 40 % (37.5:1)
		19.5 ml dH_2O
Running Buffer for EF-Tu.EF-Ts Binding	:	100 ml 10 x Gel Buffer
		900 ml dH_2O
		(prepared ice-cold)

3.5.13. Buffers for Ni-NTA Affinity Chromotography

Buffer A	:	50 mM Hepes KOH (pH 7.5)
		0.15 M KCl

		10 mM $MgCl_2$
		10 % Glycerol
		7 mM β-Mercaptoethanol
		10 μM GDP
		0.1 mM PMSF
Buffer C	:	50 mM Hepes KOH (pH 7.5)
		0.15 M KCl
		10 mM $MgCl_2$
		10 % Glycerol
		7 mM β-Mercaptoethanol
		10 μM GDP
		0.1 mM PMSF
		1 M NH_4Cl
		10 mM imidazole

3.5.14. Polymix Buffer for Dialysis, Storage and Assay

1X Polymix Buffer	:	95 mM KCl
		5 mM NH_4Cl
		5 mM $MgAc_2$. $4H_2O$
		0.5 mM $CaCl_2.2H_2O$
		8 mM Putrescine
		1 mM Spermidine
		5 mM K-phosphate pH 7.5
		1 mM DTE

10 X Polymix buffer is prepared without K-phosphate to avoid precipitation and DTE. In assays, correct working strength of 1X Polymix buffer is obtained by mixing 10X Polymix buffer appropriately with 100 mM K-phosphate and 50 mM DTE. For dialysis DTE is freshy added as solid.

3.6. Expression Vector

In figure 3.1. the original vector used in this study was shown. The other two vectors mentioned below was obtained from this vector.

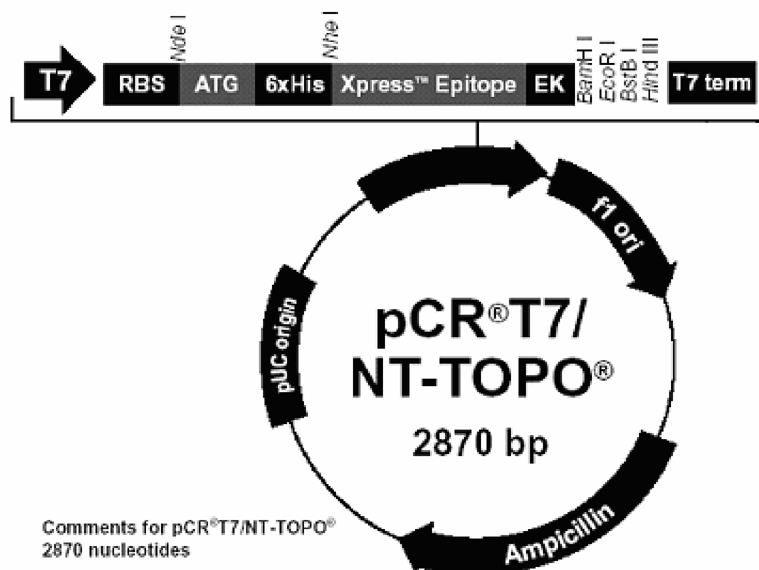

Figure 3.1. Map of pCR® T7/NT-TOPO® vector
(From http://www.invitrogen.com/content/sfs/vectors/pcrt7nttopo_map.pdf)

pCR® T7/NT-TOPO®	:	INVITROGEN, USA
pT7D3	:	Derived from pCR®T7/ NT-TOPO® after Nde I and Csp45 I restriction endonuclease reactions (this work)
pT7D3Tu	:	Derived from pT7D3 after ligation of PCR fragment of 6XHis-tag containing *Geobacillus*

		anatolicus tuf gene (this work)

3.7. PCR Product Purification

Wizard SV Gel and PCR Clean-Up System	:	Promega, USA

3.8. Equipments

Autoclave	:	Model MAC-601 (Eyela, Japan)
Balance	:	MonoBloc Model PB602, USA
Centrifuges	:	Beckman Microfuge E,Germany Beckman L7-55 Ultracentrifuge, USA Beckman J2-MC Centrifuge, USA Beckman J2-21 Centrifuge, USA Centurion K2R, United Kingdom
Rotors	:	Beckman JS-7.5A, JA-14, Germany Beckman 70 Ti, USA
Deep Freezers (-20°C)	:	Arçelik 2031D and 2020D, Turkey
Deep Freezer (-80°C)	:	Sanyo Ultra Low, UK
Refrigerator (4°C)	:	Arçelik 1071D, Turkey
Documentation System	:	BioRAD, USA
Gel Drier	:	Honeywell, UK

Magnetic Stirrer	:	Jankey δ Kunkel KM02, USA
Ovens	:	Microwave Oven (Vestel, Turkey) 110°C (Gallenkamp 300, UK)
Incubator	:	65°C (LEEK, USA)
Power Supplies	:	Betherde Research Model 200 (BMC, Uppsala)
Spectrophotometer	:	Agilent 8453, UV-Visible, USA
Thermocyclers	:	Techne (Progene, UK) Techne Gradient (BioRAD, USA)
UV Transilluminator	:	Chromato-Vue Transilluminator, Model 1TM-20UVP (USA)
Water Baths	:	RE100B, Grant, UK and Innova 3100, USA
Heat-block	:	Grant FE10, Cambridge, UK
pH Meter	:	Jenway 3010, USA
Vortex	:	MS2 Minishaker IKA, USA
Beat Beater	:	Model 1107900, Biospec Products,USA
Liquid Chromotography system	:	Pharmacia, Sweden

4. METHODS

4.1. Extraction of the Genomic DNA from *Geobacillus anatolicus*

Geobacillus anatolicus glycerol stock was taken from -80°C freezer and cells were streaked out on agar plates containing the growth medium at pH 8.5. Agar plates were incubated at 65°C overnight. A toothpick of cells was taken from agar plates and resuspended in a 1.5 ml sterile eppendorf tube containing 750 µl TE buffer. Then 23 µl of 20 % SDS and 9.3 µl of 10 mg/ml proteinase K were added to the tube. The solution was vortexed well and incubated at 37°C for 1 h. For the deproteinization of DNA, approximately one volume of TE-saturated phenol/chloroform mixture (1:1) was added to the tube. The tube was vortexed and centrifuged at RT for 2 min. The upper phase was transferred into another tube carefully avoiding denaturated proteins that are found at the aqueous/phenol interface. Extraction step was repeated. The upper phase from the last extraction step was transferred into another tube and one volume of chloroform was added. After vortexing, the tube was centrifuged at RT for 2 min. The upper phase was transferred into another tube and 0.1 volume sodium acetate (pH = 5.2) was added. After vortexing 0.6 volume of isopropanol was added. Then the tube was placed on ice for 15 min. After that the tube was centrifuged at 4 °C for 5 min. The upper phase was discarded and DNA pellet was washed with 1 ml of ice-cold %70 ethanol in order to remove salts. Then the tube was centrifuged at 4 °C for 5 min. The upper phase was discarded. DNA pellet was dried under N_2 and after drying 750 µl TE buffer was added into the tube. RNase-A was then added to a final concentration of 0.1 mg/ml and the tube was incubated at 37°C for 30 min in order to remove RNA. After that one volume of phenol was added to the tube. The tube was vortexed and centrifuged at RT for 2 min. After phenol extraction the upper phase was transferred into another tube. To remove phenol from the sample, the extraction step was repeated with one volume of chloroform two times. Then the upper phase was transferred into another tube and 0.1 volume sodium acetate (pH = 5.2) was added. After vortexing 0.6 volume of isopropanol was added. Then the tube was placed on ice for 15 min. After that the tube was centrifuged at 4 °C for 5 min. The supernatant was discarded and DNA pellet was washed with 1 ml of ice-cold %70 ethanol in order to remove salts. Then the tube was centrifuged at 4 °C for 5 min. The upper phase was discarded. DNA pellet was dried under N_2. DNA pellet was dissolved in 100 µl TE buffer at RT overnight. DNA was stored in

-20°C freezer. The concentration of the genomic DNA was determined spectrophotometrically. The calculation was based on the fact that 50 μg of double stranded DNA has an absorbance of 1.0 at $OD_{260\ nm}$. 260 nm/280 nm ratio of the absorbances were used for DNA quality estimations.

4.2. Analysis of Extracted DNA by Agarose Gel Electrophoresis

Based on the DNA ladder types used, 1-2 % agarose gels in 1X TAE buffer were used to visualize and analyze the extracted genomic DNA. 50 ml TAE buffer containing %1-2 agarose was melted in microwave oven, cooled down below 55 °C and ethidium bromide was added to a final concentration of 0.5 μg/ml under a hood. The solution was poured into the electrophoresis plate. A comb was placed onto the electrophoresis plate. Gel was polymerized at RT under a hood. Then gel was placed into an electrophoresis chamber containing 1X TAE buffer. The comb was removed. 1 μl 6X loading dye and 5 μl of samples were mixed and applied to the wells. A DNA ladder was applied to one of the wells also. The electrophoresis was at 150 Volt for 15-20 min. The gels were then analyzed under UV_{254nm} light and the images were recorded digitally.

4.3. Analysis of *Geobacillus anatolicus* Elongation Factor Gene (*tuf*)

4.3.1. Degenerate Primer Design for Unknown *G. anatolicus tuf* Gene

Because there were no nucleotide or amino acid sequence data for *G. anatolicus tuf* gene at the beginning of this study, to obtain partial sequence of *tuf* gene, two degenerate primers were designed based on nucleotide sequences of EF-Tu genes which come from various bacteria that are closely related to *G. anatolicus*. By using CLUSTALW multiple alignment technique, the sequences were aligned and two conserved regions were determined. Based on these conserved regions, two degenarate primers called TuF413 and TuR1174 were designed (See Table 3.2.).

4.3.2. Primer Design for Partially Known *G. anatolicus tuf* Gene

After sequencing of PCR product obtained with TuF413 and TuR1174, partial sequence of *tuf* gene was determined. Based on these data, within this region two new primers, TuF888 and TuR766 were designed (See Table 3.2.). These primers were used directly in sequencing reactions to confirm the obtained sequence data and reveal flanking regions of *tuf* gene.

4.3.3. Degenerate Primer Design for Adjacent Genes *fus* and *rplC*

To confirm sequence data obtained with, TuF413, TuF888, TuR766 and TuR1174, two primers were designed for adjacent genes *fus* and *rplC* (See Figure 1.7.). Because there were no nucleotide or amino acid sequence data for *G. anatolicus fus* and *rplC* genes at the beginning of this study, two degenerate primers were designed based on nucleotide sequences of *fus* and *rplC* genes which come from various bacteria that are closely related to *G. anatolicus*. By using CLUSTALW multiple alignment software, the sequences were aligned and one conserved region was determined for each gene. Based on these conserved regions, two degenarate primers called FusF and RplC were designed (See Table 3.2.).

4.3.4. PCR Amplifications of *tuf* and Adjacent Genes *fus* and *rplC*

Three PCR amplifications were performed on the *G. anatolicus* DNA. First of all, *tuf* gene was partially amplified using TuF413 and TuR1174 degenerate primers. PCR reactions were performed in 25 µl volume in PCR buffer containing 2.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 1 pmol of each primer, 50 ng of genomic DNA and 0.125 units of *Taq* DNA polymerase. PCR cycles were as follows:

94°C for 2 minutes 30 seconds
95°C for 1 minutes 30 seconds (denaturation)
50-60°C gradient for 1 minute 30 seconds (annealing) } X 30
72°C for 2 minutes 20 minutes (extension)
72°C for 7 minutes

Secondly, by using degenerate FusF and RplC primers, the fragment that contains complete *tuf* was amplified. PCR reactions were performed in 25 μl volume in PCR buffer containing 2.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 1 pmol of each primer, 50 ng of genomic DNA and 0.125 units of *Taq* DNA polymerase. PCR cycles were as follows:

94°C for 2 minutes 30 seconds
95°C for 1 minutes 30 seconds (denaturation)
50-60°C gradient for 1 minute 30 seconds (annealing) } X 35
72°C for 4 minutes (extension)
72°C for 7 minutes

Thirdly, in order to clone *tuf* gene, by using TuF2 and TuR2 primers, complete *tuf* gene was amplified. PCR reactions were performed in 25 μl volume in PCR buffer containing 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 1 pmol of each primer, 50 ng of genomic DNA and 0.125 units of *Taq* DNA polymerase. PCR cycles were as follows:

94°C for 2 minutes 30 seconds
95°C for 1 minutes 30 seconds (denaturation)
50-65°C gradient for 1 minute 30 seconds (annealing) } X 35
72°C for 2 minutes (extension)
72°C for 7 minutes

After amplification, 5 μl of each PCR product was mixed with 1 μl 6 X loading dye and electrophoretically analized on 1 % agarose gel.

4.3.5. Purification of PCR Products

Purification of PCR products was performed by using Wizard SV PCR Clean-Up system (Promega) as follows: An equal volume of membrane binding solution was added to the PCR reaction mix. After vortexing, the mixture was applied onto a SV minicolumn assembly (a minicolumn which was placed on a 2 ml collection tube) and incubated at RT for 1 min. Than the tube was centrifuged at 5000xg for 1 min, the flowthrough was discarded and the minicolumn was reinserted into a new 2 ml collection tube. To wash the

column, 700 µl membrane washing solution was added onto the column. After centrifugation for 1 min at 5000xg, the flowthrough was discarded and the column was inserted into a new 2 ml collection tube. The column was washed with 500 µl membrane washing solution and the tube was centrifuged for 5 min at 5000xg. The flowthrough was discarded and a new sterile 2 ml collection tube was placed under the minicolumn. After adding 50 µl nuclease-free water to the minicolumn, the column was left at room temperature for 1 min then centrifuged for 5 min at 5000xg to collect the purifed DNA. Purified PCR products were electrophoretically analized on 1 % agarose gel.

4.3.6. DNA sequencing

The fragment, which was amplified by using TuF413 and TuR1174 was sequenced with the same primers. The fragment, which was amplified by using FusF and RplC was sequenced with FusF, RplC, TuF888, TuR776. The six sequence data that comes from sequencing reactions was aligned and used to construct complete *tuf* sequence. DNA sequencing was carried out by cycle sequencing method using DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences, USA) using ABI 9700 Thermocycler (Applied Biosystems) at Acıbadem Hospital, Istanbul. Purified PCR products (or plasmids) were mixed with the sequencing mixture containing 5 pmol primer, 8 µl of sequencing reagent premix and water to a total reaction volume of 20 µl. The sequencing reaction was performed for both forward and reverse primers separately. The reaction conditions are 25 cycles of 95°C for 20 s, 50°C for 15 s and 60°C for 60 s. After cycle sequencing, the unbound dyes were removed by using DyeEx 2.0 Dye removal Kit (Qiagen). The purified products were analysed on the ABI 3100 Genetic Analyzer (Applied Biosystems). The sequencing results were analyzed by using CHROMAS software.

4.4. Cloning of *G. anatolicus tuf* Gene

4.4.1. Restriction Enzyme Digestions of Plasmid pCR® T7/NT-TOPO® and PCR Fragment Obtained Using TuF2 and TuR2 Primers

Restriction enzyme digestions were performed as a two-step procedure. Firstly, 20 µl of reaction mixture that contains Buffer D, acetylated BSA, 2 µl of either PCR fragment

(obtained using TuF2 and TuR2) or plasmid and 20u of NdeI restriction enzyme was incubated at 37 ^{o}C for 4 hours. Secondly, Buffer B, acetylated BSA and 40u of Csp45 I restriction enzyme were added into reaction mixture and reaction volume was adjusted to 40 µl. The reaction mixture was incubated at 37 ^{o}C for 4 hours again. As a control, 2µg of plasmid pCR® T7/NT-TOPO® was digested using the same procedure. 5 µl of samples were taken from reaction mixtures for analyzing in agarose gel electrophoresis to confirm digestion. pCR® T7/NT-TOPO® was named as pT7D3 after successful double-digestion.

4.4.2. Purification of the Digested PCR Fragment and Plasmid pT7D3

To obtain purified double digested PCR fragment, Wizard SV Gel and PCR Clean-Up System (Promega, USA) was used. Because DNA fragments more than approximately 100 bp aren't retained on the column in this system, PCR fragment could be seperated from the digested parts of the fragment (digested parts were 7-9 bp long).

To obtain purified plasmid, 40 µl reaction mixture was loaded on 1% agarose gel. After the run, the longer of the two bands appearing in the gel was cut out of the gel. This gel slice was applied to Wizard SV Gel and PCR Clean-Up System (Promega, USA) for the extraction of the digested DNA from gel.

4.4.3. Ligation of the Digested PCR fragment and Plasmid pT7D3

Content of 10 µl of ligation mixture to ligate the digested PCR fragment and pT7D3 was as follows: T4 DNA Ligase Buffer, 100ng of digested plasmid pT7D3, 1u of T4 DNA Ligase 41,2 ng of digested PCR fragments. The ligation reaction were carried out at 11ºC, overnight. Plasmid contaning the PCR fragment was named pT7D3Tu.

As a ligation control reaction, Csp45I-digested plasmid was used. At this same 30 µl of reaction volume, 7u T4 DNA Ligase and 700 ng plasmid were used. The rest of the procedure was same as in the ligation experiment mentioned above.

4.5. Transformation of the *Escherichia coli* JM109 (DE3) Cells

4.5.1. Preparation of *Escherichia coli* JM109 (DE3) Competent Cells

Escherichia coli JM109(DE3) cells were streak out on LB plates and incubated at 37 °C overnight. Then cells were subcultured in 20 ml liquid LB medium at 37 °C to OD_{600} = 0.1. Cells were grown until OD_{600} = 0.5 at 200 rpm; the cell growth was monitored spectrophotometrically by taking 1 ml samples and measuring the OD_{600} values. When the absorbance reached to OD_{600} = 0.5, cells were subcultured 1:100 in 20 ml LB liquid medium and grown until OD_{600} = 0.4 at 200 rpm. Than cells were collected by centrifugation using Beckman JA-14 rotor at 4 °C at 5000 rpm for 10 min. The supernatant was discarded and 2 ml of ice cold TSS buffer was added onto cell pellet. Cells were resuspended on ice gently stirring cells. Than cells were dispensed into ice cold eppendorf tubes in 100 µl aliquots in cold room. The tubes were stored in -80 °C freezer.

4.5.2. Transformation

JM109(DE3) competent cells were thawed on ice and 2 µl of ligation reaction containing 50 ng of pT7D3Tu, or (as a positive control) 2 µl of solution containing 50 ng of pCR®T7/NT-TOPO® plasmid or (as a negative control) 2 µl of water were added onto cells (See results section 5.2.4). The cells were incubated on ice for 30 min, then heat shocked at 42°C for 1 min 30 sec, then immediately transferred onto ice and kept on ice for 2 min. 1 ml of ice cold SOC buffer was added onto the cells. The cells were incubated at 37 ^{o}C for 50 min for recovery and for the expression of the plasmid encoded genes. For each transformation 1 ml of each sample were layered on LB-agar plates containing 150 µg/ml ampicilin, and incubated overnight at 37°C. Colonies that had been grown on the plates were re-streaked on new plates twice in order to obtain homogeneous populations of transformants.

4.6. Verification of Existence of *tuf* Gene in the Plasmid pT7D3Tu

4.6.1. Plasmid Purification

After obtaining homogeneous populations of transformants, a single colony streaked out on an agar LB-ampicillin plate and incubated overnight at 37°C. A tooth-pick of cells were taken from the plate and resuspended in ice cold 200 µl of TEL buffer by vortexing vigorously. Then 400 µl of NAS solution was added into the suspension. Cells were lyzed by gently inverting the microtube for about ten times. The mixture was stored on ice for 5 min. 300 µl of KAC solution was added gently mixed and the mixture was kept on ice for 5 min to precipate the chromosomal DNA. The mixture was centrifuged for 5 min at 20000xg. Approximately 750 µl of the supernatant was transferred into a new tube and the plasmid DNA was precipitated by adding two volume of ice-cold 100% ethanol. The mixture was stored on ice for 5 min and then centrifuged for 5 min at 20000xg. Supernatant was discarded and the plasmid DNA pellet was washed with ice cold 70 % ethanol. Then the solution was centrifuged again for 2 min at 20000xg. DNA pellet was air-dried for 15 min, dissolved in 50 µl TE buffer and stored at –20°C.

4.6.2. Confirmation of *tuf* Fragment in the Plasmid pT7D3Tu by Using PCR

To confirm that *tuf* gene is present in the pT7D3Tu, a PCR reaction was performed using two primers: T7 which complements to a region located on the plasmid and TuR1174 which complements to a region located on the *tuf* sequence. PCR reaction was performed in 25 µl volume in PCR buffer containing 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 1 pmol of each primer, 50 ng of plasmid DNA and 0.25 units of *Taq* DNA polymerase. PCR cycles were as follows:

94°C for 2 minutes 30 seconds
95°C for 1 minutes 30 seconds (denaturation) }
52°C for 1 minute 30 seconds (annealing) } X 30
72°C for 2 minutes 30 seconds (extension) }
72°C for 7 minutes

After amplification, 5 µl of each PCR product was mixed with 1 µl 6 X loading dye and electrophoretically analized on a 1 % agarose gel.

4.6.3. PCR Amplification of the Plasmid pT7D3Tu for DNA Sequencing

After confirmation of the presence of the *tuf* gene in the pT7D3Tu, a fragment that contains the complete *tuf* gene was amplified by using T7 and pRSET primer pair. PCR reaction was performed in 100 µl volume in PCR buffer containing 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 1 pmol of each primer, 50 ng of plasmid DNA and 1 units of *Taq* DNA polymerase. PCR cycles were as follows:

94°C for 2 minutes 30 seconds
95°C for 1 minutes 30 seconds (denaturation) ⎫
52°C for 1 minute 30 seconds (annealing) ⎬ X 30
72°C for 2 minutes 30 seconds (extension) ⎭
72°C for 7 minutes

After amplification, 5 µl of the PCR product was mixed with 1 µl 6 X loading dye and electropheretically analized on 1 % agarose gel. The rest of the PCR product was purified and prepared for sequencing.

4.6.4. Storage of JM109(DE3) Cells Containing pT7D3Tu.

The JM109(DE3) cells containing pT7D3Tu plasmid were grown overnight on agar LB plate containing ampicillin. A toothpick of cells were resuspunded in 1 ml liquid LB medium containing 50 % glycerol. Cells were resuspended until a homogeneous suspension was obtained and stored at -80°C freezer.

4.7. Expression of *G. Anatolicus tuf* Gene in JM109(DE3) Cells

JM109(DE3) cells containing pT7D3Tu carrying the *tuf* gene and original plasmid without *tuf* gene (as a control) were grown on agar plates overnight. 20 ml of LB liquid medium containing 100 µg/ml ampicilin were inoculated from these plates. Cells were

grown until OD_{600} : 0.6 at 37 °C at 200 rpm. At this point, IPTG was added to the flasks to the final concentration of 2 mM and from this point on 500 µl of samples were taken to eppendorf tubes at the beginning and after 2 and 5 hours. Each sample is processed immediately: samples were centrifuged at 20000xg, supernatants were discarded and the pellets were resuspended in 150 µl of protein lysis solution. Samples were placed on ice until use in the SDS-polyacrylamide gel electropheresis (SDS-PAGE) analysis.

To facilitate denaturation of proteins, samples were boiled at 90 °C for 5 min before applying to SDS-PAGE. 22,5 µl of each samples, or 15 µl of *E.coli* Tu or 15 µl of protein molecular weight marker were loaded onto 10% SDS-PAGE. Electrophoresis was performed at 10 mA overnight. After electrophoresis, the gel was stained with Coomassie Brilliant Blue staining solution for minimum 2 h then destained until the protein bands become visible.

4.8. Isolation of *G. anatolicus* EF-Tu

For the isolation of the recombinant protein, JM109(DE3) cells containing pT7D3Tu were grown in 800 ml of LB liquid medium containing 150 µg/ml ampicilin at 37°C. Cells were grown until OD_{600}: 0.6 at 200 rpm. At this point, IPTG was added to the final concentration of 2 mM. Cells were grown for additional 4 hours. The cells were harvested by centrifugation at 10000 rpm at 4°C for 15 min in Beckman JS-14 rotor.

Total 16 g of harvested cells were resuspended in 16 ml of ice cold Buffer A (homogenization buffer). Then equal volume of glass-beads were added to this suspension. Mixture was placed into the chamber of the bead-beater. Ice was placed around the chamber to prevent warming. Cells were disrupted in the Bead-beater for three times 30 second runs with minimum of 2 min waiting intervals in between to prevent warming. The cell lysate was centrifuged at 16000 rpm at 4°C for 30 min using Beckman JA-20 fixed-angle rotor. The supernatant was tranferred to a clean centrifuge tube and centrifuged at 45000 rpm at 4°C for 1 h 30 min. The supernatant was saved and then applied directly onto Ni^{+2}-affinity column.

4.9. Ni-Affinity Column Chromatography

4.9.1. Ni-Affinity Column Regeneration

Ni-NTA chromatography matrix (Pharmacia) was packed into a glass column (column dimensions: 0.9x6 cm). Column was washed with 50 ml of Buffer A containing 250 mM imidazole. Washing of the column was repeated with 50 ml of dH_2O, 50 ml of 100 mM $NiSO_4$, 50 ml of dH_2O and 50 ml of buffer A, respectively.

4.9.2. Purification Step

Before applying the samples to the column, the column was pre-equilibrated with 50 ml of equilibration buffer containing 10 mM imidazole. Then, the supernatant was applied to the column with a speed of 0.7 ml/min. To wash the unbound proteins, the column was washed with approximately 100 ml of buffer A containing 1 M NH_4Cl and 10 mM imidazole until OD_{260nm} reaches the base-line. His-tagged recombinant *Geobacillus anatolicus* elongation Factor Tu was eluted with a 100 ml of linear gradient from 25 mM to 100 mM imidazole. Fractions were collected in 7 ml of samples.

After purification step, from each fraction tube 20 µl of samples were taken, mixed with 5 µl 5X SDS-PAGE sample buffer and applied to the 10 % SDS-PAGE for analysis. Electrophoresis was run for 4 h at 30 mA.

4.10. Dialysis and Storage of *G. anatolicus* EF-Tu protein

Fractions containing pure *G.anatolicus* EF-Tu were pooled and placed on ice. 0,5 gr ammonium sulphate per ml of this sample was added slowly under continuous magnetic stirring. After 30 min, the sample was centrifuged at 16.000 rpm for 20 min using Beckman JA-20 fixed-angle rotor. Precipitated protein was dissolved in approximately 3 ml of 1x polymix buffer and dialyzed overnight at 4^oC against 1 lt 1x polymix buffer. Dialysis buffer was changed after 4 h with a fresh buffer and dialysis was continued overnight. After dialysis, to discard the particles, sample was centrifuged at 12.000 at 4 °C

for 5 min. Then, sample was divided into 50 and 100 µl aliquots in microcentrifuge tubes, freezed and stored at -80 °C until use.

4.11. Determination of Protein Concentration

Protein concentration was determined using Bradford assay (Bradford, 1976). Bovine Serum Albumin (BSA) was used as the protein standard. Using 1 mg/ml stock solution of BSA, several dilutions were prepared containing from 1 µg to 8 µg of BSA in 0.15 M NaCl in each tube. Final volume of dilutions were 100 µl. Apropriate dilutions of EF-Tu in 0.15M NaCl was also prepared. Then, 1 ml of Bradford's reagent was added into each tube. Tubes were vortexed and incubated at RT for 10 min. The absorbance of each sample at OD_{595nm} was measured. A plot of the absorbance at 595nm versus the volume of the protein sample was obtained after exactly 10 min. A plot of the absorbance at 595nm versus the known amount (from 1 µg to 8 µg) of the BSA protein was obtained. By using BSA plot as a standard, protein concentration of the EF-Tu was calculated.

4.12. Mobility-Shift Assay on Non-Denaturing PAGE for EF-Tu·EF-Ts complex formation

A 10% non-denaturing polyacrylamide gel (37.5:1) was prepared in 8 mM Tricine buffer at pH 8.2 and cooled in a cold room after polymerization . From 0 pmol to 400 pmol of *G. anatolicus* EF-Tu was prepared in 20 µl of the reaction mix in the presence of 300 pmol *G.anatolicus* EF-Ts. As a control experiment, same titrations were done for *G.anatolicus* EF-Ts and *E.coli* EF-Tu. Samples were prepared on ice and incubated at 37 °C for 10 min then transferred onto ice for at least 10 min before applying to the gel. 20 µl samples were mixed with 5 µl loading dye (phenol red in 50% glycerol) from which 20 µl were applied to 10% non-denaturing PAGE gel. Electrophoresis was done in a cold room at 4ºC in the same buffer. Before applying the samples, the gel was prerun for 1 hour at 250 V. Samples were then applied and electrophorized for 3 hrs at 250V. Protein bands were visualized by Commasie staining and destaining.

4.13. Mobility-Shift Assay on Non-Denaturing PAGE for Ternary Complex Formation between EF-Tu·GTP·Phe-$tRNA^{Phe}$

A non-denaturing 5% polyacrylamide gel (19:1) in ternary complex buffer which was given in Materials section (10 mM MOPS, 10 mM Mg-acetate, pH 6.5, containing freshly added 100 µM GTP and 1 mM DTE) was prepared. Gel plate was 0.5 mm thick and 20 cm long. After polymerization, gel was transferred to cold room at 4°C. The gel was prerun for 90 min at 40 mA (about 100 V) before applying the samples. 200 pmol of *G.anatolicus* EF-Tu or, as a control, 200 pmol of *E.coli* EF-Tu in 80 µl of polymix buffer containing 30 nmol of [^{14}C]-Phe, 0.1 µmol of ATP and 1 µmol of PEP, 0.1 µmol of GTP, 5 µg of PK, 0.3 µg of MK and 1 u of Phe-tRNA synthetase was mixed with varying amounts of $tRNA^{Phe}$ in 20 µl (from 50 pmol to 400 pmol) and incubated at 37 °C for 15 min . After incubation, samples were cooled on ice for at least 15 min. Then 20 µl of samples were mixed with 2 µl of sample buffer containing 50 % glycerol and traces of Bromophenolblue as a marker. Samples were applied onto pre-electrophoresed 5 % non-denaturing PAGE gel The electrophoresis was performed at 60 mA (about 150 V) for 3 hours. After electrophoresis, protein bands were visualized by Commasie staining and destaining. Duplicate samples were used to calculate the amount of the Phe-tRNA. For this purpose, after the incubations, 10% ice-cold TCA was added onto samples and kept on ice until filtering. Samples were filtered through glass fiber filters (Whatman, GF/C), washed with aproximately 5 ml of each ice-cold 10% TCA and isopropanol. Filters were dried at 110 °C for 15 min. 4 ml of scintillation cocktail was added onto the samples. The radioactivity retained on the filters were counted using a liquid scintillation counter.

4.14. Thermostability Test

At first, a binding mix, 90 µl for each sample, in polymix buffer containing 0.5 mM ^{3}H-GDP was prepared (specific activity 6.6 cpm/pmol). The binding mixes were stored on ice until use. Then 10 µl of samples containing 68 pmol/µl of EF-Tu in polymix buffer were incubated at between 20-90 °C in 10 °C intervals for 10 min. After incubation the samples were cooled on ice for at least 30 min. Then the samples were mixed with 90 µl of the binding mix and incubated at 37 °C for 15 min. After incubation the samples were placed on ice for 10 min. To each sample approximately 3 ml of ice-cold polymix buffer

was added and the samples were immediately filtered through nitrocellulose filters (Schleicher and Schüll, BA85). Nitrocellulose filters were washed twice, each time with 3 ml of polymix buffer and dried at 110 °C for 15 min. 4 ml of scintillation cocktail was added onto the samples. The radioactivity retained on the filters were counted using a liquid scintillation counter.

5. RESULTS

5.1. Determination of *Geobacillus anatolicus* Elongation Factor Tu Gene (*tuf*) Sequence

5.1.1. Extraction of DNA from *Geobacillus anatolicus*

DNA extraction (see section 4.1) from *Geobacillus anatolicus* appears as a single band when analyzed on agarolse gels (Figure 5.1).

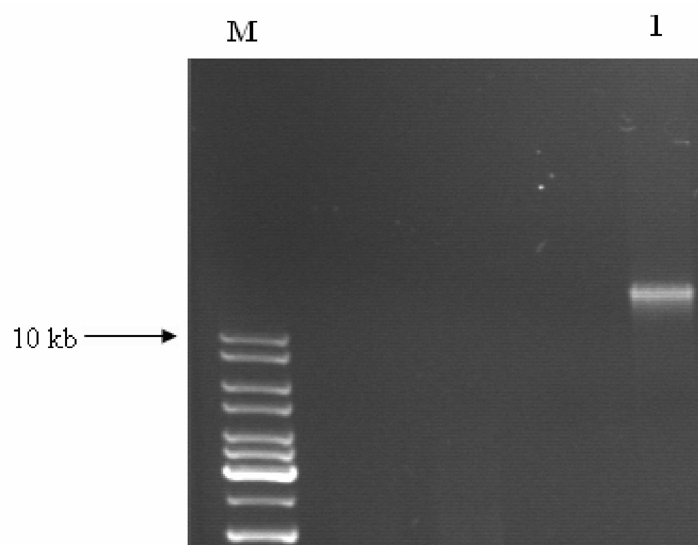

Figure 5.1. Genomic DNA of *Geobacillus anatolicus*. M, DNA Molecular marker; 1, *G. anatolicus* genomic DNA

5.1.2. Primer design

All primers that were used in this study were checked for hairpin structure, homodimer and heterodimer by using OligoAnalyzer software (Integrated DNA Technologies http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/). ClustalW multiple aligment results were shown in Table 5.1 and Figure 5.2. for *tuf* gene, in Table 5.2. and Figure 5.3. for *fus* gene and in Table 5.3. and Figure 5.4. for *rplC* gene respectively.

Table 5.1. Bacterial *tuf* genes used for degenerate primer design

Organism name	**GeneBank ID**
Geobacillus stearothermophilus	2661391
Bacillus licheniformis DSM 13	52346471
Oceanobacillus iheyensis HTE831	1015206
Streptococcus pneumoniae TIGR4	14972978
Clostridium acetobutylicum ATCC 824	15026203
Mycoplasma genitalium G-37	3845045
Escherichia coli CFT073	26110363
Aquifex pyrophilus	2687828
Thermus thermophilus HB27	46199632
Thermotoga maritima MSB8	4982068

```
              ....|....| ....|....| ....|....| ....|....| ....|....|
                      5         15         25         35         45
Geobacillu    AAGCATCCTA TACTTACCTA ACTATACAAA TTCATATTTT ACTTTTAAGG
Bacillus l    ---------- ---------- ---------- ---------- ----------
Oceanobaci    ---------- ---------- ---------- ---------- ----------
Streptococ    ---------- ---------- ---------- ---------- ----------
Clostridiu    ---------- ---------- ---------- ---------- ----------
Mycoplasma    ---------- ---------- ---------- ---------- ----------
Escherichi    ---------- -------GTG CTCTCTCCTG AAGGGGAGAG CACTATAGTA
Aquifex py    ---------- ---------- ---------- ---------- ----------
Thermus th    ---------- ---------- ---------- ---------- ----------
Thermotoga    ---------- ---------- ---------- ---------- ----------
Consensus     AAGCATCCTA TACTTACSTR MYYWYWCMWR WWSRKRWKWK MMYTWTARKR

              ....|....| ....|....| ....|....| ....|....| ....|....|
                     55         65         75         85         95
Geobacillu    AGGATCTTTC TCATGGCTAA AGCGAAATTT GAGCGCACGA AACCGCACGT
Bacillus l    ---------- --ATGGCTAA AGAAAAATTC GACCGTTCCA AATCACATGC
Oceanobaci    ---------- --ATGGCTAA AGAAAAATTC GATCGCTCCA AAAGTCACGT
Streptococ    ---------- --ATGGCAAA AGAAAAATAC GATCGTAGTA AACCACACGT
Clostridiu    ---------- --ATGGCAAA GGAAAAATTT GAAAGAACGA AACCACATGT
Mycoplasma    ---------- --ATGGCAAG AGAGAAATTT GACCGTTCCA AACCACATGT
Escherichi    AGGAATATAG CCGTGTCTAA AGAAAAATTT GAACGTACAA AACCGCACGT
Aquifex py    ---------- --ATGGCCAA GGAGAAGTTT GAGAGGACAA AAGAACACGT
Thermus th    ---------- --ATGGCGAA GGGCGAGTTT ATTCGGACGA AGCCTCACGT
Thermotoga    ---------- --ATGGCGAA GGAAAAATTT GTGAGAACAA AACCGCATGT
Consensus     AGGAWYWTWS YCRTGKCNAR RGVVRARTWY RWNMGNWSNA ARNVDCAYGY
```

Figure 5.2. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.1. Primer sites (TuF413 and TuR1174) were shown as in blue and consensus sequence was shown in red (Continued).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                   105        115        125        135        145
Geobacillu      CAACATTGGC ACGATCGGCC ACGTTGACCA TGGGAAAACG ACGCTGACAG
Bacillus l      CAACATTGGT ACAATTGGAC ACGTTGACCA TGGTAAAACT ACATTAACAG
Oceanobaci      TAACGTTGGT ACTTTAGGAC ACGTTGACCA TGGTAAAACT ACATTAACAG
Streptococ      TAACATTGGT ACTATCGGAC ACGTTGACCA CGGTAAAACT ACCCTAACTG
Clostridiu      AAACATAGGA ACAATAGGAC ACGTAGATCA CGGTAAAACA ACATTAACAG
Mycoplasma      CAATGTTGGT ACCATTGGTC ACATTGACCA TGGTAAAACC ACTTTAACAG
Escherichi      TAACGTTGGT ACTATCGGCC ACGTTGACCA CGGTAAAACT ACTCTGACCG
Aquifex py      AAACGTGGGT ACCATAGGTC ACGTTGACCA CGGAAAATCT ACGCTTACAT
Thermus th      GAACGTGGGG ACGATTGGGC ACGTGGACCA CGGGAAGACG ACGCTGACGG
Thermotoga      TAACGTTGGA ACGATTGGAC ATATCGACCA CGGAAAATCC ACACTGACAG
Consensus       NAAYRTDGGN ACNWTHGGNC AYRTNGAYCA YGGDAARWCN ACNYTDACNK

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   155        165        175        185        195
Geobacillu      CTGCGATCAC GACGGTTCTT ---GCGAAAC AAGGG----- ----AAAGCC
Bacillus l      CTGCTATCAC TACTGTGCTT CATAAGAAAT CTGGT----- ----AAAGGT
Oceanobaci      CTGCTATCAC TACTGTACTT ---GCAAAAC ACGGC----- ----GGTGGA
Streptococ      CAGCTATCAC AACTGTTTTG -GCACGTCGC TTGCCTTCA- -TCAGTTAAC
Clostridiu      CAGCAATAAC AACAATATTA ---GCAAAAG AAGGA----- ----AAAGCA
Mycoplasma      CTGCTATCTG TACAGTTTTA ---GCAAAGG AAGGA----- ----AAATCA
Escherichi      CTGCAATCAC CACCGTACTG ---GCTAAAA CCTAC----- ----GGCGGT
Aquifex py      CAGCTATAAC ATGCGTTCTC GCAGCAGGAC TCGTTGAAGG TGGTAAAGCT
Thermus th      CGGCGTTGAC GTA------- --TGTGGCGG CGGCGGAGAA TCCGAATGTA
Thermotoga      CCGCTATAAC AA-------- --AGTACCTT TCTCTCAAGG GACTT--GCC
Consensus       CNGCDWTVWS NWVNRTDYTN SVHRHDNVDN HNKNNBWVRR KNSDDDHDNH

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   205        215        225        235        245
Geobacillu      GAAGCAAAAG CGTACGACCA AATCGACGCA GCTCCGGAAG AGCGTGAACG
Bacillus l      ACTGCTATGG CATACGACCA AATCGACGGT GCTCCAGAAG AGCGCGAGCG
Oceanobaci      GAAGCACGTG CGTATGACCA AATTGATGGT GCGCCTGAAG AAAGAGAACG
Streptococ      CAACCTAAAG ACTATGCGTC TATCGATGCT GCTCCAGAAG AACGCGAACG
Clostridiu      AAAGCATTCA ATTACGAAGA GATTGATAAA GCACCAGAGG AAAAAGAAAG
Mycoplasma      GCTGCAACGC GTTATGATGA AATTGATAAA GCCCCTGAAG AAAAAGCAAG
Escherichi      GCTGCTCGTG CATTCGACCA GATCGATAAC GCGCCGGAAG AAAAAGCTCG
Aquifex py      AAGTGCTTCA AATACGAAGA AATTGACAAA GCTCCAGAAG AAAAAGAAAG
Thermus th      GAGGTTAAGG ACTACGGGGA GATTGACAAG GCGCCGGAGG AGCGTGCGCG
Thermotoga      CAGTATATTC CTTACGACCA GATCGACAAG GCCCCTGAAG AAAAGGCAAG
Consensus       VMDBNHHNNV VNTWYGVNBM DATYGAYRVN GCNCCDGARG ARMRNGMDMG
```

Figure 5.2. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.1. Primer sites (TuF413 and TuR1174) were shown as in blue and consensus sequence was shown in red (Continued).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      255        265        275        285        295
Geobacillu      CGGGATCACG ATTTCGACGG CCCACGTCGA GTATGAAACA GAGGCTCGTC
Bacillus l      CGGTATCACA ATTTCTACTG CACACGTTGA GTACGAAACT GACAACCGTC
Oceanobaci      TGGTATCACA ATCTCAACTG CACACGTAGA GTATGAAACT GAAACTCGTC
Streptococ      CGGTATCACT ATCAACACTG CGCACGTTGA GTACGAAACT GAAAAACGTC
Clostridiu      AGGAATAACA ATCAACACAG CACACGTTGA GTATGAAACA GAGAATAGAC
Mycoplasma      GGGAATCACA ATTAACTCTG CACACGTAGA ATATTCTTCT GACAAACGTC
Escherichi      TGGTATCACC ATCAACACTT CTCACGTTGA ATACGACACC CCGACCCGTC
Aquifex py      AGGTATAACC ATAAACATCA CTCACGTTGA GTACGAAACC GCTAAGAGAC
Thermus th      GGGGATTACG ATCAACACGG CGCACGTGGA GTACGAGACG GCGAAGCGGC
Thermotoga      AGGAATCACC ATCAACATCA CACACGTTGA GTATGAGACC GAAAAGAGAC
Consensus       NGGDATHACN ATHWMNWYND CNCACGTNGA RTAYKMNWCN SMNRMNMGDC


                ....|....| ....|....| ....|....| ....|....| ....|....|
                      305        315        325        335        345
Geobacillu      ACTACGCGCA CGTTGACTGC CCGGGCCACG CTGACTACGT CAAAAACATG
Bacillus l      ACTATGCACA CGTTGACTGC CCAGGACACG CTGACTATGT TAAAAACATG
Oceanobaci      ACTATGCACA CGTTGACTGC CCAGGTCACG CTGACTATGT TAAAAACATG
Streptococ      ACTACGCTCA CATCGACGCT CCAGGACACG CGGACTACGT TAAAAACATG
Clostridiu      ACTATGCTCA CGTTGACTGC CCAGGACATG CTGACTATGT AAAGAACATG
Mycoplasma      ACTATGCCCA TGTTGACTGT CCTGGACATG CTGACTACAT TAAAAATATG
Escherichi      ACTACGCGCA CGTAGACTGC CCGGGGCACG CCGACTATGT TAAAAACATG
Aquifex py      ACTACGCACA CGTTGACTGC CCCGGACACG CGGACTACAT CAAGAACATG
Thermus th      ACTATTCCCA CGTGGATTGC CCTGGGCACG CGGACTACAT CAAGAACATG
Thermotoga      ACTACGCTCA TATTGACTGT CCCGGTCACG CGGACTACAT CAAGAACATG
Consensus       ACTAYKCNCA YRTNGAYKSY CCNGGNCAYG CBGACTAYRT HAARAAYATG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                      355        365        375        385        395
Geobacillu      ATCACGGGCG CAGCGCAAAT GGACGGTGCA ATCCTTGTTG TATCGGCAGC
Bacillus l      ATCACTGGTG CTGCGCAAAT GGACGGTGCG ATCCTTGTTG TTTCTGCTGC
Oceanobaci      ATCACTGGTG CTGCACAAAT GGACGGAGCT ATCCTTGTAG TATCTGCTGC
Streptococ      ATCACTGGTG CTGCTCAAAT GGACGGAGCT ATCCTTGTAG TAGCTTCAAC
Clostridiu      ATTACAGGAG CAGCGCAAAT GGATGGAGCA ATCCTAGTAG TAAGTGCAGC
Mycoplasma      ATCACAGGTG CTGCACAAAT GGATGGAGCT ATTCTAGTTG TTTCAGCAAC
Escherichi      ATCACCGGTG CTGCTCAGAT GGACGGCGCG ATCCTGGTAG TTGCTGCGAC
Aquifex py      ATTACCGGTG CGGCCCAGAT GGACGGGGCT ATACTCGTGG TTTCCGCAGC
Thermus th      ATCACGGGTG CCGCGCAGAT GGACGGGGCG ATCCTTGTGG TGTCGGCGGC
Thermotoga      ATCACAGGAG CAGCTCAGAT GGACGGAGCC ATCCTTGTTG TTGCCGCAAC
Consensus       ATYACNGGHG CNGCNCARAT GGAYGGNGCN ATHCTNGTDG TDDSNKCDRC
```

Figure 5.2. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.1. Primer sites (TuF413 and TuR1174) were shown as in blue and consensus sequence was shown in red (Continued).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                   405        415        425        435        445
Geobacillu      TGACGGTCCG ATGCCGCAAA CGCGCGAACA CATTCTTCTC TCCCGCCAAG
Bacillus l      TGACGGCCCA ATGCCACAAA CTCGTGAGCA CATCCTTCTT TCTCGTAACG
Oceanobaci      TGATGGCCCA ATGCCACAAA CTCGTGAGCA CATCCTTCTA TCTCGTAACG
Streptococ      TGACGGACCA ATGCCACAAA CTCGTGAGCA CATCCTTCTT TCACGTCAGG
Clostridiu      AGATGGTCCA ATGCCACAAA CAAGAGAGCA CATACTATTG GCATCAAGAG
Mycoplasma      TGATAGTGTG ATGCCCCAAA CCCGCGAGCA CATCTTACTT GCCCGCCAAG
Escherichi      TGACGGCCCG ATGCCGCAGA CTCGTGAGCA CATCCTGCTG GGTCGTCAGG
Aquifex py      TGACGGCCCG ATGCCCCAAA CCAGAGAACA CGTCCTTCTC GCAAGACAGG
Thermus th      GGACGGGCCG ATGCCGCAGA CGCGGGAGCA CATTTTGCTG GCGCGGCAGG
Thermotoga      CGATGGTCCC ATGCCCCAGA CAAGAGAGCA CGTGCTTCTC GCAAGACAGG
Consensus       NGAYRGNSYV ATGCCVCARA CNMGNGARCA CRTNYTDYTN KSNHSNMRVG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   455        465        475        485        495
Geobacillu      TCGGCGTTCC GTACATCGTT GTTTTCTTGA ACAAATGCGA CATGGTGGAC
Bacillus l      TAGGTGTGCC TTACATCGTA GTATTCTTGA ACAAATGCGA CATGGTAGAC
Oceanobaci      TTGGGGTACC TGCTTTCGTA GTATTCCTTA ACAAAACAGA TATGGTTGAC
Streptococ      TTGGTGTTAA ACACCTTATC GTCTTCATGA ACAAAGTTGA CTTGGTTGAC
Clostridiu      TTGGTGTTGA ATATATAGTA GTATTCTTAA ATAAAGCAGA CCAAGTAGAT
Mycoplasma      TAGGGGTTCC TAAAATGGTA GTTTTTCTAA ACAAGTGTGA TATTGCTAGT
Escherichi      TAGGCGTTCC GTACATCATC GTGTTCCTGA ACAAATGCGA CATGGTTGAT
Aquifex py      TTAACGTTCC CTACATAGTC GTTTTCATGA ACAAGTGTGA TATGGTTGAC
Thermus th      TGGGGGTGCC GTACATTGTG GTGTTCATGA ACAAGGTGGA CATGGTGGAC
Thermotoga      TTGAGGTTCC CTACATGATC GTCTTCATAA ACAAGACAGA CATGGTTGAC
Consensus       TNRRBGTDVM NNMHHTNRTN GTNTTYHTDA AYAARDBNGA YHWDGYDRRY

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   505        515        525        535        545
Geobacillu      GACGAAGAAT TGCTTGAACT CGTTGAAATG GAAGTTCGCG ACCTTCTCTC
Bacillus l      GACGAAGAGC TTCTTGAGCT AGTTGAAATG GAAGTTCGCG ATCTTCTTTC
Oceanobaci      GATGAAGAGC TACTTGAATT AGTAGAAATG GAAGTTCGTG ACCTATTAAC
Streptococ      GACGAAGAAT TGCTTGAATT GGTTGAAATG GAAATCCGTG ACCTATTGTC
Clostridiu      GATCCAGAAT TAATCGACTT AGTAGAAATG GAAGTAAGAG AGTTATTAAA
Mycoplasma      GATGAAGAGG TACAAGAACT TGTTGCTGAA GAAGTACGTG ATCTGTTAAC
Escherichi      GACGAAGAGC TGCTGGAACT GGTTGAAATG GAAGTTCGTG AACTTCTGTC
Aquifex py      GACGAAGAGC TCCTTGAGCT CGTTGAGCTT GAAGTGAGAG AGCTTCTCAG
Thermus th      GACCCCGAGT TGCTGGACCT GGTGGAGATG GAGGTGCGGG ACCTTTTGAA
Thermotoga      GATCCTGAGC TCATCGACCT CGTCGAGATG GAAGTGAGAG ACCTTCTGAG
Consensus       GAYSMHGARB TNMWNGAVYT NGTNGMDVWD GARRTNMGNG ANYTDYTNWV
```

Figure 5.2. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.1. Primer sites (TuF413 and TuR1174) were shown as in blue and consensus sequence was shown in red (Continued).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                   555        565        575        585        595
Geobacillu      TGAATACGAC TTCCCGGGCG ATGAAGTGCC GGTTATCAAA GGTTCGGCAT
Bacillus l      TGAGTATGAG TTCCCTGGTG ACGATGTACC AGTTATCAAA GGTTCTGCTC
Oceanobaci      TGAGTACGAT TTCCCTGGTG ACGATCTACC AGTAATCAAA GGTTCTGCAC
Streptococ      AGAATACGAC TTCCCAGGTG ACGATCTTCC AGTTATCCAA GGTTCAGCAC
Clostridiu      CGAATATGGA TTCCCAGGCG ATGATACACC AATCGTAGTA GGAAGTGCGT
Mycoplasma      TTCCTATGGT TTTGATGGTA AGAACACTCC TATTATTTAT GGCTCAGCTT
Escherichi      TCAGTACGAC TTCCCGGGCG ACGACACTCC GATCGTTCGT GGTTCTGCTC
Aquifex py      CAAATACGAG TATCCCGGAG ACGAAGTTCC CGTAATAAGA GGTTCCGCTC
Thermus th      CCAGTACGAG TTTCCTGGGG ACGAGGTTCC GGTGATTCGG GGGAGTGCTC
Thermotoga      CCAGTACGGT TACCCTGGAG ACGAAGTGCC AGTCATAAGA GGTTCTGCTC
Consensus       HNMVTAYGRN TWYSMNGGNR ABRANVYDCC NRTNRTHNDD GGNWSNGCDY

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   605        615        625        635        645
Geobacillu      TAAAAGCGCT C--GAAGGCG ------ATCC GAAATGGGAA ----GAAAAA
Bacillus l      TTAAAGCTCT T--GAAGGAG ------ATGC TCAGTACGAA ----GAAAAA
Oceanobaci      TTAAAGCTCT A--GAAGGCG ------TTGC AGAATACGAA ----GAAAGA
Streptococ      TTAAAGCTCT T--GAAGGTG ------ACTC TAAATACGAA ----GACATC
Clostridiu      TAAAAGCATT A--CAGAATC C-----AGAT GATGCAGAAG C---AATAAA
Mycoplasma      TAAAAGCATT G--GAAGGTG ------ATCC AAAGTGGGAG ----GCTAAG
Escherichi      TGAAAGCGCT G--GAAGGCG ------ACGC AGAGTGGGAA ----GCGAAA
Aquifex py      TGGGAGCCCT TCAGGAGCTT GAACAAAACT CTCCCGGAAA GTGGGTAGGA
Thermus th      TTTTGGCGCT TGAGCAGATG CACAGGAACC CGAAGACGAG GCGTGGGGAG
Thermotoga      TGAAAGCCGT C--GAAGCTC CTAACGATCC GAATCACGAA GCTTACAA-A
Consensus       TDDDRGCNBT NSASVRRVHB SWMMVRWNNY NNHNBRSRAR SYKKRNNRDV

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   655        665        675        685        695
Geobacillu      ATCA------ --------TT ----GAATTG ATGAACGCGG TTGATGAGTA
Bacillus l      ATCT------ --------TT ----GAACTG ATGGCTGCTG TTGACGAGTA
Oceanobaci      ATTC------ --------TT ----GAGTTA ATGGCTGCAG TTGATGAGTA
Streptococ      GTTA------ --------TG ----GAATTG ATGAACACAG TTGATGAGTA
Clostridiu      ACCA------ --------AT AAAAGACTTA ATGGCAGAAG TAGATGCATA
Mycoplasma      ATCC------ --------AT ----GATTTG ATTAAAGCAG TTGATGAATG
Escherichi      ATCC------ --------TG ----GAACTG GCTGGCTTCC TGGATTCTTA
Aquifex py      AGC------- --------AT AAAGGAACTT CTGAACGCGA TGGACGAGTA
Thermus th      AACGAGTGGG TGGACAAGAT TTGGGAGCTG TTGGACGCGA TTGACGAGTA
Thermotoga      CCC------- --------AT CCAGGAGCTC CTCGACGCTA TGGATAACTA
Consensus       VNYNAGTGGG TGGACAAGWK HHRRGANYTN NYBRVHDHNV TDGAYDMNTR
```

Figure 5.2. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.1. Primer sites (TuF413 and TuR1174) were shown as in blue and consensus sequence was shown in red (Continued).

```
                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     705        715        725        735        745
Geobacillu       CATCCCGACT CCGCAACGTG AAGTTGACAA ACCGTTCATG ATGCCGATTG
Bacillus l       CATCCCAACT CCAGAGCGTG AAACTGACAA GCCATTCATG ATGCCTGTTG
Oceanobaci       TATCCCAACT CCAGAACGTG ACAAAGAAAA ACCATTCATG ATGCCAGTTG
Streptococ       TATCCCAGAA CCAGAACGTG ACACTGACAA ACCATTGCTT CTTCCAGTCG
Clostridiu       CATCCCAACA CCAGAAAGAC CAACAGATAA AGCATTCTTA ATGCCAATCG
Mycoplasma       GATTCCAACT CCTACACGTG AAGTAGATAA ACCTTTCTTA TTAGCAATTG
Escherichi       CATTCCGGAA CCAGAGCGTG CGATTGACAA GCCGTTCCTG CTGCCGATCG
Aquifex py       CATACCGACA CCTGAGAGAG AGGTTGACAA GCCCTTCCTG ATGCCCATAG
Thermus th       CATTCCCACG CCGGTGCGGG ACGTGGACAA GCCGTTCTTG ATGCCGGTGG
Thermotoga       CATTCCTGAT CCTCAGAGAG ACGTCGATAA GCCGTTCCTC ATGCCCATCG
Consensus        BATHCCNRMD CCDVHRMGDS MVRHNGAHAA RSCNTTSHTN HTDSCNRTNG

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     755        765        775        785        795
Geobacillu       AGGACGTTTT CTCGATCACA GGCCGTGGTA CGGTTGCGAC GGGCCGTGTT
Bacillus l       AGGACGTATT CTCAATCACA GGTCGTGGAA CAGTTGCTAC TGGACGCGTA
Oceanobaci       AGGATGTATT CTCAATCACT GGTCGTGGAA CAGTTGCAAC TGGACGTGTT
Streptococ       AGGACGTATT CTCAATCACT GGACGTGGTA CAGTTGCTTC AGGACGTATC
Clostridiu       AAGATGTCTT CACAATAACA GGAAGAGGAA CAGTTGCAAC AGGAAGAGTT
Mycoplasma       AAGATACGAT GACCATTACT GGTAGAGGTA CAGTTGTTAC AGGAAGAGTT
Escherichi       AAGACGTATT CTCCATCTCC GGTCGTGGTA CCGTTGTTAC CGGTCGTGTA
Aquifex py       AAGACGTGTT CAGCATATCG GGACGTGGAA CGGTTGTGAC CGGAAGAGTG
Thermus th       AGGACGTGTT TACGATCACG GGTCGTGGGA CGGTGGCCAC GGGTCGGATT
Thermotoga       AAGACGTGTT CTCCATCACA GGAAGAGGAA CGGTTGTTAC AGGAAGAATA
Consensus        ARGAYRYNWT BWSVATHWCN GGHMGWGGDA CVGTKGYNWC NGGHMGNRTN

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     805        815        825        835        845
Geobacillu       GAGCGTGGTA CGCTCAAAGT CGGTGACCCG GTTGAAATCA TCGGTCTTTC
Bacillus l       GAGCGCGGAC AAGTTAAAGT CGGTGACGAA GTTGAAATCA TCGGTCTTCA
Oceanobaci       GAGCGTGGAG AAGTTAAAGT TGGAGATGAA GTTGAAATCA TCGGTCTTGC
Streptococ       GACCGTGGTA TCGTTAAAGT CAACGACGAA ATCGAAATCG TTGGTATCAA
Clostridiu       GAAACTGGAA CATTAAAAGT TGGAGACGAA GTAGAAATCG TTGGAATGAA
Mycoplasma       GAAAGAGGTG AACTCAAAGT AGGTCAAGAA GTTGAAATTG TTGGTTTAAA
Escherichi       GAACGCGGTA TCATCAAAGT TGGTGAAGAA GTTGAAATCG TTGGTATC--
Aquifex py       GAAAGGGGCG TGTTGAGACC CGGAGATGAG GTAGAGATAG TAGGACTGAG
Thermus th       GAGCGGGGCA AGGTGAAGGT TGGGGACGAG GTGGAGATTG TGGGCCTTGC
Thermotoga       GAAAGAGGAA GAATCAGACC CGGTGATGAA GTTGAGATCA TAGGTCTCAG
Consensus        GAVMSNGGHV NVNTNARRSY HRRNSAHSMR RTNGARATHR TNGGHHTNNV
```

Figure 5.2. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.1. Primer sites (TuF413 and TuR1174) were shown as in blue and consensus sequence was shown in red (Continued).

```
             ....|....| ....|....| ....|....| ....|....| ....|....|
                   855        865        875        885        895
Geobacillu   TGACGAGCCG AAAGCCACGA CGGTTACGGG TGTTGAAATG TTCCGTAAAC
Bacillus l   AGAAGAAAAC AAGAAAACAA CTGTTACAGG TGTTGAAATG TTCCGCAAGC
Oceanobaci   TGAAGATGCA AGCAAAACAA CTGTAACTGG TGTTGAAATG TTCCGTAAGC
Streptococ   AGAAGAAACT CAAAAAGCAG TTGTTACTGG TGTTGAAATG TTCCGTAAAC
Clostridiu   AGATGAAATA ACAAAAGTAG TAGTAACAGG CGTAGAAATG TTCAGAAAAA
Mycoplasma   ACC---AATT AGAAAAGCAG TTGTTACTGG AATTGAAATG TTCAAAAAGG
Escherichi   -AAAGAGACT CAGAAGTCTA CCTGTACTGG CGTTGAAATG TTCCGCAAAC
Aquifex py   GGAAGAGCCC TTAAAGACTG TGGCAACGTC CATAGAGATG TTCAGGAAGG
Thermus th   TCCGGAGACG CGGAAGACGG TGGTGACGGG TGTGGAGATG CACCGGAAGA
Thermotoga   CTACGAGATC AAGAAGACCG TTGTGACGAG TGTGGAAATG TTCAGAAAGG
Consensus    NNMNGADVHN HNVRMVDYNR YNKBDACDDS HRTDGARATG YWCMRNAARV

             ....|....| ....|....| ....|....| ....|....| ....|....|
                   905        915        925        935        945
Geobacillu   TTCTTGACCA AGCAGAAGCT GGAGACAACA TCGGTGCGCT TCTCCGCGGT
Bacillus l   TTCTTGACTA TGCAGAAGCT GGAGACAACA TCGGTGCACT TCTTCGCGGT
Oceanobaci   TTCTAGATTA TGCTGAAGCT GGAGATAACA TTGGTGCACT TCTTCGTGGG
Streptococ   AACTTGACGA AGGTCTTGCT GGAGATAACG TAGGTGTCCT TCTTCGTGGT
Clostridiu   TACTTGATAG TGCATTAGCA GGAGATAACA TCGGAGCATT ATTAAGAGGA
Mycoplasma   AACTTGATTC AGCAATGGCT GGTGACAATG CTGGGGTATT ATTACGTGGT
Escherichi   TGCTGGACGA AGGCCGTGCT GGTGAGAACG TAGGTGTTCT GCTGCGTGGT
Aquifex py   TGCTTGATGA GGCGCTTCCC GGAGACAATA TTGGAGTTCT TCTGAGGGGA
Thermus th   CCCTGCAGGA GGGGATTGCT GGGGACAATG TGGGGGTGCT CCTGCGGGGT
Thermotoga   AACTCGATGA AGGAATCGCA GGAGACAACG TTGGATGTCT GCTCAGAGGA
Consensus    HNCTNSABNV DGSNNDNSCH GGDGABAAYR YNGGDKBNYT NYTNMGNGGD

             ....|....| ....|....| ....|....| ....|....| ....|....|
                   955        965        975        985        995
Geobacillu   GTATCGCGCG ACGAAGTTGA GCGTGGCCAA GTATTGGCAA AACCGGGCTC
Bacillus l   GTATCTCGTG AAGAAATCCA ACGCGGTCAA GTACTTGCTC AACCAGGTAC
Oceanobaci   GTTTCTCGTG AAGACATCAA CCGTGGTCAA GTATTAGCTA AACCAGGTTC
Streptococ   GTTCAACGTG ATGAAATCGA ACGTGGACAA GTTATCGCTA AACCAGGTTC
Clostridiu   GTACAGAGAG AAGACATCGA AAGAGGTCAG GTATTAGCAA AACCAGGTTC
Mycoplasma   GTTGAACGTA AAGAAGTTGA AAGAGGTCAA GTTTTAGCAA AACCAGGCTC
Escherichi   ATCAAACGTG AAGAAATCGA ACGTGGTCAG GTACTGGCTA AGCCGGGCAC
Aquifex py   GTAGGAAAGG ACGACGTAGA GAGGGGACAG GTGCTTGCCC AGCCCGGGAG
Thermus th   GTGAGCCGGG AGGAGGTGGA GCGGGGGCAG GTGCTGGCGA AGCCTGGGAG
Thermotoga   ATCGACAAGG ATGAAGTTGA AAGAGGACAG GTTCTCGCAG CTCCCGGAAG
Consensus    RTNNVNMRNR ANGAVRTNVA VMGNGGNCAR GTDHTNGCNV MDCCNGGNWS
```

Figure 5.2. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.1. Primer sites (TuF413 and TuR1174) were shown as in blue and consensus sequence was shown in red (Continued).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     1005       1015       1025       1035       1045
Geobacillu      AATCACGCCG CATACGAAAT TTAAAGCACA AGTTTACGTT CTGACGAAAG
Bacillus l      AATCACTCCG CACAAAAAAT TCAAAGCTGA AGTTTACGTT CTTTCTAAAG
Oceanobaci      AATCACTCCA CATACTAACT TCAAAGCTGA AGTTTATGTA TTATCAAAAG
Streptococ      AATCAACCCA CACACTAAAT TCAAAGGTGA AGTCTACATC CTTACTAAAG
Clostridiu      AATAACTCCA CATAATAAAT TCGTAGGTCA AGTATACGTA TTAAAGAAAG
Mycoplasma      TATTAAACCG CACAAGAAAT TTAAAGCTGA GATCTATGCT TTAAAGAAAG
Escherichi      CATCAAGCCG CACACCAAGT TCGAATCTGA AGTGTACATT CTGTCCAAAG
Aquifex py      CGTGAAGGCA CACAGAAAGT TCAGGGCACA GGTATACGTT CTTAGCAAGG
Thermus th      CATTACGCCG CACACGAAGT TTGAGGCCTC GGTGTACGTG TTGAAGAAGG
Thermotoga      CATCAAACCT CACAAGAGGT TCAAGGCTCA GATCTACGTT TTGAAGAAGG
Consensus       HRTNAMNSCD CAYAVNARVT TYRDRKSHBM RRTNTAYRYN YTDWVNAARG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                     1055       1065       1075       1085       1095
Geobacillu      AGGAAGGCGG ACGCCATACT CCGTTCTTCT CGAACTATCG TCCGCAATTC
Bacillus l      AAGAGGGTGG ACGTCACACT CCATTCTTCT CTAACTACCG CCCTCAGTTC
Oceanobaci      AAGAAGGTGG ACGTCATACT CCATTCTTCT CTAACTACCG TCCACAGTTC
Streptococ      AAGAAGGTGG ACGTCACACT CCATTCTTCA ACAACTACCG TCCACAATTC
Clostridiu      AAGAAGGCGG AAGACATACA CCATTCTTCA ATGGATATAG ACCACAATTC
Mycoplasma      AAGAAGGTGG TAGACACACT GGTTTTTTAA ACGGTTACCG TCCTCAATTC
Escherichi      ATGAAGGCGG TCGTCATACT CCGTTCTTCA AAGGCTACCG TCCGCAGTTC
Aquifex py      AGGAAGGAGG AAGGCACACG CCTTTCTTTG TGAACTACAG GCCACAGTTT
Thermus th      AGGAGGGTGG ACGGCACACG GGGTTTTTTT CGGGGTACCG TCCGCAGTTT
Thermotoga      AAGAGGGAGG AAGACATACA CCGTTCACAA AAGGCTACAA GCCTCAGTTC
Consensus       ADGARGGHGG WMGNCAYACD SSDTTYWYHD HNRRNTAYMR NCCDCARTTY

                ....|....| ....|....| ....|....| ....|....| ....|....|
                     1105       1115       1125       1135       1145
Geobacillu      TACTTCCGCA CAACGGACGT AACAGGC--- ATCATCACGC TTCCGGAAGG
Bacillus l      TACTTCCGTA CAACTGACGT AACTGGT--- ATCATTCAGC TTCCAGAAGG
Oceanobaci      TACTTCCGTA CTACGGACGT AACTGGT--- GTTATTGAAT TACCAGAAGG
Streptococ      TACTTCCGTA CTACTGACGT TACAGGT--- TCAATCGAAC TTCCAGCAGG
Clostridiu      TATTTCAGAA CAACAGACGT AACAGGA--- TCAATCCAGT TACCAGATGG
Mycoplasma      TATTTCCGTA CCACTGATGT AACTGGT--- TCTATTGCTT TAGCTGAAAA
Escherichi      TACTTCCGTA CTACTGACGT GACTGGT--- ACCATCGAAC TGCCGGAAGG
Aquifex py      TACTTCAGGA CCGCTGATGT AACGGGAACT GTAGTGAAGC TACCTGAAGG
Thermus th      TACTTTCGGA CGACGGACGT GACGGGG--- GTGGTGCAGT TGCCTCCGGG
Thermotoga      TACATAAGAA CCGCTGACGT TACAGGAGAA ATCGTAGGAC TTCCTGAAGG
Consensus       TAYWTHMGNA CNRCDGAYGT DACDGGNRMW DYNRTNVVDY TDSCDSMDRR
```

Figure 5.2. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.1. Primer sites (TuF413 and TuR1174) were shown as in blue and consensus sequence was shown in red (Continued).

```
               ....|....| ....|....| ....|....| ....|....| ....|....|
                   1155       1165       1175       1185       1195
Geobacillu     CGTTGAAATG GTCATGCCTG GCGACAACGT TGAGATGACG GTTGAGCTGA
Bacillus l     CGTAGAAATG GTTATGCCTG GAGACAACAT CGAAATGACT GTTGAATTGA
Oceanobaci     AACTGAAATG GTAATGCCTG GTGATAACAT TGAAATGACA GTAGAACTTA
Streptococ     TACTGAAATG GTAATGCCTG GTGATAACGT GACAATCGAC GTTGAGTTGA
Clostridiu     AGTAGAAATG GTAATGCCAG GAGACCACAT AGATATGACA GTTGAATTAA
Mycoplasma     TACTGAAATG GTTCTACCTG GTGATAATGC TTCTATTACT GTTGAGTTAA
Escherichi     CGTAGAGATG GTAATGCCGG GCGACAACAT CAAAATGGTT GTTACCCTGA
Aquifex py     CGTTGAGATG GTGATGCCCG GGGATAACGT TGAGCTTGAG GTAGAGCTTA
Thermus th     CGTGGAGATG GTGATGCCTG GGGACAACGT GACGTTTACG GTGGAGCTGA
Thermotoga     TGTCGAAATG GTCATGCCTG GAGACCACGT CGAAATGGAA ATAGAACTCA
Consensus      HRYNGARATG GTNMTRCCNG GNGAYMAYRY NDMDHTBRHN RTDRMVYTNA

               ....|....| ....|....| ....|....| ....|....| ....|....|
                   1205       1215       1225       1235       1245
Geobacillu     TCGCGCCGAT CGCGATCGAG GAAGGAACGA AATTCTCGAT CCGTGAAGGC
Bacillus l     TTTCAACAAT CGCGATCGAA GACGGAACTC GCTTCTCTAT CCGTGAAGGC
Oceanobaci     TTTCTCCAAT CGCGATTGAA GACGGAACTC GTTTTTCAAT CCGTGAAGGT
Streptococ     TTCACCCAAT CGCCGTAGAA CAAGGTACTA CATTCTCTAT CCGTGAGGGT
Clostridiu     TAACAAAAGT AGCAATGGGA GACAACTTAA GATTCGCTAT CAGAGAAGGC
Mycoplasma     TTGCTCCTAT CGCTTGTGAA AAAGGTAGTA AGTTCTCAAT TCGTGAAGGT
Escherichi     TCCACCCGAT CGCGATGGAC GACGGTCTGC GTTTCGCAAT CCGTGAAGGC
Aquifex py     TAGCGCCTGT AGCACTTGAG GAAGGACTGA GGTTTGCGAT AAGGGAAGGT
Thermus th     TCAAGCCGGT GGCGCTGGAG GAGGGTTTGC GGTTTGCCAT CCGTGAGGGT
Thermotoga     TCTACCCTGT CGCTATCGAA AAGGGACAGA GATTCGCTGT AAGGGAAGGC
Consensus      THNMNMMDRT VGCNNKNGRV VAVRRHHNDM VNTTYKCNRT HMGDGARGGY

               ....|....| ....|....| ....|....| ....|....| ....|....|
                   1255       1265       1275       1285       1295
Geobacillu     GGCCGTACAG TTGGCGCTGG TTCCGTATCG GAAATCATCG AGTAATCAAG
Bacillus l     GGACGTACAG TTGGTTCTGG CGTTGTTTCT TCAATCATCG AATAA-----
Oceanobaci     GGACGTACTG TAGGTTCTGG CGTTGTTTCT TCTATCCAAA AATAA-----
Streptococ     GGACGTACTG TTGGTTCAGG TATGGTTACA GAAATCGAAG CTTAA-----
Clostridiu     GGAAGAACAG TTGGATCAGG AGTTGTTACT AGTATAATAG AGTAA-----
Mycoplasma     GGTAGAACTG TAGGGGCAGG CACTGTAACA GAAGTTCTAG AATAG-----
Escherichi     GGCCGTACCG TTGGCGCGGG CGTTGTTGCT AAAGTTCTGG GCTAA-----
Aquifex py     GGAAGGACTG TAGGTGCTGG TGTCGTTACC AAAATCCTTG ACTGA-----
Thermus th     GGGCGGACCG TGGGCGCCGG CGTCGTCACC AAGATCCTGG AGTAA-----
Thermotoga     GGAAGAACAG TTGGAGCTGG TGTGGTTACA GAAGTCATCG AGTGA-----
Consensus      GGNMGDACHG TDGGNKCNGG HDYBGTHDCN DVDRTHVWNR VNTRRTCAAG
```

Figure 5.2. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.1. Primer sites (TuF413 and TuR1174) were shown as in blue and consensus sequence was shown in red (Continued).

```
               ....|....| ....|....| ....|....| ....|....| ....|..
                    1305       1315       1325       1335       1345
Geobacillu     AAAAAAGGAT GTCCAATCGT TGGACATCCT TTTTCTTTAT CCTGGCT
Bacillus l     ---------- ---------- ---------- ---------- -------
Oceanobaci     ---------- ---------- ---------- ---------- -------
Streptococ     ---------- ---------- ---------- ---------- -------
Clostridiu     ---------- ---------- ---------- ---------- -------
Mycoplasma     ---------- ---------- ---------- ---------- -------
Escherichi     ---------- ---------- ---------- ---------- -------
Aquifex py     ---------- ---------- ---------- ---------- -------
Thermus th     ---------- ---------- ---------- ---------- -------
Thermotoga     ---------- ---------- ---------- ---------- -------
Consensus      AAAAAAGGAT GTCCAATCGT TGGACATCCT TTTTCTTTAT CCTGGCT
```

Figure 5.2. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.1. Primer sites (TuF413 and TuR1174) were shown as in blue and consensus sequence was shown in red.

Table 5.2. Bacterial *fus* genes used for degenerate primer design.

Organism name	**GeneBank ID**
Bacillus anthracis str. Sterne	2851108
Bacillus subtilis subsp. subtilis str. 168	936826
B. thuringiensis serovar konkukian str. 97-27	2853707
Bacillus licheniformis ATCC 14580	3029537
Bacillus halodurans C-125	892222
Geobacillus kaustophilus HTA426	3184849
Oceanobacillus iheyensis HTE831	1015203
Aquifex aeolicus VF5	1192533

```
              ....|....| ....|....| ....|....| ....|....| ....|....|
                  5         15         25         35         45
Bacillus a    ATGGCAAGAG AGTTCTCTTT AGAAAACACT CGTAATATTG GTATCATGGC
Bacillus s    ATGGCAAGAG AGTTCTCCTT AGAAAAAACT CGTAATATCG GTATCATGGC
Bacillus t    ATGGCAAGAG AGTTCTCTTT AGAAAACACT CGTAATATTG GTATCATGGC
Bacillus l    ATGGCAAGAG AGTTCTCCTT AGAAAAAACT CGTAATATCG GAATCATGGC
Bacillus h    ATGGCAAGAG AGTTCTCCTT AGAAAATACG CGTAATATCG GGATCATGGC
Geobacillu    ATGGCAAGAG AGTTCTCCTT GGAAAAGACT CGCAACATAG GGATTATGGC
Oceanobaci    ATGGCTAGAG AGTTCTCCTT GGAAAAGACG CGTAATATTG GTATTATGGC
Aquifex ae    ATGGCGAGAG AGGTGCCTAT AGAGAAATTG AGAAACATAG GTATAGTTGC
Consensus     ATGGCDAGAG AGKTSYCYWT RGARAANWYK MGHAAYATHG GDATHRTKGC

              ....|....| ....|....| ....|....| ....|....| ....|....|
                  55         65         75         85         95
Bacillus a    TCACATCGAT GCTGGTAAAA CAACAGCTAC TGAGCGTATT CTGTACTACA
Bacillus s    TCACATCGAT GCCGGTAAAA CGACTACTAC AGAACGTATC TTGTTCTATA
Bacillus t    TCACATCGAT GCTGGTAAAA CAACAGCTAC TGAGCGTATT CTGTACTACA
Bacillus l    TCACATCGAT GCCGGTAAAA CGACAACAAC TGAACGTATC TTGTTCTACA
Bacillus h    TCACATCGAT GCCGGTAAGA CAACAACAAC AGAGCGGATT TTGTTCTACA
Geobacillu    GCACATTGAC GCCGGGAAAA CGACGACGAC TGAACGGATC TTGTTCTATA
Oceanobaci    CCACATTGAT GCGGGTAAAA CAACCACTAC TGAGCGTATT CTTTTCTATA
Aquifex ae    TCACATTGAC GCGGGTAAAA CTACGACTAC CGAGAGAATT CTCTATTACA
Consensus     BCACATYGAY GCBGGKAARA CDACNRCDAC HGARMGDATY YTBTWYTAYA

              ....|....| ....|....| ....|....| ....|....| ....|....|
                  105        115        125        135        145
Bacillus a    CAGGACGTAT TCACAAAATC GGTGAAACTC ACGAAGGTGC ATCTCAGATG
Bacillus s    CTGGTCGTAT CCACAAAATT GGTGAAACTC ATGAAGGAGC TTCACAAATG
Bacillus t    CAGGACGTAT TCACAAAATC GGTGAAACTC ACGAAGGTGC ATCTCAGATG
Bacillus l    CAGGCCGTAT CCATAAAATT GGTGAAACTC ACGAAGGGGC TTCCCAAATG
Bacillus h    CAGGACGTAT CCATAAAATT GGTGAGACTC ACGAAGGTGC TTCTCAAATG
Geobacillu    CCGGCCGCGT TCATAAAATC GGGGAAGTGC ATGAAGGTGC GGCAACGATG
Oceanobaci    CAGGACGCAT TCACAAAATT GGTGAAACTC ACGAAGGTGC TTCTCAAATG
Aquifex ae    CGGGTAAGAC TTACAAGATA GGTGAAGTTC ACGAAGGTGC TGCAACGATG
Consensus     CNGGHMRBRY YYAYAARATH GGKGARRYKC AYGAAGGDGC DKCHMMRATG

              ....|....| ....|....| ....|....| ....|....| ....|....|
                  155        165        175        185        195
Bacillus a    GACTGGATGG AGCAAGAGCA AGAGCGTGGT ATCACAATTA CTTCTGCTGC
Bacillus s    GACTGGATGG AGCAGGAACA AGAACGTGGT ATCACAATCA CTTCCGCTGC
Bacillus t    GACTGGATGG AGCAAGAGCA AGAGCGTGGT ATCACAATTA CTTCTGCTGC
Bacillus l    GACTGGATGG AGCAGGAACA AGAACGCGGT ATCACAATCA CATCTGCTGC
Bacillus h    GACTGGATGG AGCAGGAGCA AGAGCGTGGA ATCACGATCA CATCTGCTGC
Geobacillu    GACTGGATGG AACAAGAACA AGAGCGCGGA ATTACGATCA CGTCGGCGGC
Oceanobaci    GACTGGATGG AGCAAGAGCA AGAACGTGGT ATTACAATCA CTTCTGCTGC
Aquifex ae    GACTGGATGC CCCAGGAAAA GGAAAGAGGT ATAACCATAA CCGTTGCAAC
Consensus     GACTGGATGS MVCARGARMA RGARMGHGGW ATHACVATHA CNKYBGCDRC
```

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2. Primer site (FusF) was shown in blue and consensus sequence was shown in red (Continued).

```
                 ....|....| ....|....| ....|....| ....|....| ....|....|
                    205        215        225        235        245
Bacillus a    AACTACAGCA CAATGGA--- ---------A AGGTCACCGT GTAAACATCA
Bacillus s    GACTACTGCA CAGTGGA--- ---------A AGGATACCGC GTAAACATCA
Bacillus t    AACTACAGCA CAATGGA--- ---------A AGGTCACCGT GTAAACATCA
Bacillus l    AACAACTGCA CAGTGGA--- ---------A AGGCTACCGC GTAAACATCA
Bacillus h    AACAACAGCT CAATGGA--- ---------A GAACAATCGT ATCAACATCA
Geobacillu    GACAACGGCG CAATGGA--- ---------A AGGCCATCGC ATCAACATCA
Oceanobaci    GACAACAGCT GCTTGGA--- ---------G AGACCACCGT ATTAACATCA
Aquifex ae    GACCGCATGT TATTGGACGA GAAACGGGGA GAGGTATCAA ATAAACATAA
Consensus     RACHRCDKSD BMDTGGACGA GAAACGGGGR RRRNHAYCRH RTHAACATMA

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                    255        265        275        285        295
Bacillus a    TTGACACTCC AGGTCACGTA GATTTCACAG TAGAAGTAGA ACGTTCTTTA
Bacillus s    TCGATACTCC AGGACACGTA GACTTCACAG TTGAAGTTGA ACGTTCTCTT
Bacillus t    TTGACACTCC AGGTCACGTA GATTTCACAG TAGAAGTAGA ACGTTCTTTA
Bacillus l    TCGACACACC AGGACACGTA GACTTCACAG TTGAGGTTGA ACGTTCCCTG
Bacillus h    TCGATACTCC TGGACACGTA GACTTTACCG TTGAGGTTGA ACGTTCTTTA
Geobacillu    TCGACACGCC GGGGCACGTC GACTTTACGG TCGAGGTTGA GCGTTCGTTG
Oceanobaci    TCGATACTCC GGGACACGTA GACTTCACTG TAGAAGTTGA GCGTTCCCTG
Aquifex ae    TTGACACACC CGGACACGTT GACTTCTCCG TTGAAGTTGT ACGTTCCATG
Consensus     TYGAYACDCC NGGDCACGTH GAYTTYWCNG THGARGTWGW RCGTTCBHTD

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                    305        315        325        335        345
Bacillus a    CGCGTACTTG ATGGCGCAGT AGCAGTACTT GATGCACAAT CTGGTGTAGA
Bacillus s    CGTGTACTTG ATGGTGCTGT TGCTGTACTT GACGCGCAAT CAGGCGTTGA
Bacillus t    CGCGTACTTG ATGGCGCAGT AGCAGTACTT GATGCACAAT CTGGTGTAGA
Bacillus l    CGCGTACTTG ATGGTGCCGT TGCCGTTCTT GACGCACAAT CAGGCGTTGA
Bacillus h    CGTGTATTAG ATGGAGCTGT TGCGGTACTT GATGCTCAAT CAGGTGTTGA
Geobacillu    CGCGTATTGG ACGGGGCCAT CACAGTCTTA GACGCGCAAT CGGGTGTTGA
Oceanobaci    CGTGTACTTG ATGGTGCGGT GACTGTACTA GATGCACAAT CTGGGGTGGA
Aquifex ae    AAAGTTCTCG ACGGAATAGT TTTCATATTC TCCGCGGTTG AAGGTGTGCA
Consensus     MRHGTWYTNG AYGGNRYNRT NDYNRTHYTH KMYGCDSWWK MDGGBGTDSA

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                    355        365        375        385        395
Bacillus a    ACCACAAACA GAAACTGTTT GGCGTCAGGC TACTACTTAC GGCGTACCTC
Bacillus s    GCCTCAAACA GAAACTGTTT GGCGTCAGGC AACAACTTAT GGAGTACCGC
Bacillus t    ACCACAAACA GAAACTGTTT GGCGTCAGGC TACTACTTAC GGCGTACCTC
Bacillus l    GCCGCAAACT GAAACAGTTT GGCGTCAGGC GACAACTTAC GGAGTACCTC
Bacillus h    GCCGCAAACA GAGACAGTTT GGCGTCAAGC AACAACTTAC GGTGTTCCTC
Geobacillu    ACCGCAAACG GAAACCGTTT GGCGCCAAGC GACGACATAT GGCGTCCCAC
Oceanobaci    GCCACAAACT GAAACAGTAT GGCGCCAAGC AACAACTTAT GGTGTTCCAC
Aquifex ae    ACCTCAGTCC GAAGCAAACT GGAGATGGGC GGACAGGTTC CAAGTTCCGA
Consensus     RCCDCARWCN GARRCHRWHT GGMGHYRRGC DRMNASDTWY SRHGTHCCDM
```

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2. Primer site (FusF) was shown in blue and consensus sequence was shown in red (Continued).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                   405        415        425        435        445
Bacillus a     GTATCGTATT CGTTAACAAA ATGGATAAAA TCGGTGCAGA TTTCT-TATA
Bacillus s     GTATCGTTTT TGTTAACAAA ATGGACAAAA TCGGTGCAGA CTTCC-TTTA
Bacillus t     GTATCGTATT CGTTAACAAA ATGGATAAAA TCGGTGCAGA TTTCT-TATA
Bacillus l     GTATCGTATT TGTCAACAAG ATGGACAAAA CGGGTGCGGA CTTCC-TTTA
Bacillus h     GTGTCGTGTT CGTTAACAAA ATGGACAAAA CAGGAGCAGA CTTCT-TATA
Geobacillu     GGATCGTGTT CGTCAACAAA ATGGATAAAA TCGGCGCCGA TTTCT-TGTA
Oceanobaci     GTATCGTATT CATCAATAAA ATGGATAAAA CAGGTGCGGA CTTCT-TGTA
Aquifex ae     GGATAGCCTT CATAAACAAG ATGGACCGTC TGGGTGCGGA TTTTTACAGA
Consensus      GKRTMGYNTT YRTHAAYAAR ATGGAYMRWM YVGGHGCVGA YTTYYAYDKA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   455        465        475        485        495
Bacillus a     CTCTGTAGGA ACAATCCACG ATCGTTTACA AGCAAACGCA CACCCAATTC
Bacillus s     CTCTGTAGGA ACTTTAAGAG ACCGTCTTCA AGCAAACGCT CATGCAATTC
Bacillus t     CTCTGTAGGA ACAATCCACG ATCGTTTACA AGCAAACGCA CACCCAATTC
Bacillus l     CTCTGTAGGA ACTCTTAGAG ACCGTCTTGA GGCGAACGCT CACGCGATCC
Bacillus h     TTCCGTAAGC ACTCTTCATG ACCGCCTTCA AGCAAACGCA CATCCGATCC
Geobacillu     TTCGGTGAAA ACGCTCCATG ACCGCCTGCA AGCGAACGCC CATCCGGTGC
Oceanobaci     TTCTACAGGT ACATTAAAAG AGCGCCTAGG TGCTAATGCA CATCCGATTC
Aquifex ae     GTGTTTAAGG A-AATAGAAG AAAAGCTAAC CATAAAGCCC GTTGCCATTC
Consensus      BTSBDYRRRN ACDHTHVRHG ANMRBYTDVV NRYDAABSCH SWYSCVRTBC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   505        515        525        535        545
Bacillus a     AGTTACCAAT CGGTGCTGAA GATGAGTTCA ATGGTATCAT TGACCTTGTT
Bacillus s     AATTGCCGAT CGGTGCTGAA GATAACTTCG AAGGTATCAT TGACCTTGTA
Bacillus t     AGTTACCAAT CGGTGCTGAA GATGAGTTCA ATGGTATCAT TGACCTTGTT
Bacillus l     AATTACCGAT TGGCGCCGAA GACAATTTTG AAGGGATCAT CGACCTTGTA
Bacillus h     AACTACCAAT TGGTGCGGAA GACAATTTTG AAGGAATTAT CGACTTAGTA
Geobacillu     AACTGCCGAT CGGCGCTGAA GATCAATTCT CCGGCATTAT CGACTTGGTT
Oceanobaci     AGATGCCAAT TGGTGCTGAA GATGAGTTTC ACGGAATCAT TGACTTAATC
Aquifex ae     AAATACCCCT GGGAGCGGAG GACCAGTTTG AAGGTGTTAT AGATCTAATG
Consensus      ARHTRCCVMT BGGHGCBGAR GAYVANTTYN MHGGNRTYAT HGAYYTDRTN

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   555        565        575        585        595
Bacillus a     GAAGAATGTG CTT---ACAT GTACGGTAAC GATTTAGGAA CAGACATTCA
Bacillus s     GAAAACGTTG CGT---ATTT CTACGAAGAT GACCTCGGAA CTCGTTCTGA
Bacillus t     GAAGAATGTG CTT---ACAT GTACGGTAAC GATTTAGGAA CAGACATTCA
Bacillus l     GAAAACGTTG CGT---ACTA CTATGAAGAT GACCTTGGAA CACGCTCTGA
Bacillus h     GATATGGTCG CTT---ATTT CTATGAAGAT GATCTCGGAA CAAGAACAGA
Geobacillu     GAAATGTGCG CGT---ACCA TTACCATGAT GAGCTTGGCA AAAACATTGA
Oceanobaci     ACAATGGATG CTT---ACTT CTATTTAGAT GATTTAGGTC AACGTTCCGA
Aquifex ae     GAAATGAAGG CAATAAGGTG GCTCGAAGAA ACCCTCGGAG CTAAATACGA
Consensus      RMWRWVDDBG CDWTAARBHD BYWYBDWRAH RMBYTHGGHV MWVRHWHHSA
```

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2. Primer site (FusF) was shown in blue and consensus sequence was shown in red (Continued).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      605        615        625        635        645
Bacillus a      ACGTGTTGAA ATTCCTGAAG AGCATAAAGA ACTAGCTGAA GAATACCGTG
Bacillus s      TGCAAAAGAA ATCCCTGAAG AGTACAAAGA GCAAGCTGAA GAGCTTCGCA
Bacillus t      ACGTGTTGAA ATTCCTGAAG AGCATAAAGA ACTAGCTGAA GAATACCGTG
Bacillus l      AGCGCGTGAA ATCCCTGCTG AGTACAAAGA CAAAGCTGAA GAGCTTCGCG
Bacillus h      AGCGAAGGAA ATTCCTGACG AGTATAAAGA GCAAGCGCAA GAGTATCATG
Geobacillu      GCGCATTGAC ATTCCGGAAG AATACCGCGA TATGGCAGAA GAGTACCATA
Oceanobaci      AACACGTGAA ATTCCTGCTG AGTACAAAGA CGAAGCAGAA GAGTTACGTG
Aquifex ae      AGTAGTAGAC ATTCCTCCAG AATACCAGGA AAAGGCTCAA GAATGGCGCG
Consensus       DVBNVDDGAM ATYCCKSMHG ARYAYMRVGA NVWRGCDSAA GARYDNCRYR

                ....|....| ....|....| ....|....| ....|....| ....|....|
                      655        665        675        685        695
Bacillus a      GAAAACTTAT TGAAGCGGTA GCTGAGCTTG ATGAAGAAAT GATGATGAAG
Bacillus s      ACAGCCTTAT TGAAGCTGTA TGTGAGCTTG ATGAAGAGCT TATGGATAAG
Bacillus t      GAAAACTTAT TGAAGCGGTA GCTGAGCTTG ATGAAGAAAT GATGATGAAG
Bacillus l      CTAGCTTGAT TGAAGCTGTT GCTGAGCTTG ATGAAGAGCT TATGATGAAG
Bacillus h      AGAAACTAGT GGAAGCTGCT GCTGAACTAG ATGAAGAACT CATGATGAAA
Geobacillu      ACAAGCTGAT TGAGGCGGTC GCTGAGCTGG ACGAAGAATT AATGATGAAA
Oceanobaci      CAAGCTTAAT CGAAGCTGTT GCAGAAACTG ACGAAGAACT TATGATGAAA
Aquifex ae      AAAAGATGAT AGAAACCATC GTAGAAACCG ACGACGAGTT AATGGAAAAG
Consensus       VNARVHTDRT NGARRCBRYH KBWGARMYNG AYGAMGARHT NATGRWDAAR

                ....|....| ....|....| ....|....| ....|....| ....|....|
                      705        715        725        735        745
Bacillus a      TACCTAGAAG GTGAAGAAAT CACTGTAGAA GAGCTTAAAG CTGGTATCCG
Bacillus s      TACCTTGAGG GTGAAGAACT TACAATTGAC GAGTTAAAAG CAGGAATTCG
Bacillus t      TACCTAGAAG GTGAAGAAAT CACTGTAGAA GAGCTTAAAG CTGGTATCCG
Bacillus l      TACTTGGAAG GTGAAGAAAT CACAGTTGAC GAGCTGAAAG CAGCAATCCG
Bacillus h      TACTTAGAAG GAGAAGAACT TACAAAGGAC GAATTAAAAG CTGCGATCCG
Geobacillu      TATTTGGAAG GGGAAGAAAT CACGACGGAA GAGCTGAAGG CCGCGATCCG
Oceanobaci      TATCTTGAAG GAGAAGAAAT CTCCAACGAT GAATTGAAAA CTGCTATTCG
Aquifex ae      TACTTAGAAG GACAGGAAAT ATCTATAGAT GAACTAAGAA AAGCTTTAAG
Consensus       TAYYTDGARG GDSARGAAMT HWCNRHNGAH GARYTDARRR MHGSDWTHMG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                      755        765        775        785        795
Bacillus a      TAAGGCTACA ACTTCTGTAG AATTCTTCCC AGTAATCTGT GGTTCTGCAT
Bacillus s      TAAAGGTACA TTAAATGTTG AATTCTACCC TGTTCTTGTT GGTTCTGCCT
Bacillus t      TAAGGCTACA ACTTCTGTAG AATTCTTCCC AGTAATCTGT GGTTCTGCAT
Bacillus l      CAAAGGAACT TGTAACGTTG AGTTCTATCC GGTTCTATGT GGTTCAGCTT
Bacillus h      TAAAGGAACG TGTAATGTAG AATTCTATCC AGTCCTTTGT GGTTCGGCAT
Geobacillu      CAAAGCGACG ATCAGCGTTG AATTCTTCCC GGTCTTCTGC GGTTCGGCAT
Oceanobaci      TCAAGCAACT CTTAATGTTG ATTTTTATCC AGTACTTTGC GGTTCAGCAT
Aquifex ae      AAAGGCAACA ATAGAGAGAA AGCTCGTTCC CGTTCTTTGC GGTTCTGCAT
Consensus       HMARGSDACD HBHDVBRKWR ADYTYKWYCC NGTHHTHKKY GGTTCDGCHT
```

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2. Primer site (FusF) was shown in blue and consensus sequence was shown in red (Continued).

```
               ....|....| ....|....| ....|....| ....|....| ....|....|
                     805        815        825        835        845
Bacillus a     TCAAAAACAA AGGTGTTCAA ATTCTGTTAG ACGCAGTTAT CGACTACCTA
Bacillus s     TCAAAAACAA AGGTGTTCAA CTAGTACTTG ATGCTGTTCT TGACTACCTT
Bacillus t     TCAAAAACAA AGGTGTTCAA ATTCTGTTAG ACGCAGTTAT CGACTACCTA
Bacillus l     TCAAAAACAA AGGTGTTCAA TTAGTGCTTG ACGCAGTTCT TGACTACCTG
Bacillus h     TTAAGAACAA AGGTGTTCAG CTTATGCTAG ACGCTGTTCT TGATTACCTC
Geobacillu     TCAAAAACAA AGGTGTTCAG CTGCTTCTTG ACGGCGTTGT CGACTACTTG
Oceanobaci     TTAAGAACAA AGGTGTTCAG TTAATGCTAG ACGCAGTACT TGATTATCTA
Aquifex ae     TCAAGAACAA AGGTGTTCAA CCCCTTCTTG ACGCAGTTAT AGATTACCTG
Consensus      TYAARAACAA AGGTGTTCAR HYNVTDYTWG AYGSHGTWVT HGAYTAYYTN

               ....|....| ....|....| ....|....| ....|....| ....|....|
                     855        865        875        885        895
Bacillus a     CCATCTCCAT TAGACGTACC TGCTATTAAA GGTATCGTTC CTGATACAGA
Bacillus s     CCTGCACCAA CTGACGTTGC TGCAATAAAA GGTACACGTC CTGATACAAA
Bacillus t     CCATCTCCAT TAGACGTACC TGCTATTAAA GGTATCGTTC CTGATACAGA
Bacillus l     CCAGCTCCAA CTGACGTACC AGCAATCAAA GGTACTCTGC CTGATTCTGA
Bacillus h     CCTTCTCCAC TTGATGTACC TGCAATTAAA GGACATGTTC CAGATACAGA
Geobacillu     CCGTCTCCGG TTGACATCCC GGCGATTCGC GGCGTCGTTC CGGATACCGA
Oceanobaci     CCTGCACCAA CAGATGTACC TCCAATCGAA GGTATCATCC CTGGTACAGA
Aquifex ae     CCTTCTCCTA TAGACCTTCC TCCCGTTAAG GGGACAAATC CCAAGACCGG
Consensus      CCDKCWCCDN YWGAYVTHSC DSCNRTHVRV GGNVHHVDBC CNRRKWCHRR

               ....|....| ....|....| ....|....| ....|....| ....|....|
                     905        915        925        935        945
Bacillus a     TGAAGAAGTA GAACGTAAGT CTAGCGATGA AGAACCATTC GCAGCTCTAG
Bacillus s     TGAAGAGATT GAGCGTCATT CTTCTGACGA AGAGCCATTC TCTGCATTAG
Bacillus t     TGAAGAAGTA GAACGTAAGT CTAGCGATGA AGAACCATTC GCAGCTCTAG
Bacillus l     TGAAGAGGTT ACTCGTGAGT CTTCTGACGA TGCACCTTTC TCAGCGCTTG
Bacillus h     AGAAGAAGCT GTTCGTAAAC CTGGTGACGA TCAGCCGTTT GCGGCGCTTG
Geobacillu     AGAAGAAGTG ACACGTGAAG CAAGCGATGA CGCTCCGTTT GCCGCTTTGG
Oceanobaci     GGAAAAAGTA ACTCGTCCAT CTTCTGATGA TGAGCCTTTC TCTGCATTAG
Aquifex ae     GGAAGAAGAG GTCAGACACC CCTCTGACGA CGAACCCTTC TGCGCTTACG
Consensus      DGAARARRHD RHNMGWVMNB CHDSYGAYGA HSMDCCNTTY KSNGCDYWNG

               ....|....| ....|....| ....|....| ....|....| ....|....|
                     955        965        975        985        995
Bacillus a     CATTCAAAAT CATGACTGAC CCTTATGTTG GTAAGTTAAC ATTCTTCCGT
Bacillus s     CATTTAAAGT TATGACTGAC CCTTACGTTG GTAAGTTGAC GTTCTTCCGT
Bacillus t     CATTCAAAAT CATGACTGAC CCTTATGTTG GTAAGTTAAC ATTCTTCCGT
Bacillus l     CATTTAAAGT TATGACTGAC CCTTACGTTG GTAAATTGAC GTTCTTCCGT
Bacillus h     CGTTTAAGGT TATGACAGAC CCTTATGTTG GTAAGCTTAC GTTCTTCCGT
Geobacillu     CATTTAAAAT CATGACTGAC CCGTATGTCG GGAAATTGAC GTTTATTCGC
Oceanobaci     CTTTCAAAGT AATGACAGAC CCTTATGTAG GTAAACTTAC ATTCTTCCGT
Aquifex ae     CCTTTAAGGT TATGTCCGAC CCGTATGCCG GACAACTTAC CTACATCAGA
Consensus      CNTTYAARRT HATGWCHGAC CCKTAYGYHG GDMARYTDAC VTWYWTYMGH
```

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2. Primer site (FusF) was shown in blue and consensus sequence was shown in red (Continued).

```
               ....|....| ....|....| ....|....| ....|....| ....|....|
                    1005       1015       1025       1035       1045
Bacillus a     GTGTACTCTG GTGTGTTAAA CTCTGGATCA TACGTGAAAA ACTCAACTAA
Bacillus s     GTGTACTCTG GAACACTTGA TTCTGGTTCA TACGTGAAAA ACTCTACTAA
Bacillus t     GTGTACTCTG GTGTGTTAAA CTCTGGATCA TACGTGAAAA ACTCAACTAA
Bacillus l     GTGTACTCAG GAACACTTGA TTCCGGTTCA TACGTAAGAA ACTCAACAAA
Bacillus h     GTATATTCTG GAACGCTAGA TTCAGGATCA TACGTTAAAA ACTCGACAAA
Geobacillu     GTCTACTCTG GAACGCTTGA TTCCGGTTCA TATGTGATGA ACACGACGAA
Oceanobaci     GTGTACTCTG GAACATTAGA TTCTGGTTCA TACGTGAAAA ACTCAGTAAA
Aquifex ae     GTGTTCTCAG GAACGCTAAA AGCGGGTTCT TACGTCTACA ACGCAACCAA
Consensus      GTVTWYTCWG GWRYRYTWRA HKCNGGWTCW TAYGTNWDVA ACDCDRYNAA

               ....|....| ....|....| ....|....| ....|....| ....|....|
                    1055       1065       1075       1085       1095
Bacillus a     AGGTAAGCGT GAGCGTGTAG GTCGTATCCT ACAAATGCAC GCTAACAGCC
Bacillus s     AGGCAAACGT GAGCGTGTTG GACGTATTCT TCAAATGCAC GCTAACAGCC
Bacillus t     AGGTAAGCGT GAGCGTGTAG GTCGTATCCT ACAAATGCAC GCTAACAGCC
Bacillus l     AGGCAAACGT GAGCGCGTCG GCCGTATCCT GCAAATGCAC GCCAACAGCC
Bacillus h     AGATAAGCGT GAGCGCGTCG GTCGTATCCT ACAAATGCAC GCGAACCACC
Geobacillu     AGGGAAGCGC GAACGGATCG GCCGCTTGCT GCAAATGCAC GCGAACCACC
Oceanobaci     AGATAAGCGT GAGCGTGTAG GACGTATTCT ACAAATGCAC GCTAACTCTC
Aquifex ae     GGACGAAAAG CAAAGGGCTG GAAGACTTCT TCTCATGCAC GCGAACTCCA
Consensus      RGRBRARMRB SARMGBRYHG GHMGHHTBCT DCWMATGCAC GCBAACHVYM

               ....|....| ....|....| ....|....| ....|....| ....|....|
                    1105       1115       1125       1135       1145
Bacillus a     GTGAAGAGAT CTCAACAGTT TACGCTGGTG ATATCGCTGC TGCTGTAGGT
Bacillus s     GTGAAGAAAT CTCTACTGTA TACGCAGGGG ATATCGCGGC TGCTGTAGGT
Bacillus t     GTGAAGAGAT CTCAACAGTT TACGCTGGTG ATATCGCTGC TGCTGTAGGT
Bacillus l     GTGAAGAAAT CTCTACTGTG TACGCAGGGG ATATCGCTGC TGCTGTAGGT
Bacillus h     GTGAGGAAAT TTCGACAGTT TACTCTGGTG ACATCGCAGC AGCTGTAGGT
Geobacillu     GTCAAGAAAT TTCGAAAGTC TATGCCGGTG ATATTGCCGC AGCGGTAGGT
Oceanobaci     GAGAAGAAAT CTCTAGAGCA TATTCAGGTG AAATTGCTGC AGCAGTTGGT
Aquifex ae     GAGAGGAAAT ACAGCAGGTT TCCGCGGGTG AAATTTGTGC AGTTGTAGGA
Consensus      GWSARGARAT HYMDMVDGYN TMYKCNGGKG AHATYKSNGC WGYDGTWGGW

               ....|....| ....|....| ....|....| ....|....| ....|....|
                    1155       1165       1175       1185       1195
Bacillus a     TTAAAAGATA CTACTACTGG TGATACTCTT TGTGACGAGA AGAGCCTTGT
Bacillus s     CTTAAAGATA CAACTACTGG AGATACTCTA TGTGATGAGA AGGATCTTGT
Bacillus t     TTAAAAGATA CTACTACTGG TGATACTCTT TGTGACGAGA AGAGCCTTGT
Bacillus l     TTGAAAGACA CGACTACAGG GGACACGCTT TGCGACGAGA AAAACCTTGT
Bacillus h     TTAAAAGATA CGTCTACAGG GGATACACTT TGTGACGAAA AGAATCTTGT
Geobacillu     TTAAAAGATA CGACGACCGG TGATACTCTA TGTGATGAGA AGCATCCTGT
Oceanobaci     TTGAAAGATA CTTCTACAGG GGATACTCTT TGTGATGAGA AGAACCTTGT
Aquifex ae     CT---AGACG CCGCAACGGG TGATACTCTC TGTGATGAAA AGCACCCCAT
Consensus      YTDAAAGAYR CNDCDACNGG DGAYACDCTH TGYGAYGARA ARVRYCYYRT
```

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2. Primer site (FusF) was shown in blue and consensus sequence was shown in red (Continued).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                   1205       1215       1225       1235       1245
Bacillus a      TATCCTTGAG TCTATGGAAT TCCCAGAGCC AGTTATCTCT GTAGCTATCG
Bacillus s      TATCCTTGAG TCAATGGAAT TCCCAGAACC AGTTATCGAC GTTGCTATTG
Bacillus t      TATCCTTGAG TCTATGGAAT TCCCAGAGCC AGTTATCTCT GTAGCTATCG
Bacillus l      TATCCTGGAA TCAATGGAAT TCCCTGAGCC TGTTATCCAC GTTGCAATCG
Bacillus h      TATTCTTGAG TCTATGGAGT TCCCAGAGCC TGTTATCCAC TTGTCCGTTG
Geobacillu      CATTTTAGAG TCGATGCAAT TCCCAGAACC GGTCATTTCG GTGGCAATCG
Oceanobaci      TATCTTAGAG TCTATGGACT TCCCAGAACC AGTTATCTCT GTTGCGATTG
Aquifex ae      AATCCTTGAA AAGCTTGAAT TCCCTGACCC CGTTATATCT ATGGCTATAG
Consensus       HATYYTDGAR WMDMTKSAVT TCCCWGAVCC NGTYATHBMB DTDKCNRTHG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   1255       1265       1275       1285       1295
Bacillus a      AACCAAAATC AAAAGCTGAC CAAGATAAAA TGGGTACAGC ATTATCTAAG
Bacillus s      AGCCTAAATC AAAAGCTGAC CAAGATAAAA TGGGTATCGC TTTAGCGAAA
Bacillus t      AACCAAAATC AAAAGCTGAC CAAGATAAAA TGGGTACAGC ATTATCTAAG
Bacillus l      AGCCTAAATC TAAGGCTGAC CAAGACAAAA TGTCTACTGC ACTGGCTAAA
Bacillus h      AGCCTAAATC AAAAGCGGAC CAAGATAAAA TGGGCTTAGC GCTTGCGAAG
Geobacillu      AGCCGAAGTC GAAAGCCGAT CAAGACAAAA TGAGCCAAGC GCTGCAAAAA
Oceanobaci      AACCAGAAAC TAAAGCTGAC CAAGATAAAA TGGCAATTGC TTTATCTAAG
Aquifex ae      AGCCAAAGAC CAAGAAGGAC CAAGAAAAAC TCTCACAAGT TCTCAACAAG
Consensus       ARCCDRARWC NAARRMBGAY CAAGAHAAAM TSDSHHHHGY DYTNNMNAAR

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   1305       1315       1325       1335       1345
Bacillus a      CTTTCTGAAG AAGATCCAAC ATTCCGTGCT CACACTGACC AAGAAACTGG
Bacillus s      CTTGCTGAAG AAGATCCTAC ATTCCGTACT CAAACAAACC CAGAAACTGG
Bacillus t      CTTTCTGAAG AAGATCCAAC ATTCCGTGCT CACACTGACC AAGAAACTGG
Bacillus l      TTGGCTGAAG AGGATCCAAC ATTCCGTGCG CATACTGACC CTGAAACAGG
Bacillus h      CTTGCTGAAG AGGATCCAAC ATTCAAAACA CACACAGATG AAGAAACTGG
Geobacillu      TTGCAAGAAG AAGACCCGAC TTTCCGCGCT CACACCGACC CGGAAACGGG
Oceanobaci      TTAGCTGAAG AGGATCCAAC TTTCCGTACT GAAACTAACC CAGAAACTGG
Aquifex ae      TTCATGAAAG AGGATCCAAC CTTCAGGGCA ACAACCGATC CCGAAACTGG
Consensus       YTNNHDRAAG ARGAYCCDAC HTTCMRNRCD VMHACHRAYS MNGAAACDGG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   1355       1365       1375       1385       1395
Bacillus a      CCAAACAATC ATCGCTGGTA TGGGTGAACT TCACCTTGAT ATCATCGTTG
Bacillus s      CCAAACGATC ATCTCTGGTA TGGGTGAACT TCACCTTGAT ATCATTGTTG
Bacillus t      CCAAACAATC ATCGCTGGTA TGGGTGAACT TCACCTTGAT ATCATCGTTG
Bacillus l      TCAAACGATC ATCGGCGGTA TGGGTGAGCT TCACCTTGAT ATCATCGTTG
Bacillus h      GCAAACGATC ATTGCTGGAA TGGGTGAGCT TCACCTAGAC ATTATCGTTG
Geobacillu      ACAAACGATC ATCTCCGGAA TGGGCGAGTT GCATCTTGAC ATCATCGTCG
Oceanobaci      TCAAACAATT ATCTCAGGTA TGGGTGAACT TCACCTTGAT ATCATTGTTG
Aquifex ae      TCAGATACTC ATACACGGAA TGGGTGAGCT CCACCTCGAA ATAATGGTTG
Consensus       NCARAYRMTY ATHBVHGGWA TGGGYGARYT BCAYCTHGAH ATHATBGTYG
```

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2. Primer site (FusF) was shown in blue and consensus sequence was shown in red (Continued).

```
                   ....|....| ....|....| ....|....| ....|....| ....|....|
                      1405       1415       1425       1435       1445
Bacillus a     ACCGTATGCG CCGTGAATTC AAAGTTGAAG CAAACGTTGG TGCTCCTCAG
Bacillus s     ACCGTATGAA ACGCGAATTT AAAGTAGAAG CCAATGTTGG TGCTCCTCAA
Bacillus t     ACCGTATGCG CCGTGAATTC AAAGTTGAAG CAAACGTTGG TGCTCCTCAG
Bacillus l     ACCGTATGAA ACGCGAATTT AAAGTTGAAG CCAACGTCGG TGCTCCTCAA
Bacillus h     ACCGTCTTCG CCGTGAGTTT AAAGTGGAAG CCAATGTTGG TGCACCTCAA
Geobacillu     ACCGGATGCG CCGCGAGTTC AAAGTCGAAG CGAACGTCGG GGCGCCGCAA
Oceanobaci     ACCGTTTGAA GCGTGAGTTT AAAGTAGGAG CTCAAGTAGG GGCACCTCAA
Aquifex ae     ACAGAATGAA GAGGGAATAC GGAATTGAAG TGAACGTCGG TAAACCGCAG
Consensus      ACMGDHTKMR VMGBGARTWY RRARTNGRAG YNMAHGTHGG KRMDCCKCAR

                   ....|....| ....|....| ....|....| ....|....| ....|....|
                      1455       1465       1475       1485       1495
Bacillus a     GTAGCATACC GTGAGACTTT CCGCGCTGCT GCGAAAGTTG AAGGTAAGTT
Bacillus s     GTTGCGTACC GTGAAACATT CCGTACCGGT GCAAAAGTTG AAGGTAAATT
Bacillus t     GTAGCATACC GTGAGACTTT CCGCGCTGCT GCGAAAGTTG AAGGTAAGTT
Bacillus l     GTTGCGTATC GTGAAACATT CCGTGCGAGC GCGCAAGTTG AAGGTAAATT
Bacillus h     GTTTCTTACC GTGAAACGAT TCGTCAAGCA GCTCAAGTCG AAGGTAAGTT
Geobacillu     GTTGCCTACC GCGAAACGTT CCGCAAATCG GCGCAAGTCG AAGGCAAATT
Oceanobaci     GTTTCTTATC GTGAAACATT CCGTGGTTCT GCTGAAGTCG AAGGTAAATT
Aquifex ae     GTTGCTTACA AGGAAACCAT CAGGAAAAAG GCAATTGGTG AGGGTAAGTT
Consensus      GTWKCNTAYM RBGARACNWT YMGBVVNDVN GCDVWWGKYG ARGGYAARTT

                   ....|....| ....|....| ....|....| ....|....| ....|....|
                      1505       1515       1525       1535       1545
Bacillus a     CGCTCGTCAA TCTGGTGGAC GTGGACAATT CGGTCACGTT TGGATTGAGT
Bacillus s     CGTACGTCAG TCTGGTGGAC GCGGTCAGTT CGGACACGTT TGGATCGAAT
Bacillus t     CGCTCGTCAA TCTGGTGGAC GTGGACAATT CGGTCACGTT TGGATTGAGT
Bacillus l     CGTACGTCAG TCTGGTGGAC GCGGACAGTT CGGACACGTT TGGATCGAAT
Bacillus h     CGTACGTCAA TCTGGTGGTC GTGGTCAATA CGGTCACGTG TGGATTGAAT
Geobacillu     CATCCGTCAA TCCGGCGGCC GCGGTCAATA CGGCCACGTT TGGATCGAAT
Oceanobaci     CGTACGTCAG TCTGGTGGAC GCGGACAATT CGGTCACGTT TGGGTTAAAT
Aquifex ae     CATCAAGCAA ACTGGTGGTA GAGGGCAGTA CGGTCACGCG ATAATCGAAA
Consensus      CRYHMRKCAR WCYGGYGGHM GHGGDCARTW CGGHCACGYK WKRRTYRARW

                   ....|....| ....|....| ....|....| ....|....| ....|....|
                      1555       1565       1575       1585       1595
Bacillus a     TTGAACCTAA TGAAGAAGGT AAAGGATTCG AATTCGAAAA CAAGATCGTC
Bacillus s     TCGAACCAAA CGAAGAAGGC GCTGGCTTTG AATTTGAAAA TGCGATTGTC
Bacillus t     TTGAACCTAA TGAAGAAGGT AAAGGATTCG AATTCGAAAA CAAGATCGTC
Bacillus l     TCTCTCCAAA CGAAGAAGGA AAAGGCTTTG AGTTCGAAAA CGCAATCGTC
Bacillus h     TCTCTCCTAA CGAAGAAGGT GCCGGTTTCG AATTTGTGAA TGGTATCGTC
Geobacillu     TCTCGCCGAA CGAGCGCGGC AAAGGCTTCG AATTCGAGAA CGCCATCGTC
Oceanobaci     TCGAACCAAA CGAAGAAGGC GCTGGATTTG AATTTGAAAA CAAAATTGTT
Aquifex ae     TCGAACCCCT CCCCAGAGGT GCGGGATTTG AATTCATAGA CGACATTCAC
Consensus      TYKMDCCNMW YSMVVRMGGH RMNGGHTTYG ARTTYRWRRA YRVNATYSWY
```

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2. Primer site (FusF) was shown in blue and consensus sequence was shown in red (Continued).

```
              ....|....| ....|....| ....|....| ....|....| ....|....|
                   1605       1615       1625       1635       1645
Bacillus a    GGTGGTGTTG TTCCACGTGA ATACATCCCA GCTGTAGGCG CAGGTCTTGA
Bacillus s    GGTGGGGTCG TTCCTCGTGA ATACATCCCA GCTGTTCAAG CAGGTCTGGA
Bacillus t    GGTGGTGTTG TTCCACGTGA ATACATCCCA GCTGTAGGCG CAGGTCTTGA
Bacillus l    GGTGGTGTCG TTCCTCGTGA ATACATTCCG GCTGTTCAAG CAGGACTTGA
Bacillus h    GGAGGGGTAG TACCAAGAGA ATACATCCCA TCTGTTCAAG CTGGTCTTGA
Geobacillu    GGTGGGGTTG TTCCGAAAGA GTATGTGCCG GCCGTCCAAG CTGGATTGGA
Oceanobaci    GGTGGTGTAG TTCCTCGTGA ATACATTCCG GCAGTTCAAC AAGGTATCCA
Aquifex ae    GGAGGAGTTA TCCCCAAAGA ATTCATACCC TCCGTTGAGA AGGGTGTAAA
Consensus     GGWGGDGTHR THCCNMRWGA RTWYRTNCCV KCHGTHSRVV MDGGWNTNVA

              ....|....| ....|....| ....|....| ....|....| ....|....|
                   1655       1665       1675       1685       1695
Bacillus a    AGATGCACTT AAAAATGGTG TACTTGCTGG ATATCCAGTA GTAGACATTA
Bacillus s    AGATGCACTC GAAAATGGTG TACTAGCAGG CTTCCCGCTT ATCGACATTA
Bacillus t    AGATGCACTT AAAAATGGTG TACTTGCTGG ATATCCAGTA GTAGACATTA
Bacillus l    AGATGCACTT CAAAACGGTG TAGTCGCAGG TTACCCTGTT ATCGACATCA
Bacillus h    AGAAGCGCTT GAAAATGGTC TTCTTGCAGG TTACCCAGTG ATTGACATTA
Geobacillu    AGAAGCGATG CAAAACGGCG TCTTAGCCGG CTATCCGGTT GTCGACATTA
Oceanobaci    AGAGTCTATG GAAAATGGGG TTCTTGCAGG ATATCCGTTG ATCGATATCA
Aquifex ae    GGAAGCTATG CAAAACGGAA TTCTCGCAGG ATACCCCGTT GTTGACGTTA
Consensus     RGADKCDMTB VAAAAYGGNV THBTHGCHGG HTWYCCNBTD RTHGAYRTYA

              ....|....| ....|....| ....|....| ....|....| ....|....|
                   1705       1715       1725       1735       1745
Bacillus a    AAGCTGCATT AGTTGACGGA TCTTACCATG ATGTCGATTC ATCTGAGATG
Bacillus s    AAGCTAAACT TTTCGACGGT TCTTACCATG ATGTTGACTC AAACGAAATG
Bacillus t    AAGCTGCATT AGTTGACGGA TCTTACCATG ATGTCGATTC ATCTGAGATG
Bacillus l    AAGCGAAACT TTTTGACGGT TCTTACCACG ATGTCGACTC TAACGAAATG
Bacillus h    AGGCGAAGCT TTTTGACGGT TCTTACCATG ATGTCGACTC AAGTGAGATG
Geobacillu    AAGCGAAACT GTTTGACGGA TCGTACCACG ATGTCGACTC GAGCGAAATG
Oceanobaci    AAGCAACACT TTATGATGGT AGCTACCATG ATGTCGACTC AAACGAGATG
Aquifex ae    GAGTTAGACT CTTTGACGGT TCTTACCACG AAGTTGACTC TTCGGACATA
Consensus     RRGYDRVRYT NKWYGAYGGW WSBTACCAYG AWGTYGAYTC DWVBGAVATR

              ....|....| ....|....| ....|....| ....|....| ....|....|
                   1755       1765       1775       1785       1795
Bacillus a    GCGTTCAAAA TCGCTGCATC TATGGCACTT AAAGCTGCGG TTTCTAAATG
Bacillus s    GCATTTAAAG TCGCTGCTTC TATGGCATTG AAAAATGCAG TCAGCAAATG
Bacillus t    GCGTTCAAAA TCGCTGCATC TATGGCACTT AAAGCTGCGG TTTCTAAATG
Bacillus l    GCGTTTAAAA TCGCTGCTTC CATGGCGCTT AAAAACGCGG CAAGCAAGTG
Bacillus h    GCGTTTAAGA TTGCTGCTTC TATGGCACTT AAGAACGCAA AATCAAAATG
Geobacillu    GCATTCAAAA TCGCTGCTTC CTTGGCATTG AAAAACGCCG CGACGAAGTG
Oceanobaci    GCATTTAAAG TTGCAGCATC TATGGCGCTA AAAGCTGCTA AAAACAAATG
Aquifex ae    GCATTCCAGG TTGCGGGTTC CTTGGCATTC AAAGATGCAG CCAAAAAGGC
Consensus     GCRTTYMARR TYGCDGSWTC YWTGGCRYTN AARRMYGCNR HNWVNAARKS
```

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2. Primer site (FusF) was shown in blue and consensus sequence was shown in red (Continued).

```
               ....|....| ....|....| ....|....| ....|....| ....|....|
                    1805       1815       1825       1835       1845
Bacillus a     TAACCCAGTA ATTCTTGAGC CAATGATGAA AGTTGAAGTT GTAATTCCTG
Bacillus s     TAACCCAGTT CTTCTTGAAC CAATCATGAA AGTGGAAGTG GTTATTCCAG
Bacillus t     TAACCCAGTA ATTCTTGAGC CAATGATGAA AGTTGAAGTT GTAATTCCTG
Bacillus l     TAACCCAGTT ATTCTTGAGC CGATCTCCAA AGTTGAAGTC GTTATCCCAG
Bacillus h     TAACCCAGTT CTTCTAGAAC CAATGATGAA AGTGGAAGTC GTTGTACCTG
Geobacillu     CGATCCGGTT CTTTTGGAAC CGATCATGAA AGTCGAAGTC GTCATCCCTG
Oceanobaci     TAAACCAGTT CTTCTTGAAC CAATGATGAG AGTTGAAGTT GTTGTTCCTG
Aquifex ae     AGATCCCGTT CTTCTGGAAC CCATAATGGA AGTTGAAGTG GAAACTCCCG
Consensus      HRAHCCVGTW MTTYTDGARC CVATVWYSRR AGTBGAAGTB GWHRYHCCHG

               ....|....| ....|....| ....|....| ....|....| ....|....|
                    1855       1865       1875       1885       1895
Bacillus a     AAGAGTACAT GGGTGACATT ATGGGTGACG TAACATCTCG TCGTGGACGT
Bacillus s     AAGAATATAT GGGAGATATC ATGGGTGATA TCACTTCACG TCGCGGACGC
Bacillus t     AAGAGTACAT GGGTGACATT ATGGGTGACG TAACATCTCG TCGTGGACGT
Bacillus l     AAGAATACAT GGGAGACATC ATGGGCGACA TTACGTCACG CCGCGGTCGC
Bacillus h     AAGAGTACAT GGGTGACGTT ATGGGTGACA TCACTTCACG CCGTGGTCGC
Geobacillu     AAGAGTACCT CGGCGACATT ATGGGTGACA TCACGTCCCG CCGCGGCCGC
Oceanobaci     AAGAATACAT GGGAGACATT ATGGGTGACG TAACATCCCG TCGTGGACGC
Aquifex ae     AAAAGTACGT GGGTGACGTT ATAGGTGACC TTAACTCCAG AAGAGGAAAG
Consensus      AARARTAYVT SGGHGAYRTY ATRGGYGAYV THAMNTCHMG HMGHGGHMRB

               ....|....| ....|....| ....|....| ....|....| ....|....|
                    1905       1915       1925       1935       1945
Bacillus a     GTAGAAGGTA TGGAAGCTCG CGGTAACGCT CAAGTTGTTC GCGCTATGGT
Bacillus s     GTAGAAGGTA TGGAAGCACG CGGTAACGCT CAAGTTGTTC GTGCTATGGT
Bacillus t     GTAGAAGGTA TGGAAGCTCG CGGTAACGCT CAAGTTGTTC GCGCTATGGT
Bacillus l     GTAGAAGGTA TGGAAGGACG CGGAAACGCT CAGGTCGTTC GCGCGATGGT
Bacillus h     GTAGAAGGTA TGGAAGCTCG CGGTAATGCT CAAGTGGTTA AAGCATTCGT
Geobacillu     ATCGAGGGGA TGGAAGCGCG CGGCAACGCC CAAGTTGTTC GTGCAATGGT
Oceanobaci     GTAGAAGGTA TGGAAGCACG CGGTACAGCA CAAGTTGTTA AAGCGTTCGT
Aquifex ae     ATTATGGGAA TGGAAAACAA GGGAGTTATA ACAGTCATAA AGGCTCACGT
Consensus      RTHRWRGGDA TGGAARVNMR SGGHRHHRYH MMRGTBRTWM RNGCDHWSGT

               ....|....| ....|....| ....|....| ....|....| ....|....|
                    1955       1965       1975       1985       1995
Bacillus a     TCCACTTTCT GAAATGTTCG GTTATGCAAC GTCATTACGT TCTAACACTC
Bacillus s     TCCTCTTGCT GAAATGTTCG GTTACGCTAC TGCACTTCGT TCTAATACGC
Bacillus t     TCCACTTTCT GAAATGTTCG GTTATGCAAC GTCATTACGT TCTAACACTC
Bacillus l     TCCACTTTCT GAAATGTTCG GTTACGCTAC GGCTCTCCGT TCAAACACAC
Bacillus h     TCCTCTTGCT GAAATGTTCG GTTACGCAAC CTCTTTGCGT TCTCGTACAC
Geobacillu     GCCGATGGCC GAAATGTTCG GTTATGCGAC ATCGCTCCGT TCGAACACGC
Oceanobaci     ACCACTTTCA GAAATGTTTG GTTATGCAAC TGCCCTACGT TCGAACACAC
Aquifex ae     TCCCCTCGCA GAGATGTTCG GATACGCTAC GACGCTCAGG AGCTTGACAC
Consensus      DCCNMTBKCH GARATGTTYG GWTAYGCDAC NDCNYTNMGK WSNHDBACDC
```

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2. Primer site (FusF) was shown in blue and consensus sequence was shown in red (Continued).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                   2005       2015       2025       2035       2045
Bacillus a      AAGGACGCGG AACATTCTCT ATGGTGTTTG ACCACTATGA AGAAGTACCA
Bacillus s      AAGGTCGCGG TACGTTCACT ATGCACATGG ATCACTACGG AGAAGTGCCG
Bacillus t      AAGGACGCGG AACATTCTCT ATGGTGTTTG ACCACTATGA AGAAGTACCA
Bacillus l      AAGGACGCGG TACGTTCACA ATGGTCTTTG ACCACTATGA AGAAGTGCCT
Bacillus h      AAGGACGTGG AACTTACACA ATGTTCTTCG ATCACTACGA GGAAGTACCT
Geobacillu      AAGGACGCGG AACGTTCTCG ATGGTGTTTG ACCACTATGA AGAAGTTCCG
Oceanobaci      AAGGTCGCGG TGTTTACACA ATGCACTTTG ATCACTACGA AGAAACACCG
Aquifex ae      AAGGTAGGGG AACCTTTATA ATGAAATTTT CCCACTACGA CGAAGTTCCG
Consensus       AAGGWMGBGG WRYNTWYWYD ATGNWVWTBK MYCACTAYGR VGAARYDCCD

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   2055       2065       2075       2085       2095
Bacillus a      AAGTCTGTTT CTGAAGAAAT TATCAAAAAA AATAAAGGTG AATAA-----
Bacillus s      AAGAGCGTCG CAGAAGAAAT TATCAAAAAA AATAAAGGCG AATAA-----
Bacillus t      AAGTCTGTTT CTGAAGAAAT TATCAAAAAA AATAAAGGTG AATAA-----
Bacillus l      AAGAGCATCG CTGATGAGAT CATCAAAAAA AATCAAGGTG AATAA-----
Bacillus h      AAGAGCATTT CTGAAGAAAT TATTAAGAAA AACTCTGGTG AATAA-----
Geobacillu      AAAAACATCG CCGATGAAAT TATCAAAAAA AATAAAGGCG AATAA-----
Oceanobaci      AAAAGCATTA CTGAAGAAAT TATTAAGAAA AATGCTGGTG AATAA-----
Aquifex ae      CAGCAAATTG CGGAAAAGAT TATCGGCGAA AG-AATGGCC GGTAAGAGCT
Consensus       MARHVHRTYD CNGAWRARAT YATYRRVRAA ARYNMWGGYS RRTAAGAGCT

                ....|
                   2105
Bacillus a      -----
Bacillus s      -----
Bacillus t      -----
Bacillus l      -----
Bacillus h      -----
Geobacillu      -----
Oceanobaci      -----
Aquifex ae      CTTAA
Consensus       CTTAA
```

Figure 5.3. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.2. Primer site (FusF) was shown in blue and consensus sequence was shown in red.

Table 5.3. Bacterial *rplC* genes used for degenerate primer design.

Organism name	**GeneBank ID**
Bacillus licheniformis ATCC 14580	3027551
Geobacillus kaustophilus HTA426	3183452
Oceanobacillus iheyensis HTE831	1015209
Bacillus cereus ZK	3022952
Bacillus anthracis str. 'Ames Ancestor'	2818190
Bacillus clausii KSM-K16	3201528

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     5          15         25         35         45
Bacillus l      ATGACCAAAG GAATCTTAGG AAGAAAAATT GGTATGACGC AAGTTTTCGC
Geobacillu      ATGACGAAAG GAATCTTAGG AAGAAAGATC GGTATGACGC AAATATTTGC
Oceanobaci      ATGACGAAAG GAATCTTAGG TCGTAAAATC GGCATGACTC AGCTATTCTC
Bacillus c      ATGACCAAAG GAATCTTAGG AAGAAAGATC GGTATGACTC AAGTATTTGC
Bacillus a      ATGACCAAAG GAATCTTAGG AAGAAAGATC GGTATGACTC AAGTATTTGC
Bacillus c      ATGACCAAAG GAATCTTAGG TAGAAAAATC GGTATGACTC AAGTGTTTGC
Consensus       ATGACSAAAG GAATCTTAGG WMGWAARATY GGYATGACKC ARVTDTTYKC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                     55         65         75         85         95
Bacillus l      AGAAAACGGT GATCTTATTC CGGTAACTGT TATCGAAGCT GCTCCTAACG
Geobacillu      GGAAAATGGC GATTTGATTC CGGTAACCGT CATCCATGCG ACGCCGAACG
Oceanobaci      TGATAATGGA GAGCTAATTC CAGTAACAGT AATTCAAGCT GAGCCAAATG
Bacillus c      TGAGAACGGT GAGTTAATCC CAGTAACAGT TATCGCTGCT AATCCAAACG
Bacillus a      TGAGAACGGT GAGTTAATCC CAGTAACAGT TATCGCTGCT AATCCAAACG
Bacillus c      AGAAAACGGA GAAGTCATTC CAGTAACAGT AATTGAAGCA GAACCAAATA
Consensus       DGADAAYGGH GADBTNATYC CRGTAACHGT HATYSMWGCD RMDCCDAAYR

                ....|....| ....|....| ....|....| ....|....| ....|....|
                     105        115        125        135        145
Bacillus l      TTGTGCTTCA GAAGAAAACT TCTGAAAACG ACGGCTATGA AGCGATTCAG
Geobacillu      TCGTCCTGCA AAAGAAAACG ATCGAAAACG ACGGTTACGA AGCGATCCAA
Oceanobaci      TTGTATTGCA AAAGAAAACA TTAGAAAATG ATGGTTACGA AGCGTTACAA
Bacillus c      TTGTTCTTCA AAAGAAAACA ACTGAAACTG ATGGCTACAA CGCAATTCAG
Bacillus a      TTGTTCTTCA AAAGAAAACA ACTGAAACTG ATGGCTACAA CGCAATCCAG
Bacillus c      TTGTCCTTCA AAAGAAGACA GTGGAATCTG ATGGCTATGA GGCGATTCAG
Consensus       TYGTNYTKCA RAAGAARACD DYNGAAWMYG AYGGYTAYRA VGCRWTHCAR

                ....|....| ....|....| ....|....| ....|....| ....|....|
                     155        165        175        185        195
Bacillus l      ATCGGTTTTG ACGACAAGCG TGAAAAGCTT GCCAACAAAC CAGAAAAAGG
Geobacillu      TTAGGTTTTG AAGATATTAG CGAAAAACGC GCCAATAAAC CGCAAATTGG
Oceanobaci      ATCGGTTTTG CTGATAAAAA GGAATCACGT ACTAATAAAG CGGAAAAAGG
Bacillus c      TTAGGATTTG AAGATAAACG TGAAAAGTTA ACTAACAAAC CTGAACAAGG
Bacillus a      TTAGGATTTG AAGATAAACG TGAAAAGTTA ACTAACAAAC CTGAACAAGG
Bacillus c      ATCGGTTTTG CTGATGCGAA AAAGCCGAAC AAACCAGCTA CTGGCCATGC
Consensus       WTMGGWTTTG MHGAYRHDMR NRARHMRHDH RMHMMHRMWV CDSRMMWWGS

                ....|....| ....|....| ....|....| ....|....| ....|....|
                     205        215        225        235        245
Bacillus l      ACACGTTGCA AAAGCGGAAA CTGCTCCTAA GCGCTTCGTT AAGGAATTAC
Geobacillu      CCATGCTGCC AAAGCGAACA CGGCACCTAA GCGCTTCATT CGTGAAATCC
Oceanobaci      TCATGCTGAG AAAGCAGGCA CTGCCCCTAA GCGCTACGTT CGTGAAATCC
Bacillus c      CCACACTGCT AAAGCATCTA CAACTCCTAA GCGCTTCATT CGCGAAATCC
Bacillus a      CCACACTGCT AAAGCATCTA CAACTCCTAA GCGCTTCATT CGCGAAATCC
Bacillus c      CGCTAAAGCT GAAACGGCTC CTAAGCGCTT CATTAAGGAA ATCCGTGGCG
Consensus       HSMYRHWGMN RAARCRDVHM CDRMNCSYWW SMKYWWSRWW MDBSRWDKMS
```

Figure 5.4. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.3. Primer site (RplC) was shown in blue and consensus sequence was shown in red (Continued).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      255        265        275        285        295
Bacillus l      GCGGAGTGGA TATGGATGCG TATGAAGTTG GTCAGGAAGT CAAAGTTGAT
Geobacillu      GCGGCGCCAA CATCGATGAA TATGAAGTTG GCCAAGAAGT GAAAGTTGAC
Oceanobaci      GTGGCGCAGA GGTTAATAGC TTTGAAGTTG GGCAAGAGAT CAAAGTCGAC
Bacillus c      GCGATGCAGA CGTGGACGGA TTAGAGGTTG GTCAAGAGGT AAAAGTTGAC
Bacillus a      GCGATGCAGA CGTGGACGGA TTAGAGGTTG GTCAAGAGGT AAAAGTTGAC
Bacillus c      TTAATCTTGA CGAAGTTGAA GTAGGTCAAG CGATTAACGT AACAACATTT
Consensus       KYRRHSYNRA BRWNRWYRVV KWWGRDSWWG SBMWDRAVRT VAMARYHKWY

                ....|....| ....|....| ....|....| ....|....| ....|....|
                      305        315        325        335        345
Bacillus l      ATTTTCTCTA ACGGAGAAAT CGTAGATGTA ACAGGAACAT CGAAAGGTAA
Geobacillu      ATTTTCAGCG AAGGCGAAAT TGTCGATGTC ACGGGCATTT CGAAAGGGAA
Oceanobaci      ATCTTTGAGT CTGGAGAAAA AATCGATGTA ACTGGTACAT CTAAAGGGAA
Bacillus c      GTATTCGCTA CAGGTGAAAT CGTTGATGTA ACAGGAATTT CTAAAGGTAA
Bacillus a      GTATTCGCTA CAGGTGAAAT CGTTGATGTA ACAGGAATTT CTAAAGGTAA
Bacillus c      GCAGCAGGAG ATCTTGTTGA CGTAACAGGA ACATCTAAAG GGAAAGGGTT
Consensus       RYHKYHDVND MHSKHGWWRW HRTHRMWGKM ACDKSHAHWK SKAAAGGKWW

                ....|....| ....|....| ....|....| ....|....| ....|....|
                      355        365        375        385        395
Bacillus l      AGGTTTCCAA GGCGCTATCA AACGCCACGG ACAATCTCGC GGACCAATGT
Geobacillu      AGGTTTCCAA GGGGCGATCA AACGCCACGG TCAATCACGC GGACCAATGG
Oceanobaci      AGGTTTCCAG GGTGTAATTA AGCGTCATGG TCAACAACGT GGACCAACTA
Bacillus c      AGGTTTCCAA GGTGTTATCA AACGCCACGG ACAATCTCGC GGACCTATGT
Bacillus a      AGGTTTCCAA GGTGTTATCA AACGCCACGG ACAATCTCGC GGACCTATGT
Bacillus c      CCAAGGCGCG ATCAAACGCC ATAACCAATC CCGTGGACCA ATGTCACACG
Consensus       MSRWKKCSMR RKBRHDMKYM ADMRYCAHKS HCRWBVWCSH RKRYCWMHBD

                ....|....| ....|....| ....|....| ....|....| ....|....|
                      405        415        425        435        445
Bacillus l      CTCACGGTTC ACGCTACCAC CGCCGTCCTG GTTCAATGGG ACCGGTAGAC
Geobacillu      CTCACGGTTC CCGTTATCAT CGCCGTCCGG GTTCGATGGG GTCCATCGCG
Oceanobaci      CACACGGTTC TCATTTCCAT AGAGCTCCAG GTGCAATGGG TGTAATTGAT
Bacillus c      CTCATGGTTC TCGCTATCAC CGTCGTCCAG GTTCAATGGG CCCAGTTGCT
Bacillus a      CTCATGGTTC TCGCTATCAC CGTCGTCCAG GTTCAATGGG CCCAGTTGCT
Bacillus c      GCTCTCGTTA CCACCGTCGT CCAGGTTCAA TGGGCCCTGT AGCGCCAAAC
Consensus       SHYMYSGTTM HCRYYDYCRY MSHSSTYCDR KKKSVMYKGK NBYVVYHRMB

                ....|....| ....|....| ....|....| ....|....| ....|....|
                      455        465        475        485        495
Bacillus l      CCTAACCGTG TATTCAAAGG AAAGCTGTTG CCTGGACGTA TGGGCGGAGA
Geobacillu      CCAAACCGCG TTTTCAAAAC GAAAAACTTG CCGGGCCGCA TGGGCGGTGA
Oceanobaci      CCAATGCGTG TATTCAAAGG TAAAAAACTA CCAGGTCACA TGGGTTCTGA
Bacillus c      CCGAACCGTG TATTCAAAGG CAAAAAACTT GCTGGACGTA TGGGTGGAGA
Bacillus a      CCGAACCGTG TATTCAAAGG CAAAAAACTT GCTGGACGTA TGGGTGGAGA
Bacillus c      CGTGTATTTA AAGGAAAAGC GCTTCCTGGC CGTATGGGCG GAGAACAAAT
Consensus       CSDRWVYKYR WWKKMAAARS NMWDMHNBKN SSDRKNSRYR KRGRHBVWRW
```

Figure 5.4. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.3. Primer site (RplC) was shown in blue and consensus sequence was shown in red (Continued).

```
               ....|....| ....|....| ....|....| ....|....| ....|....|
                     505        515        525        535        545
Bacillus l     GCAAATCACT GTTCAAAACC TTGAAATCGT AAAAGTAGAT GCAGAGCGCA
Geobacillu     GCGCGTGACG ATCCAAAACT TGAAAATCGT GAAAGTGGAT CCGGAGCGCA
Oceanobaci     ACAAGTAACG ATCCAAAACC TTGAAGTTGT AAGTGTAGAT ACGGATAAGA
Bacillus c     CCAAGTTACT ATCCAAAACT TAGAAATCGT TCAAGTTGAC ACTGAGCGCA
Bacillus a     CCAAGTTACT ATCCAAAACT TAGAAATCGT TCAAGTTGAC ACTGAGCGCA
Bacillus c     TACAATGCAA AACCTTGAGA TCGTTCGCGT TGACGAAGAG CGCAACTTGC
Consensus      NMVMRTNMMD RWYCWWRASH TNRWWVKYGT DVRHGWDGAB VSNRABHDSM

               ....|....| ....|....| ....|....| ....|....| ....|....|
                     555        565        575        585        595
Bacillus l     ACCTTCTTTT AGTTAAAGGT AACGTTCCTG GTGCCAAAAA ATCACTAGTC
Geobacillu     ACTTGCTGCT CATTAAAGGC AATGTGCCGG GTCCGAGAAA AGGCTTAGTC
Oceanobaci     ATCTACTACT AATTAAAGGT AACGTGCCTG GCGCGAAAAA ATCTTACGTT
Bacillus c     ACTTATTATT AGTAAAAGGT AACGTTCCAG GTGCTAAGAA ATCTCTTGTA
Bacillus a     ACTTATTATT AGTAAAAGGT AACGTTCCAG GTGCTAAGAA ATCTCTTGTA
Bacillus c     TTCTCGTTAA AGGCAACGTA CCAGGGGCTA AGAAAAGCTA TGTAACTGTA
Consensus      WYYTNBTDHW MRKHAAMGKH MMHGKKSCDR RBVMNARVWA WKBHHHHGTH

               ....|....| ....|....| ....|....| ....|....| ..
                     605        615        625        635
Bacillus l     ACTGTTAAAA GTGCTGTTAA ATCTAAATAA .......... ..
Geobacillu     ATCGTCAAAA GCGCCGTGAA AGCGGCGAAG AAGGCGAAAT AA
Oceanobaci     AAAATCACAA GTGCAGTAAA AGGTAACTAA .......... ..
Bacillus c     GTTGTTCAAG GCGCTGTGAA GGTTAGCAAA TAA....... ..
Bacillus a     GTTGTTCAAG GCGCTGTGAA GGTTAGCAAA TAA....... ..
Bacillus c     CAATCTGCAG TTAAGGCGTA A......... .......... ..
Consensus      VHHDYYVMAR KYRMNGYDWA RKBKRVVWAR WARGCGAAAT AA
```

Figure 5.4. ClustalW multiple aligment result of nucleotide sequences indicated in Table 5.3. Primer site (RplC) was shown in blue and consensus sequence was shown in red.

5.1.3. Partial PCR Amplification of *tuf* Gene

By using degenerate TuF413 and TuR1174 primers, *tuf* gene was partially amplified in a gradient PCR in the range of 50-60 °C annealing temperature. As an optimal degree, 56 °C annealing temperature was chosen. According to the ClustalW multiple alignment results of the species related to *G. anatolicus*, expected PCR product size was 761 bp. At 56 °C, PCR product appeared as a single band at the expected size when analyzed electrophoretically on 1 % agarose gels (Figure 5.5.).

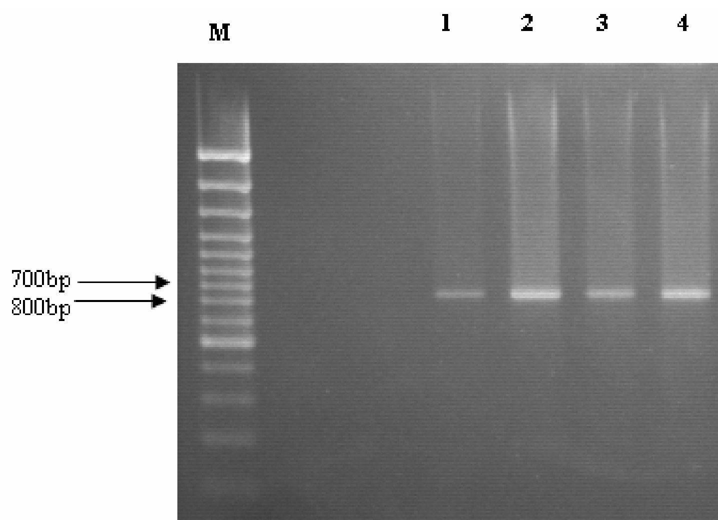

Figure 5.5. Partial PCR amplification of *G.anatolicus tuf* gene. M, molecular size marker; 1, 2, 3, 4, PCR product containing 761 bp partial *tuf* gene at 56 °C annealing temperature at 1,5 mM, 2 mM, 3 mM and 4 mMof magnesium concentrations

5.1.4. Confirmation of Partial *tuf* Sequence

After obtaining of the DNA fragment that was supposed to contain a part of a *tuf* sequence, this fragment was sequenced by using using TuF413 and TuR1174 primers,. The sequence data obtained was compared with other EF-Tu genes in GeneBank and confirmed that this fragment indeed contains part of a *tuf* gene. This sequence will later be used to design sequencing primers for the complete *Geobacillus anatolicus tuf* gene.

5.1.5. Amplification of the Fragment Containing Complete *tuf* Sequence

To amplify the fragment that contains the complete *tuf* gene, FusF and RplC primers located in the adjacent genes of *tuf* were used. Expected product size was determined as approximately 2100 bp by analyzing genomic sequences of various phylogenetically related species indicated in Table 5.2. and Table 5.3. PCR results were shown in Figure 5.6. confirming that the PCR products were at the expected size. Before sequencing of this fragment, it was confirmed, by an additional PCR amplification using TuF413 and TuR1174 primers that *tuf* gene is indeed present in this fragment (Figure 5.7.)

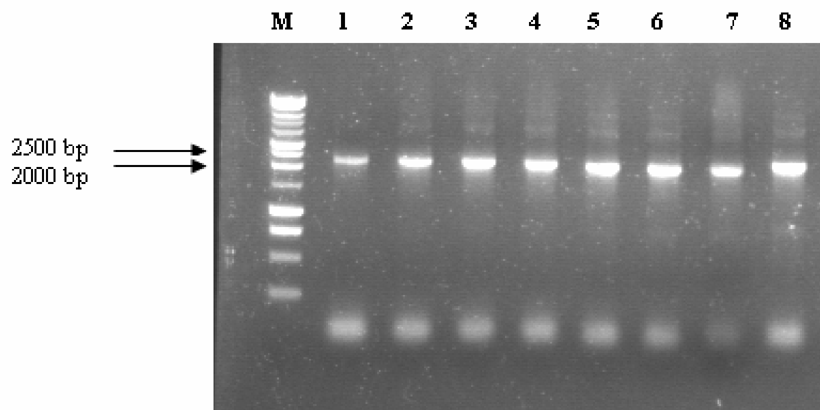

Figure 5.6. PCR amplification using FusF and RplC primers. M, molecular weight marker, 1, 2, 3, 4, 5, 6, 7, 8, PCR results in the range of 50-60 °C annealing temperature

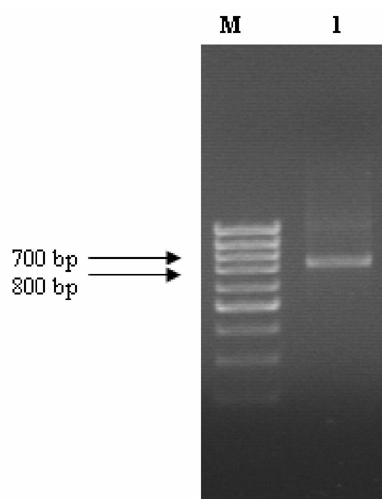

Figure 5.7. Confirmation of the presence of the *tuf* gene in the fragment amplified by using FusF and RplC primers. M, molecular weight marker; 1, PCR product by using TuF413 and TuR1174 primers

5.1.6. Complete *G.anatolicus tuf* Sequence

At first, the fragment amplified by using TuF413 and TuR1174 primers was sequenced with the same primers. After sequencing of this fragment with TuF413 and TuR1174, obtained sequence data was used to design new non-degenerate primers TuF888

and TuR776. The fragment amplified with FusF and RplC primers carrying the complete *tuf* gene was then sequenced with TuF888, TuR776, FusF and RplC primers. Obtained six sequence data was aligned and used to construct the complete *tuf* sequence. 1188 bp long complete *tuf* gene was shown in the Figure 5.8. Nucleic Acid Statistics (NASTATS) were given in Table 5.4.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                    5         15         25         35         45
EF-Tu G. a      ATGGCTAAAG CGAAATTTGA GCGTACGAAA CCGCACGTCA ACATTGGCAC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                    55         65         75         85         95
EF-Tu G. a      GATCGGCCAC GTTGACCATG GGAAAACGAC GTTGACAGCT GCGATCACAA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   105        115        125        135        145
EF-Tu G. a      CAGTTCTTGC GAAACAAGGG AAAGCGGAAG CAAGAGCGTA CGACCAAATC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   155        165        175        185        195
EF-Tu G. a      GACGCAGCTC CGGAAGAGCG TGAACGCGGA ATCACGATTT CGACAGCGCA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   205        215        225        235        245
EF-Tu G. a      CGTTGAGTAT GAAACGGACG CTCGTCACTA TGCGCACGTT GACTGCCCAG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   255        265        275        285        295
EF-Tu G. a      GCCACGCTGA CTACGTGAAA AACATGATCA CGGGCGCAGC GCAAATGGAC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   305        315        325        335        345
EF-Tu G. a      GGCGCGATCC TTGTTGTATC GGCTGCTGAC GGTCCGATGC CGCAAACGCG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   355        365        375        385        395
EF-Tu G. a      CGAACACATT CTTCTCTCCC GCCAAGTCGG TGTACCGTAC ATCGTTGTTT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   405        415        425        435        445
EF-Tu G. a      TCTTGAACAA ATGCGACATG GTGGACGACG AAGAATTGCT TGAACTCGTT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   455        465        475        485        495
EF-Tu G. a      GAAATGGAAG TTCGCGATCT TCTCTCTGAA TACGACTTCC CGGGCGACGA
```

Figure 5.8. Complete *G. anatolicus tuf* sequence. Start and stop codons were shown as bold (Continued).

```
                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     505        515        525        535        545
EF-Tu G. a       AGTGCCGGTC ATCAAAGGTT CGGCATTAAA AGCGCTCGAA GGCGACCCGC

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     555        565        575        585        595
EF-Tu G. a       AATGGGAAGA AAAAATCATT GAACTGATGA ACGCGGTTGA CGAGTACATC

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     605        615        625        635        645
EF-Tu G. a       CCGACTCCGC AACGTGAAGT AGACAAACCG TTCATGATGC CGGTTGAGGA

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     655        665        675        685        695
EF-Tu G. a       CGTTTTCTCG ATCACAGGCC GCGGTACGGT TGCGACGGGC CGTGTTGAGC

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     705        715        725        735        745
EF-Tu G. a       GCGGTACGTT AAAAGTCGGC GACCCGGTTG AAATTATCGG TCTTTCGGAC

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     755        765        775        785        795
EF-Tu G. a       GAGCCGAAAA CGACGACGGT TACGGGTGTA GAAATGTTCC GTAAGCTGCT

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     805        815        825        835        845
EF-Tu G. a       TGACCAAGCG GAAGCTGGCG ACAACATCGG TGCGCTTCTC CGCGGTGTAT

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     855        865        875        885        895
EF-Tu G. a       CGCGTGACGA AGTTGAGCGC GGCCAAGTAT TGGCGAAACC GGGCTCGATC

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     905        915        925        935        945
EF-Tu G. a       ACGCCGCACA CGAAATTTAA AGCGCAAGTT TACGTTCTGA CGAAAGAAGA

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     955        965        975        985        995
EF-Tu G. a       AGGCGGACGC CATACTCCGT TCTTCTCGAA CTACCGTCCG CAATTCTACT

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     1005       1015       1025       1035       1045
EF-Tu G. a       TCCGTACAAC GGACGTAACG GGCATCATCA CGCTTCCGGA AGGCGTAGAA

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     1055       1065       1075       1085       1095
EF-Tu G. a       ATGGTTATGC CTGGCGACAA CGTTGAAATG ACGGTTGAAC TGATCGCTCC

                 ....|....| ....|....| ....|....| ....|....| ....|....|
                     1105       1115       1125       1135       1145
EF-Tu G. a       GATCGCGATC GAGGAAGGAA CAAAATTCTC GATCCGTGAA GGCGGCCGCA

                 ....|....| ....|....| ....|....| ....|...
                     1155       1165       1175       1185
EF-Tu G. a       CGGTTGGTGC TGGTTCCGTA TCGGAAATCA TTGAGTAA
```

Figure 5.8. Complete *G. anatolicus tuf* sequence. Start and stop codons were shown as bold (Continued).

Table 5.4. NATSTATS result *G. anatolicus tuf* sequence

	Total #	**Percentage**
All nucleotides	**1188**	
A	**312**	**26.3 %**
T	**240**	**20.2 %**
G	**299**	**25.2 %**
C	**337**	**28.4 %**
A and T	**552**	**46.5 %**
C and G	**636**	**53.5 %**

By using complete *tuf* sequence obtained, *G. anatolicus* complete EF-Tu amino acid sequence was determined. This sequence was shown in Figure 5.9. The amino acid sequence of *G. anatolicus* EF-Tu was aligned and compared with other thermophilic EF-Tu sequences using ClustalW multiple alignment. The species used in this aligment were as follows: *Geobacillus stearothermophilus*, *Aquifex pyrophilus*, *Aquifex aeolicus*, *Thermus aquaticus*, *Thermus thermophilus HB27*, *Thermotoga maritima*, *Thermoanaerobacter tengcongensis*. The alignment result and conserved regions were shown in Figure 5.10.

```
MAKAKFERTK PHVNIGTIGH VDHGKTTLTA AITTVLAKQG KAEARAYDQI
DAAPEERERG ITISTAHVEY ETDARHYAHV DCPGHADYVK NMITGAAQMD
GAILVVSAAD GPMPQTREHI LLSRQVGVPY IVVFLNKCDM VDDEELLELV
EMEVRDLLSE YDFPGDEVPV IKGSALKALE GDPQWEEKII ELMNAVDEYI
PTPQREVDKP FMMPVEDVFS ITGRGTVATG RVERGTLKVG DPVEIIGLSD
EPKTTTVTGV EMFRKLLDQA EAGDNIGALL RGVSRDEVER GQVLAKPGSI
TPHTKFKAQV YVLTKEEGGR HTPFFSNYRP QFYFRTTDVT GIITLPEGVE
MVMPGDNVEM TVELIAPIAI EEGTKFSIRE GGRTVGAGSV SEIIE
```

Figure 5.9. Complete *G. anatolicus* EF-Tu amino acid sequence.

```
               ....|....| ....|....| ....|....| ....|....| ....|....|
                   5          15         25         35         45
Geobacillu     MAKAKFERTK PHVNIGTIGH VDHGKTTLTA AITTVLAK-- -QGK-AEAKA
Aquifex py     MAKEKFERTK EHVNVGTIGH VDHGKSTLTS AITCVLAAGL VEGGKAKCFK
Aquifex ae     MAKEKFERTK EHVNVGTIGH VDHGKSTLTS AITCVLAAGL VEGGKAKCFK
Thermus aq     MAKGEFIRTK PHVNVGTIGH VDHGKTTLTA ALTYVAAAE- --NPNVEVKD
Thermus th     MAKGEFIRTK PHVNVGTIGH VDHGKTTLTA ALTYVAAAE- --NPNVEVKD
Thermotoga     MAKEKFVRTK PHVNVGTIGH IDHGKSTLTA AITKYLSL-- --KGLAQYIP
Thermoanae     MAKQKFERTK PHVNVGTIGH VDHGKTTLTA AITLILSKA- ---GLAQAKG
G. anatoli     MAKAKFERTK PHVNIGTIGH VDHGKTTLTA AITTVLAK-- -QGK-AEARA
Consensus      MAK  F RTK  HVN GTIGH  DHGK TLT  A T

               ....|....| ....|....| ....|....| ....|....| ....|....|
                   55         65         75         85         95
Geobacillu     YDQIDAAPEE RERGITISTA HVEYETEARH YAHVDCPGHA DYVKNMITGA
Aquifex py     YEEIDKAPEE KERGITINIT HVEYETAKRH YAHVDCPGHA DYIKNMITGA
Aquifex ae     YEEIDKAPEE KERGITINIT HVEYETAKRH YAHVDCPGHA DYIKNMITGA
Thermus aq     YGDIDKAPEE RARGITINTA HVEYETAKRH YSHVDCPGHA DYIKNMITGA
Thermus th     YGEIDKAPEE RARGITINTA HVEYETAKRH YSHVDCPGHA DYIKNMITGA
Thermotoga     YDQIDKAPEE KARGITINIT HVEYETEKRH YAHIDCPGHA DYIKNMITGA
Thermoanae     YDEIDKAPEE KARGITINTT HVEYETAKRH YAHVDCPGHA DYVKNMITGA
G. anatoli     YDQIDAAPEE RERGITISTA HVEYETDARH YAHVDCPGHA DYVKNMITGA
Consensus      Y  ID APEE   RGITI    HVEYET  RH Y H DCPGHA DY KNMITGA

               ....|....| ....|....| ....|....| ....|....| ....|....|
                  105        115        125        135        145
Geobacillu     AQMDGAILVV SAADGPMPQT REHILLSRQV GVPYIVVFLN KCDMVDDEEL
Aquifex py     AQMDGAILVV SAADGPMPQT REHVLLARQV NVPYIVVFMN KCDMVDDEEL
Aquifex ae     AQMDGAILVV SAADGPMPQT REHVLLARQV NVPYIVVFMN KCDMVDDEEL
Thermus aq     AQMDGAILVV SAADGPMPQT REHILLARQV GVPYIVVFMN KVDMVDDPEL
Thermus th     AQMDGAILVV SAADGPMPQT REHILLARQV GVPYIVVFMN KVDMVDDPEL
Thermotoga     AQMDGAILVV AATDGPMPQT REHVLLARQV EVPYMIVFIN KTDMVDDPEL
Thermoanae     AQMDGAILVV SAADGPMPQT REHILLARQV GVPYIVVFLN KADMVDDPEL
G. anatoli     AQMDGAILVV SAADGPMPQT REHILLSRQV GVPYIVVFLN KCDMVDDEEL
Consensus      AQMDGAILVV  A DGPMPQT REH LL RQV  VPY  VF N K DMVDD EL

               ....|....| ....|....| ....|....| ....|....| ....|....|
                  155        165        175        185        195
Geobacillu     LELVEMEVRD LLSEYDFPGD EVPVIKGSAL KALEGDP--- -------KWE
Aquifex py     LELVELEVRE LLSKYEYPGD EVPVIRGSAL GALQELEQNS PG-----KWV
Aquifex ae     LELVELEVRE LLSKYEYPGD EVPVIRGSAL GALQELEQNS PG-----KWV
Thermus aq     LDLVEMEVRD LLNQYEFPGD EVPVIRGSAL LALEEMHKNP KTKRGENEWV
Thermus th     LDLVEMEVRD LLNQYEFPGD EVPVIRGSAL LALEQMHRNP KTRRGENEWV
Thermotoga     IDLVEMEVRD LLSQYGYPGD EVPVIRGSAL KAVEAP-NDP NH-----EAY
Thermoanae     IELVEMEVRD LLNQYEFPGD DTPIVVGSAL KALECGCGKR ECQ-----WC
G. anatoli     LELVEMEVRD LLSEYDFPGD EVPVIKGSAL KALEGDP--- -------QWE
Consensus        LVE EVR  LL  Y  PGD   P   GSAL  A
```

Figure 5.10. Sequence aligment of thermophilic EF-Tu’s and G.anatolicus EF-Tu amino acid sequences. Conserved regions were shown as bold (Continued).

```
             ....|....| ....|....| ....|....| ....|....| ....|....|
                  205        215        225        235        245
Geobacillu   EKIIELMNAV DEYIPTPQRE VDKPFMMPIE DVFSITGRGT VATGRVERGT
Aquifex py   GSIKELLNAM DEYIPTPERE VDKPFLMPIE DVFSISGRGT VVTGRVERGV
Aquifex ae   ESIKELLNAM DEYIPTPQRE VDKPFLMPIE DVFSISGRGT VVTGRVERGV
Thermus aq   DKIWELLDAI DEYIPTPVRD VDKPFLMPVE DVFTITGRGT VATGRIERGK
Thermus th   DKIWELLDAI DEYIPTPVRD VDKPFLMPVE DVFTITGRGT VATGRIERGK
Thermotoga   KPIQELLDAM DNYIPDPQRD VDKPFLMPIE DVFSITGRGT VVTGRIERGR
Thermoanae   GKIWELMDVV DEYIPTPERD IDKPFLMPVE DVFTITGRGT VATGRVERGK
G. anatoli   EKIIELMNAV DEYIPTPQRE VDKPFMMPVE DVFSITGRGT VATGRVERGT
Consensus      I EL     D YIP P R   DKPF MP E DVF I GRGT V TGR ERG

             ....|....| ....|....| ....|....| ....|....| ....|....|
                  255        265        275        285        295
Geobacillu   LKVGDPVEII GLSDEPKATT VTGVEMFRKL LDQAEAGDNI GALLRGVSRD
Aquifex py   LRPGDEVEIV GLREEPLKTV ATSIEMFRKV LDEALPGDNI GVLLRGVGKD
Aquifex ae   LRPGDEVEIV GLREEPLKTV ATSIEMFRKV LDEALPGDNI GVLLRGVGKD
Thermus aq   VKVGDEVEIV GLAPETRKTV VTGVEMHRKT LQEGIAGDNV GLLLRGVSRE
Thermus th   VKVGDEVEIV GLAPETRKTV VTGVEMHRKT LQEGIAGDNV GVLLRGVSRE
Thermotoga   IRPGDEVEII GLSYEIKKTV VTSVEMFRKE LDEGIAGDNV GCLLRGIDKD
Thermoanae   VKVGDEVEII GLTTESRKTV VTGVEMFRKT LDEAQAGDNI GVLLRGIQRD
G. anatoli   LKVGDPVEII GLSDEPKTTT VTGVEMFRKL LDQAEAGDNI GALLRGVSRD
Consensus       GD VEI  GL  E   T   T  EM RK  L     GDN  G LLRG

             ....|....| ....|....| ....|....| ....|....| ....|....|
                  305        315        325        335        345
Geobacillu   EVERGQVLAK PGSITPHTKF KAQVYVLTKE EGGRHTPFFS NYRPQFYFRT
Aquifex py   DVERGQVLAQ PGSVKAHRKF RAQVYVLSKE EGGRHTPFFV NYRPQFYFRT
Aquifex ae   DVERGQVLAQ PGSVKAHKRF RAQVYVLSKE EGGRHTPFFV NYRPQFYFRT
Thermus aq   EVERGQVLAK PGSITPHTKF EASVYILKKE EGGRHTGFFT GYRPQFYFRT
Thermus th   EVERGQVLAK PGSITPHTKF EASVYVLKKE EGGRHTGFFS GYRPQFYFRT
Thermotoga   EVERGQVLAA PGSIKPHKRF KAQIYVLKKE EGGRHTPFTK GYKPQFYIRT
Thermoanae   EVERGQVLAK PGTIKPHTKF EAQVYVLTKE EGGRHTPFFN GYRPQFYFRT
G. anatoli   EVERGQVLAK PGSITPHTKF KAQVYVLTKE EGGRHTPFFS NYRPQFYFRT
Consensus     VERGQVLA  PG    H  F  A  Y L KE EGGRHT F    Y PQFY RT

             ....|....| ....|....| ....|....| ....|....| ....|....|
                  355        365        375        385        395
Geobacillu   TDVTG-IITL PEGVEMVMPG DNVEMTVELI APIAIEEGTK FSIREGGRTV
Aquifex py   ADVTGTVVKL PEGVEMVMPG DNVELEVELI APVALEEGLR FAIREGGRTV
Aquifex ae   ADVTGTVVKL PEGVEMVMPG DNVELEVELI APVALEEGLR FAIREGGRTV
Thermus aq   TDVTG-VVRL PQGVEMVMPG DNVTFTVELI KPVALEEGLR FAIREGGRTV
Thermus th   TDVTG-VVQL PPGVEMVMPG DNVTFTVELI KPVALEEGLR FAIREGGRTV
Thermotoga   ADVTGEIVGL PEGVEMVMPG DHVEMEIELI YPVAIEKGQR FAVREGGRTV
Thermoanae   TDVTG-TIQL PEGVEMVMPG DHVTLRVELI TPIAMEEGLK FAIREGGRTV
G. anatoli   TDVTG-IITL PEGVEMVMPG DNVEMTVELI APIAIEEGTK FSIREGGRTV
Consensus     DVTG    L P GVEMVMPG D V    ELI  P A E G   F  REGGRTV
```

Figure 5.10. Sequence aligment of thermophilic EF-Tu's and G.anatolicus EF-Tu amino acid sequences. Conserved regions were shown as bold (Continued).

```
               ....|....|
                  405
Geobacillu    GAGSVSEIIE
Aquifex py    GAGVVTKILD
Aquifex ae    GAGVVTKILD
Thermus aq    GAGVVTKILE
Thermus th    GAGVVTKILE
Thermotoga    GAGVVTEVIE
Thermoanae    GAGVVSAIIE
G. anatoli    GAGSVSEIIE
Consensus     GAG V
```

Figure 5.10. Sequence aligment of thermophilic EF-Tu's and G.anatolicus EF-Tu amino acid sequences. Conserved regions were shown as bold.

5.2. Cloning of *tuf* Gene of *G. anatolicus*

5.2.1. Cloning Method

In this study, pT7D3 vector was used as the cloning vector. pT7D3 was derived from a commercially available pCR®T7/NT-TOPO® vector. pCR®T7/NT-TOPO® contains the bacterial phage T7 promoter site for high-level protein expression in *Escherichia coli*, N-terminal Xpress™ epitope for detection with an Anti-Xpress™ Antibody, N-terminal polyhistidine (6xHis) tag for purification using Ni-affinity chromatography, detection site with an Anti-HisG Antibody, enterokinase cleavage site for efficient removal of the N-terminal fusion tag and ampicilin resistance site for detection of transformants carrying plasmid. In this study, original pCR®T7/NT-TOPO® vector wasn't used because of enterokinase encountered non-specific cleavage problems in previous experiments in our laboratory. Instead of this, NdeI and Csp45I double digested pCR®T7/NT-TOPO® plasmid was used as a cloning vector (Figure 5.11.). This new vector was called pT7D3.

Figure 5.11. pCR®T7/NT-TOPO® cloning sites (BstBI cleavage site is identical to Csp45I)

5.2.2. Constructing Primers for Cloning of *tsf* Gene

To ligate *tuf* gene into pT7D3 vector, two appropriate forward and reverse primers named TuF2 and TuR2 were designed. These primers contained additional sites for appropriate digestion and for ligation steps necessary during the cloning of the *tuf* gene. Primers and additional sites were shown in Figure 5.12.

TuF2: 5'- **CGG GGG GCA TAT G**GC TAA AGC GAA ATT TGA GCG TAC GAA A - 3'

1 **2**

TuR2: 5'- **CAA AAA ATT CGA ATT AAT GAT GAT GAT GAT GAT G**CT CAA TGA

1 **2** **3** **4**

TTT CCG ATA CGG AAC CAG C - 3'

Figure 5.12. TuF2: 1, 7 bp long overhang for efficent digestion; 2, NdeI cut site (it contains translation start codon ATG); the rest of primer corresponds to the beginning of *tuf* gene
TuR2: 1, 7 bp long overhang for efficent digestion; 2, Csp45I cut site; 3, stop codon; 4, 6xHis tag coding region; the rest of primer corresponds to the end of *tuf* gene

After construction of TuF2 and TuR2 primers, these primers were used to amplify the complete *G. anatolicus tuf* gene. The expected product size was approximately 1200

bp. The results of a gradient PCR in the range of 50-65 °C. as analyzed electrophoretically on a 1% agarose gel is shown in Figure 5.13.

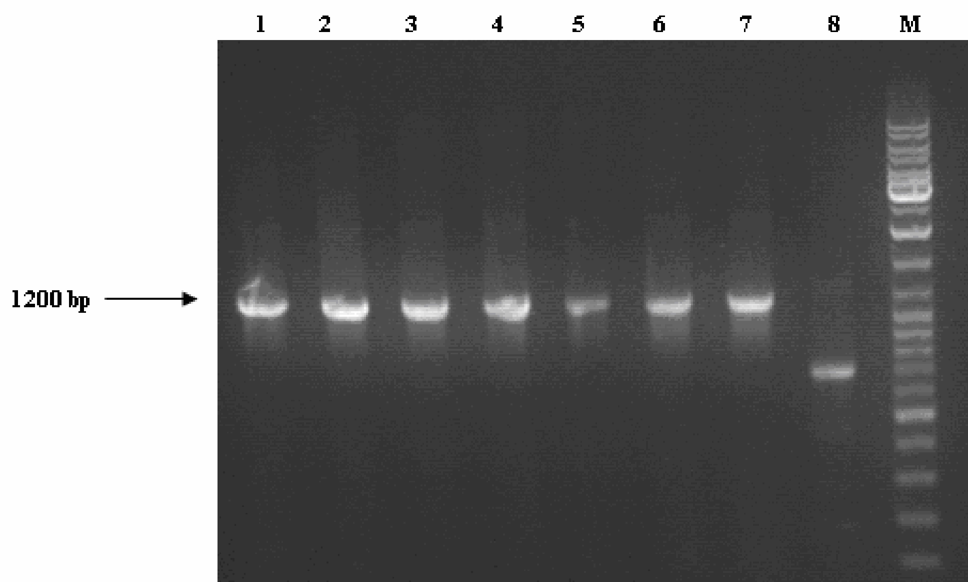

Figure 5.13. Result of the PCR amplification using TuF2 and TuR2 primers; 1, 2, 3, 4, 5, 6, 7, the fragment containing *tuf* gene with annealing temperatures in the range of 50-65°C, 8, positive control with the primers TuF413 and TuR1174, M, DNA molecular weight marker

5.2.3. Construction of pT7D3Tu

pCR®T7/NT-TOPO® vector was digested with Nde I and Csp45 I restriction enzymes. The resulting linear vector was named as pT7D3. The PCR fragment that was amplified using TuF2 and TuR2 pair was ligated into pT7D3. The resulting circular plasmid was named as pT7D3Tu. pT7D3Tu was used to transform *Escherichia coli* JM109 (DE3) cells. After transformation, plasmids were isolated from transformants. Analysis of purified plasmids on 1% agarose gel confirmed the insertion of the fragments into the vector (Figure 5.14.).

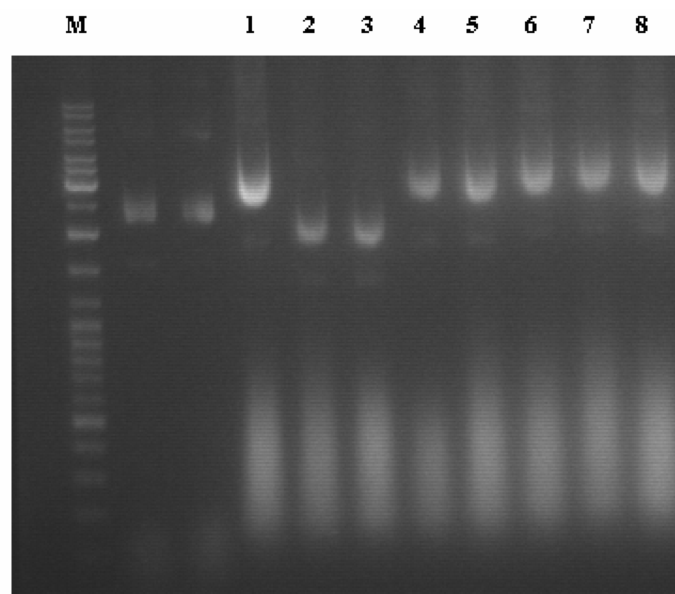

Figure 5.14. Result of plasmid isolation from transformants; M, molecular size marker; 1, 4, 5, 6, 7, 8, plasmids with *tuf* gene; 2, 3, plasmids without *tuf* gene

5.2.4. Confirmation of the presence of the *tuf* Gene in the pT7D3Tu vector

To confirm that the *tuf* gene is actually present in the pT7D3Tu vector, a PCR amplification was performed using two primers: T7 which is located on the plasmid and TuR1174 which is located on the *tuf* sequence. Expected product size of this reaction was approximately 1200 bp. Result of this PCR reaction confirmed that *tuf* gene was present in the pT7D3Tu (Figure 5.12.). At the same time, another PCR amplification was performed using TuF2 and TuR2 primer pair, for the confirmation of the correct insertion of the tuf gene by sequencing (data not shown). Expected product size in this reaction was approximately 1200 bp. Result of this PCR amplification also confirmed that *tuf* gene was present in the pT7D3Tu (Figure 5.15.).

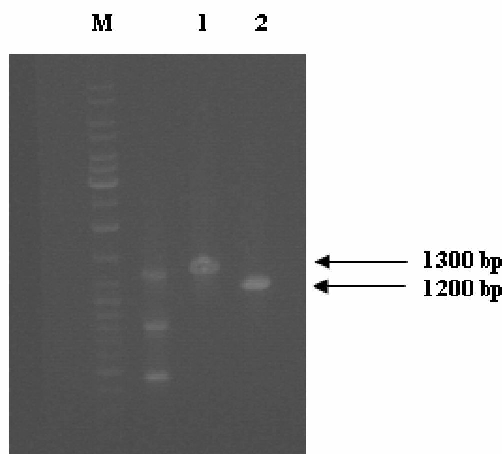

Figure 5.15. PCR results on the pT7D3Tu plasmid; M, molecular size marker; 1, PCR result with T7-pRSET primer pair; 2 PCR result with T7-TuR1174 primer pair.

5.3. Expression of *G. Anatolicus tuf* Gene in JM109(DE3) Cells

In this study, for the expression of *tuf* gene, JM109(DE3) cells were used as the plasmid host. The plasmid pCR®T7/NT-TOPO® and thus, its derivative pT7D3Tu contains T7 RNA polymerase promoter site in front of the Shine-Dalgarno sequence of the inserted gene. In the presence of IPTG, pT7D3Tu can be induced to express the recombinant *G. anatolicus tuf* protein in *E.coli* JM109(DE3) cells because this DE3 derivative of the JM109 contain T7 RNA polymerase gene. The JM109(DE3) cells were induced by IPTG to express the *tuf* gene. As a control the JM109(DE3) cells transformed with pCR®T7/NT-TOPO® (without *tuf* gene) were used. Size of the protein band appeared for the cells carrying pT7D3Tu but absent for pCR®T7/NT-TOPO® was very close to the calculated molecular weight of the protein (44.1 kDa; *G. anatolicus* EF-Tu [43.3 kDa] + 6xHis-tag) when analyzed electrophoretically by 10% SDS-PAGE. The amount of recombinant protein increased as time passed. Results of the experiment were shown in Figure 5.16.

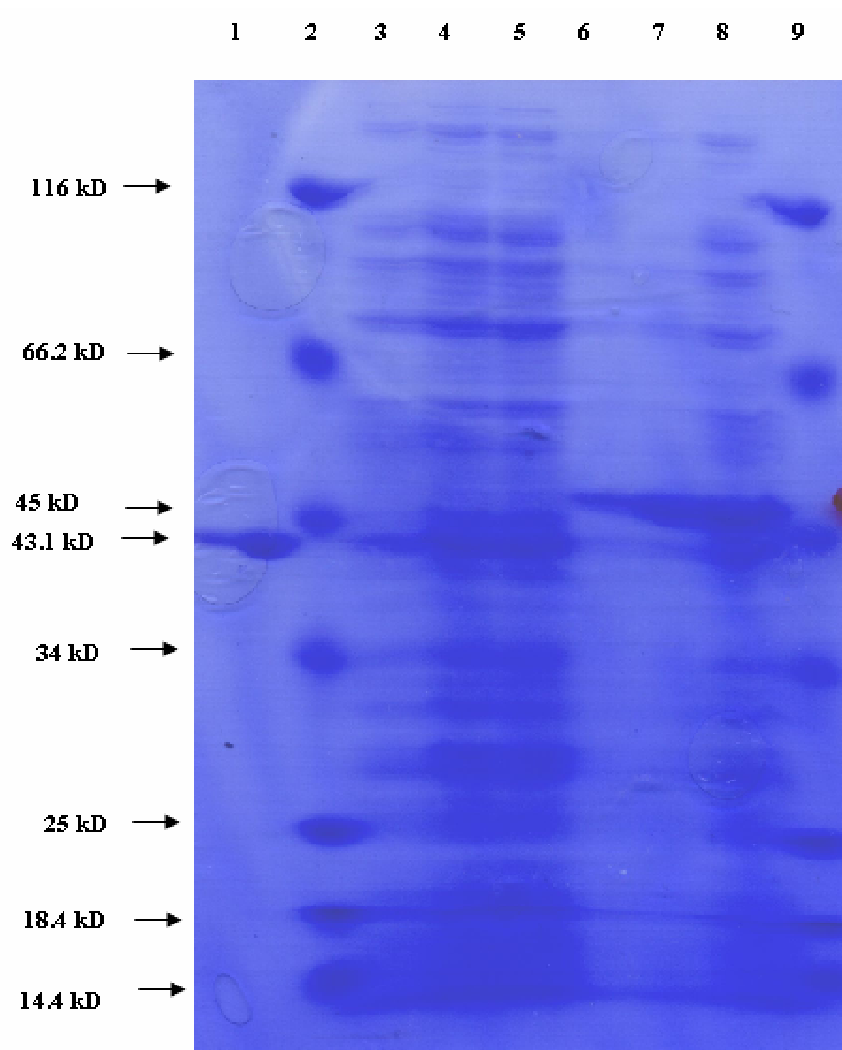

Figure 5.16. Results of expression experiment. Lane 1, *E.coli* EF-Tu (43.1 kD); Lanes 2 and 9, protein molecular weight markers; Lanes 3,4 and 5, JM109(DE3) cells carrying pCR®T7/NT-TOPO® (-*tuf*) at 0, 2, 5 hours respectively; Lanes 6,7 and 8, JM109(DE3) cells carrying pT7D3Tu at 0, 2, 5 hours, respectively

5.4. Ni-Affinity Column Chromatography

Because the recombinant G. anatolicus EF-Tu includes an additional 6 Histidine residues at its C-terminal, it is purified by using Ni-Affinity Column Chromatography. *G. anatolicus* Elongation Factor Tu was isolated from JM109(DE3) carrying the pT7D3Tu vector. For preparative purpose 800 ml of culture was induced with IPTG and 6 hrs after

induction cells were harvested, lyzed and centrifuged and applied to Ni-affinity column as explained in section 4.9. Elution profile from the Ni-Affinity column is analyzed by SDS-PAGE (Figure 5.17). G. anatolicus EF-Tu elutes from the column at about 87.5 mM imidazole as a single protein band. Pooled and concentrated fractions after dialysis of the recombinant *Geobacillus anatolicus* EF-Tu is shown in Figure 5.18. By using Bradford assay, total protein yield was determined to be 6,47 mg (from a 800 ml cell culture) and the protein concentration to be 3,41 mg/ml.

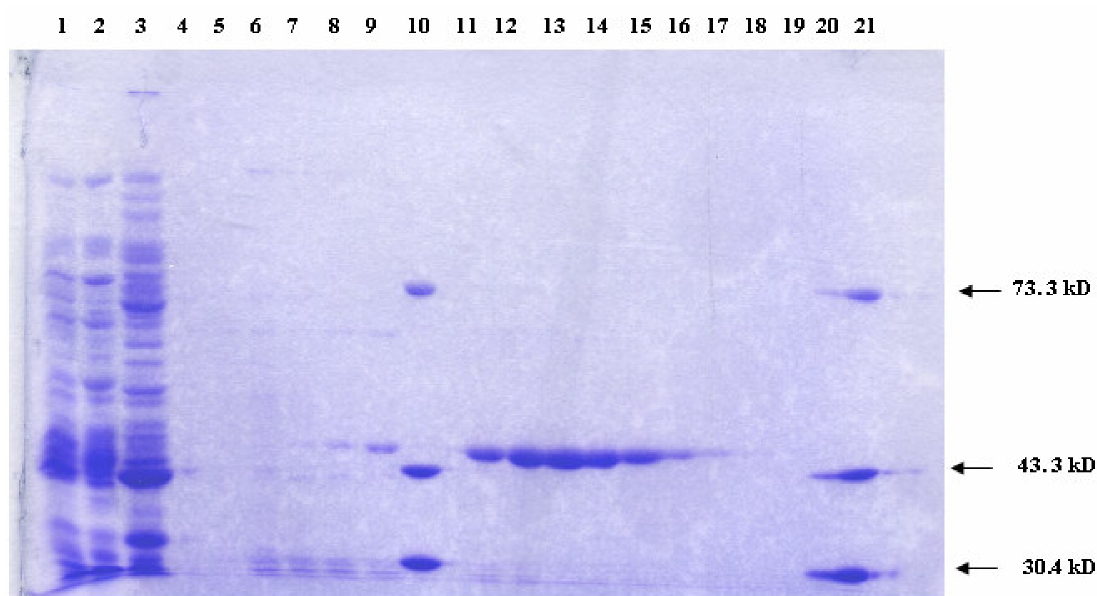

Figure 5.17. SDS analysis of Ni-Affinity choromotography; Lanes 1 and 2: samples before applying the column; lane 3: flowthrough; lanes 4 to 9: NH_4Cl wash; lane 10: protein molecular weight marker (from the top to the bottom, *E. coli* EF-G, EF-Tu and EF-Ts respectively); lanes 11 to 18: eluates with imidazole gradient, lane 20-21, protein molecular weight marker (from the top to the bottom, *E. coli* EF-G, EF-Tu and EF-Ts respectively)

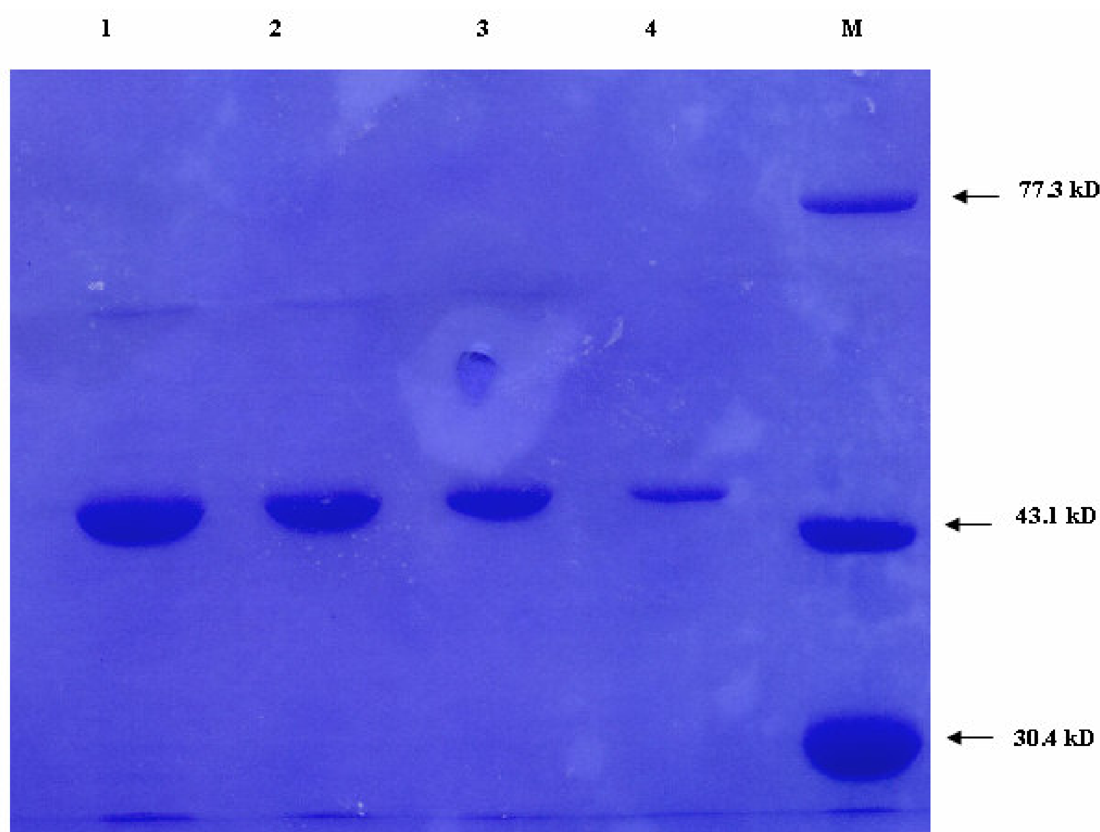

Figure 5.18. *G. Anatolicus* EF-Tu (lanes 1-4); M, protein molecular weight marker (from the top to the bottom, *E. coli* EF-G, EF-Tu and EF-Ts respectively.

5.5. Interactions of *G. anatolicus* EF-Tu with EF-Ts

In a previous study *G. anatolicus* EF-Ts has been cloned and purified (Sahin, 2006, MSci. Thesis). This recombinant *G. anatolicus* EF-Ts as well as *E. coli* EF-Ts is used in a mobility-shift assay on non-denaturing PAGE for complex formation with the purified recombinant *G. anatolicus* EF-Tu. Recombinant *G.anatolicus* EF-Tu is fully competent in complex formation with *G. anatolicus* EF-Ts (Figure 5.16., lanes 2-5) as well as with *E. coli* EF-Ts (Figure 5.19., lanes 8-11). From lane 2 to 5, intensty of the complex band increased. It suggests that with the increasing amount of EF-Tu, EF-Ts binds to more EF-Tu molecules at the subsequent lanes. Mobility of EF-Tu molecule and position of the complex band will be discussed in the Discussion section.

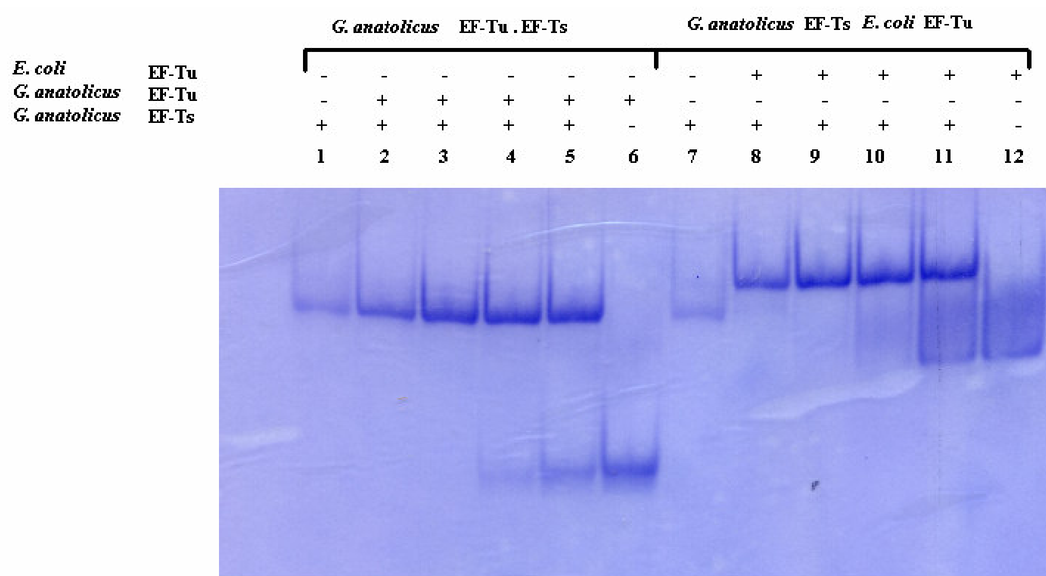

Figure 5.19. Mobility-shift assay for complex formation between EF-Tu and EF-Ts. All lanes contain 300 pmol of *G.anatolicus* EF-Ts except lane 6 and 12. *G.anatolicus* EF-Tu (lanes 1-6) and *E.coli* EF-Tu (lanes 7-12) were titrated (0 pmol, 100 pmol, 200 pmol, 300 pmol, 400 pmol, 400 pmol) , respectively.

5.6. Ternary Complex Formation between G. *anatolicus* EF-Tu, GTP and Phe-tRNA^{Phe}

Mobility shift on a non-denaturing PAGE is used in order to analyze the complex formation between G. *anatolicus* EF-Tu, GTP and Phe-tRNA^{Phe}. In this experiment it was shown that *G.anatolicus* EF-Tu can form a ternary complex with GTP and Phe-tRNA^{Phe} (Figure 5.20.). With the increasing amounts of tRNA^{Phe}, negatively charged tRNA^{Phe} binding shifts the upper EF-Tu band to a high mobility (lower) band corresponding to the ternary complex. Thus results suggest that ternary complex was formed at the bottom band. In the case of *E.coli* EF-Tu, all EF-Tu molecules could participate in ternary complex formation; however it wasn't the case for *G.anatolicus* EF-Tu, some Tu remain free (see Discussion section).

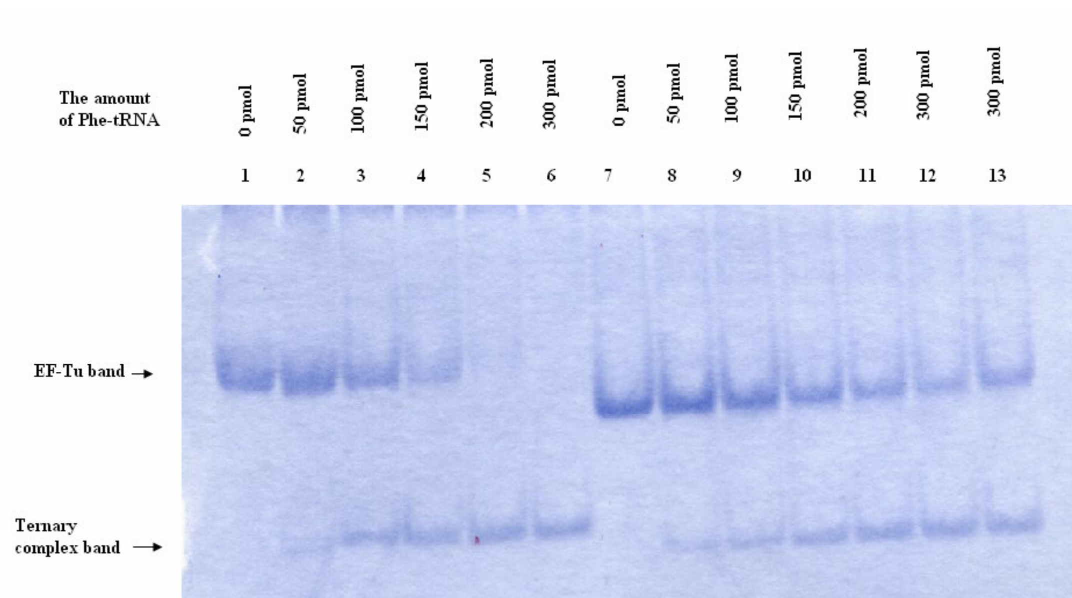

Figure 5.20. Ternary complex formation experiment. In lanes 1-6 there are 200 pmol of *E.coli* EF-Tu and in lanes 7-13 there are 200 pmol of *G.anatolicus* EF-Tu. In lanes 1-6, 0, 50, 100, 150, 200, 300 pmol of $tRNA^{Phe}$ are present respectively. In lanes 7-13, 0, 50, 100, 150, 200, 300 and 300 pmol of $tRNA^{Phe}$ are present respectively. In lane 13, incubation time was increased by 20 min.

5.7. Thermostability Test

Because *G. anatolicus* EF-Tu is a hyperthermophilic protein, it was expected that EF-Tu would be active above mesophilic temperatures. To determine thermostability profile of *G. anatolicus* EF-Tu, a thermostability test was performed based on EF-Tu's ability to bind GDP. *E. coli* EF-Tu can be inactivated significantly when subjected to 60 °C for 10 min, *G. anatolicus* EF-Tu retained 98.7 % of its activity for GDP binding at this temperature. Thermostability profile based on GDP binding ability of *G.anatolicus* EF-Tu was shown in Figure 5.21.

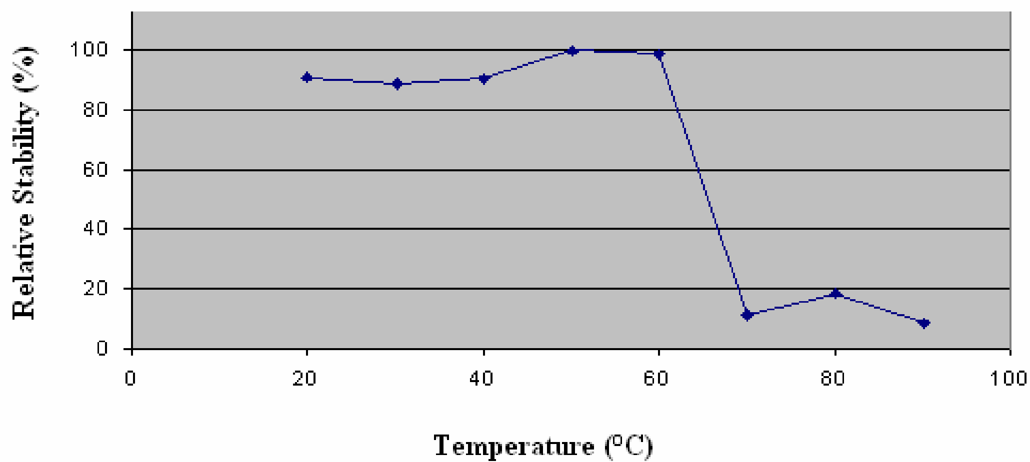

Figure 5.21. Thermostability of *G. anatolicus* EF-Tu as judged by GDP binding. GDP binding ability after incubating for 10 min at temperatures between 20-90 °C in 10 °C intervals. Because it retained the maximum amount of ^{3}H-GDP at 50 °C, this value was taken as 100%.

6. DISCUSSION

6.1. Determination of *Geobacillus anatolicus tuf* Gene Sequence

Because there is no sequence information about *Geobacillus anatolicus* EF-Tu gene (*tuf*) before this study, the species phylogenetically related to *Geobacillus anatolicus* were chosen for the sequence determination. For this purpose, the known 16S rRNA sequence that was previously obtained in our laboratory was used in BLAST search to look up related organisms. Those from which *tuf* genes were previously sequenced were chosen. By using ClustalW multiple alignment technique, the *tuf* gene sequences from these species were aligned and the conserved regions were determined. Based on this information, degenerate primers were designed to amplify a part of the *tuf* gene. After DNA sequencing of this region, partial *tuf* sequence of *Geobacillus anatolicus* was obtained. Bu using this sequence information, now specific for the *Geobacillus anatolicus,* non-degenerate primers were designed for *tuf* gene sequencing. In addition to this, to doublecheck the flanking sequences of the *tuf* gene, two additional degenerate primers were designed within the genes expected to be upstream and downstream of the *tuf* gene. This was possible, because there is a high degree of conservation at and around the tuf gene in the organisms that were chosen from the BLAST search. Using these primers, a fragment of *Geobacillus anatolicus* DNA containing the full *tuf* gene was obtained and sequenced by using six primers. *tuf* gene of *Geobacillus anatolicus* is found to be 1182 bp long. The nucleotide sequence of the *Geobacillus anatolicus tuf* gene was used to determine the EF-Tu amino acid sequence. Based on BLAST results of both the nucleotide and the amino acid sequences in the bacterial GenBank databases, it was confirmed that the amplified fragment, in fact, contained the *tuf* gene, corresponded to *Geobacillus anatolicus* EF-Tu protein.

By the course of this study, the complete genome of *Geobacillus kaustophilus* has been published. The nucleotide sequence comparisons between *Geobacillus anatolicus* and its closest relative *Geobacillus kaustophilus* revealed 93% sequence identity. This result confirmed that *Geobacillus anatolicus* is a new species based on its 16S ribosomal DNA comparisons previously reported (Uysal *et al.,* 2001) (Figure 6.1.). In a previous study in our laboratory (Sahin, 2006, MSci. Thesis) on EF-Ts from *Geobacillus anatolicus*,

it was shown that there is a 95,5% amino acid sequence identity between *Geobacillus anatolicus* and *Geobacillus kaustophilus* EF-Ts. But in the case of EF-Tu amino acid sequences, the identity was 99%. Only two amino acids were found to be different (Figure 6.2.) when EF-Tu sequences compared. EF-Tu has a more conserved aminoacid sequence. This may be because EF-Tu is a more ancient protein. In addition to this, given that *Geobacillus kaustophilus* is a thermophilic organism, this may suggest that overall amino acid sequence of EF-Tu is essential for the thermoadaptation.

```
             ....|....| ....|....| ....|....| ....|....| ....|....|
                 5          15         25         35         45
G. kaustop   ATGGCTAAAG CGAAATTTGA GCGCACGAAA CCGCACGTCA ACATTGGCAC
G. anatoli   ATGGCTAAAG CGAAATTTGA GCGTACGAAA CCGCACGTCA ACATTGGCAC
Consensus    ATGGCTAAAG CGAAATTTGA GCGYACGAAA CCGCACGTCA ACATTGGCAC

             ....|....| ....|....| ....|....| ....|....| ....|....|
                 55         65         75         85         95
G. kaustop   GATCGGCCAC GTTGACCACG GGAAAACGAC GTTGACAGCT GCGATCACGA
G. anatoli   GATCGGCCAC GTTGACCATG GGAAAACGAC GTTGACAGCT GCGATCACAA
Consensus    GATCGGCCAC GTTGACCAYG GGAAAACGAC GTTGACAGCT GCGATCACRA

             ....|....| ....|....| ....|....| ....|....| ....|....|
                105        115        125        135        145
G. kaustop   CGGTTTTGGC GAAACAAGGG AAAGCCGAAG CAAAAGCGTA CGACCAAATC
G. anatoli   CAGTTCTTGC GAAACAAGGG AAAGCGGAAG CAAGAGCGTA CGACCAAATC
Consensus    CRGTTYTKGC GAAACAAGGG AAAGCSGAAG CAARAGCGTA CGACCAAATC

             ....|....| ....|....| ....|....| ....|....| ....|....|
                155        165        175        185        195
G. kaustop   GACGCTGCTC CGGAAGAGCG TGAACGCGGG ATTACGATTT CGACGGCGCA
G. anatoli   GACGCAGCTC CGGAAGAGCG TGAACGCGGA ATCACGATTT CGACAGCGCA
Consensus    GACGCWGCTC CGGAAGAGCG TGAACGCGGR ATYACGATTT CGACRGCGCA

             ....|....| ....|....| ....|....| ....|....| ....|....|
                205        215        225        235        245
G. kaustop   CGTCGAGTAT GAAACAGATG CTCGCCACTA TGCGCACGTT GACTGCCCGG
G. anatoli   CGTTGAGTAT GAAACGGACG CTCGTCACTA TGCGCACGTT GACTGCCCAG
Consensus    CGTYGAGTAT GAAACRGAYG CTCGYCACTA TGCGCACGTT GACTGCCCRG

             ....|....| ....|....| ....|....| ....|....| ....|....|
                255        265        275        285        295
G. kaustop   GCCACGCTGA CTACGTTAAA AACATGATCA CGGGCGCGGC GCAAATGGAC
G. anatoli   GCCACGCTGA CTACGTGAAA AACATGATCA CGGGCGCAGC GCAAATGGAC
Consensus    GCCACGCTGA CTACGTKAAA AACATGATCA CGGGCGCRGC GCAAATGGAC

             ....|....| ....|....| ....|....| ....|....| ....|....|
                305        315        325        335        345
G. kaustop   GGCGCGATCC TTGTTGTATC GGCGGCTGAC GGTCCGATGC CGCAAACGCG
G. anatoli   GGCGCGATCC TTGTTGTATC GGCTGCTGAC GGTCCGATGC CGCAAACGCG
Consensus    GGCGCGATCC TTGTTGTATC GGCKGCTGAC GGTCCGATGC CGCAAACGCG
```

Figure 6.1. Nucleotide sequence comparison between G. *kaustophilus* and G. *anatolicus*. Different bases were shown in red.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                   355        365        375        385        395
G. kaustop      CGAACACATT CTTCTCTCCC GCCAAGTCGG TGTTCCGTAC ATCGTTGTAT
G. anatoli      CGAACACATT CTTCTCTCCC GCCAAGTCGG TGTACCGTAC ATCGTTGTTT
Consensus       CGAACACATT CTTCTCTCCC GCCAAGTCGG TGTWCCGTAC ATCGTTGTWT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   405        415        425        435        445
G. kaustop      TCTTGAACAA ATGCGACATG GTGGACGACG AAGAATTGCT TGAACTTGTT
G. anatoli      TCTTGAACAA ATGCGACATG GTGGACGACG AAGAATTGCT TGAACTCGTT
Consensus       TCTTGAACAA ATGCGACATG GTGGACGACG AAGAATTGCT TGAACTYGTT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   455        465        475        485        495
G. kaustop      GAAATGGAAG TTCGCGATCT TCTCTCTGAA TACGACTTCC CAGGCGATGA
G. anatoli      GAAATGGAAG TTCGCGATCT TCTCTCTGAA TACGACTTCC CGGGCGACGA
Consensus       GAAATGGAAG TTCGCGATCT TCTCTCTGAA TACGACTTCC CRGGCGAYGA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   505        515        525        535        545
G. kaustop      AGTGCCGGTC ATCAAAGGTT CGGCATTAAA AGCGCTTGAA GGCGACCCGC
G. anatoli      AGTGCCGGTC ATCAAAGGTT CGGCATTAAA AGCGCTCGAA GGCGACCCGC
Consensus       AGTGCCGGTC ATCAAAGGTT CGGCATTAAA AGCGCTYGAA GGCGACCCGC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   555        565        575        585        595
G. kaustop      AATGGGAAGA AAAAATCATT GAACTGATGA ATGCGGTTGA CGAATACATC
G. anatoli      AATGGGAAGA AAAAATCATT GAACTGATGA ACGCGGTTGA CGAGTACATC
Consensus       AATGGGAAGA AAAAATCATT GAACTGATGA AYGCGGTTGA CGARTACATC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   605        615        625        635        645
G. kaustop      CCGACGCCGC AACGTGAAGT GGACAAACCG TTCATGATGC CAATCGAGGA
G. anatoli      CCGACTCCGC AACGTGAAGT AGACAAACCG TTCATGATGC CGGTTGAGGA
Consensus       CCGACKCCGC AACGTGAAGT RGACAAACCG TTCATGATGC CRRTYGAGGA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   655        665        675        685        695
G. kaustop      CGTCTTCTCG ATCACGGGCC GCGGTACGGT TGCAACAGGC CGTGTTGAAC
G. anatoli      CGTTTTCTCG ATCACAGGCC GCGGTACGGT TGCGACGGGC CGTGTTGAGC
Consensus       CGTYTTCTCG ATCACRGGCC GCGGTACGGT TGCRACRGGC CGTGTTGARC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   705        715        725        735        745
G. kaustop      GCGGTACGTT GAAAGTTGGC GACCCGGTTG AAATCATCGG TTTGTCGGAT
G. anatoli      GCGGTACGTT AAAAGTCGGC GACCCGGTTG AAATTATCGG TCTTTCGGAC
Consensus       GCGGTACGTT RAAAGTYGGC GACCCGGTTG AAATYATCGG TYTKTCGGAY

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   755        765        775        785        795
G. kaustop      GAGCCGAAAA CGACGACCGT TACGGGTGTA GAAATGTTCC GCAAACTTCT
G. anatoli      GAGCCGAAAA CGACGACGGT TACGGGTGTA GAAATGTTCC GTAAGCTGCT
Consensus       GAGCCGAAAA CGACGACSGT TACGGGTGTA GAAATGTTCC GYAARCTKCT
```

Figure 6.1. Nucleotide sequence comparison between G. *kaustophilus* and G. *anatolicus*. Different bases were shown in red.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                   805        815        825        835        845
G. kaustop      CGATCAAGCG GAAGCCGGCG ACAACATCGG TGCGCTTCTC CGCGGCGTAT
G. anatoli      TGACCAAGCG GAAGCTGGCG ACAACATCGG TGCGCTTCTC CGCGGTGTAT
Consensus       YGAYCAAGCG GAAGCYGGCG ACAACATCGG TGCGCTTCTC CGCGGYGTAT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   855        865        875        885        895
G. kaustop      CTCGTGATGA AGTCGAGCGC GGCCAAGTAT TGGCGAAACC TGGTTCGATC
G. anatoli      CGCGTGACGA AGTTGAGCGC GGCCAAGTAT TGGCGAAACC GGGCTCGATC
Consensus       CKCGTGAYGA AGTYGAGCGC GGCCAAGTAT TGGCGAAACC KGGYTCGATC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   905        915        925        935        945
G. kaustop      ACGCCGCACA CGAAATTCAA AGCGCAAGTA TACGTCCTGA CGAAAGAAGA
G. anatoli      ACGCCGCACA CGAAATTTAA AGCGCAAGTT TACGTTCTGA CGAAAGAAGA
Consensus       ACGCCGCACA CGAAATTYAA AGCGCAAGTW TACGTYCTGA CGAAAGAAGA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   955        965        975        985        995
G. kaustop      AGGCGGACGC CATACTCCGT TCTTCTCGAA CTACCGTCCG CAATTCTACT
G. anatoli      AGGCGGACGC CATACTCCGT TCTTCTCGAA CTACCGTCCG CAATTCTACT
Consensus       AGGCGGACGC CATACTCCGT TCTTCTCGAA CTACCGTCCG CAATTCTACT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   1005       1015       1025       1035       1045
G. kaustop      TCCGTACGAC GGACGTAACG GGCATTATCA CGCTTCCGGA AGGCGTCGAA
G. anatoli      TCCGTACAAC GGACGTAACG GGCATCATCA CGCTTCCGGA AGGCGTAGAA
Consensus       TCCGTACRAC GGACGTAACG GGCATYATCA CGCTTCCGGA AGGCGTMGAA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   1055       1065       1075       1085       1095
G. kaustop      ATGGTTATGC CTGGCGACAA CGTTGAAATG ACGGTGGAAC TGATCGCCCC
G. anatoli      ATGGTTATGC CTGGCGACAA CGTTGAAATG ACGGTTGAAC TGATCGCTCC
Consensus       ATGGTTATGC CTGGCGACAA CGTTGAAATG ACGGTKGAAC TGATCGCYCC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   1105       1115       1125       1135       1145
G. kaustop      GATCGCGATC GAAGAAGGTA CGAAATTCTC GATCCGCGAA GGCGGCCGTA
G. anatoli      GATCGCGATC GAGGAAGGAA CAAAATTCTC GATCCGTGAA GGCGGCCGCA
Consensus       GATCGCGATC GARGAAGGWA CRAAATTCTC GATCCGYGAA GGCGGCCGYA

                ....|....| ....|....| ....|....| ....|...
                   1155       1165       1175       1185
G. kaustop      CGGTCGGCGC TGGCTCTGTA TCGGAAATCA TCGAGTAA
G. anatoli      CGGTTGGTGC TGGTTCCGTA TCGGAAATCA TTGAGTAA
Consensus       CGGTYGGYGC TGGYTCYGTA TCGGAAATCA TYGAGTAA
```

Figure 6.1. Nucleotide sequence comparison between G. *kaustophilus* and G. *anatolicus*. Different bases were shown in red.

```
               ....|....| ....|....| ....|....| ....|....| ....|....|
                   5         15         25         35         45
G. anatoli     MAKAKFERTK PHVNIGTIGH VDHGKTTLTA AITTVLAKQG KAEARAYDQI
G. kaustop     MAKAKFERTK PHVNIGTIGH VDHGKTTLTA AITTVLAKQG KAEAKAYDQI
Consensus      MAKAKFERTK PHVNIGTIGH VDHGKTTLTA AITTVLAKQG KAEA AYDQI

               ....|....| ....|....| ....|....| ....|....| ....|....|
                   55        65         75         85         95
G. anatoli     DAAPEERERG ITISTAHVEY ETDARHYAHV DCPGHADYVK NMITGAAQMD
G. kaustop     DAAPEERERG ITISTAHVEY ETDARHYAHV DCPGHADYVK NMITGAAQMD
Consensus      DAAPEERERG ITISTAHVEY ETDARHYAHV DCPGHADYVK NMITGAAQMD

               ....|....| ....|....| ....|....| ....|....| ....|....|
                  105        115        125        135        145
G. anatoli     GAILVVSAAD GPMPQTREHI LLSRQVGVPY IVVFLNKCDM VDDEELLELV
G. kaustop     GAILVVSAAD GPMPQTREHI LLSRQVGVPY IVVFLNKCDM VDDEELLELV
Consensus      GAILVVSAAD GPMPQTREHI LLSRQVGVPY IVVFLNKCDM VDDEELLELV

               ....|....| ....|....| ....|....| ....|....| ....|....|
                  155        165        175        185        195
G. anatoli     EMEVRDLLSE YDFPGDEVPV IKGSALKALE GDPQWEEKII ELMNAVDEYI
G. kaustop     EMEVRDLLSE YDFPGDEVPV IKGSALKALE GDPQWEEKII ELMNAVDEYI
Consensus      EMEVRDLLSE YDFPGDEVPV IKGSALKALE GDPQWEEKII ELMNAVDEYI

               ....|....| ....|....| ....|....| ....|....| ....|....|
                  205        215        225        235        245
G. anatoli     PTPQREVDKP FMMPVEDVFS ITGRGTVATG RVERGTLKVG DPVEIIGLSD
G. kaustop     PTPQREVDKP FMMPIEDVFS ITGRGTVATG RVERGTLKVG DPVEIIGLSD
Consensus      PTPQREVDKP FMMP EDVFS ITGRGTVATG RVERGTLKVG DPVEIIGLSD

               ....|....| ....|....| ....|....| ....|....| ....|....|
                  255        265        275        285        295
G. anatoli     EPKTTTVTGV EMFRKLLDQA EAGDNIGALL RGVSRDEVER GQVLAKPGSI
G. kaustop     EPKTTTVTGV EMFRKLLDQA EAGDNIGALL RGVSRDEVER GQVLAKPGSI
Consensus      EPKTTTVTGV EMFRKLLDQA EAGDNIGALL RGVSRDEVER GQVLAKPGSI

               ....|....| ....|....| ....|....| ....|....| ....|....|
                  305        315        325        335        345
G. anatoli     TPHTKFKAQV YVLTKEEGGR HTPFFSNYRP QFYFRTTDVT GIITLPEGVE
G. kaustop     TPHTKFKAQV YVLTKEEGGR HTPFFSNYRP QFYFRTTDVT GIITLPEGVE
Consensus      TPHTKFKAQV YVLTKEEGGR HTPFFSNYRP QFYFRTTDVT GIITLPEGVE

               ....|....| ....|....| ....|....| ....|....| ....|
                  355        365        375        385        395
G. anatoli     MVMPGDNVEM TVELIAPIAI EEGTKFSIRE GGRTVGAGSV SEIIE
G. kaustop     MVMPGDNVEM TVELIAPIAI EEGTKFSIRE GGRTVGAGSV SEIIE
Consensus      MVMPGDNVEM TVELIAPIAI EEGTKFSIRE GGRTVGAGSV SEIIE
```

Figure 6.2. Amino acid sequence comparison between G. *kaustophilus* and G. *anatolicus*. Different amino acids were shown in red.

Escherichia coli EF-Tu·EF-Ts crystal structure was determined to a high resolution (Kawashima *et al.,* 1996). According to this structure there are 27 amino acid residues on the EF-Tu interacting with EF-Ts. In the *G. anatolicus* EF-Tu, except one amino acid (In nucleotide binding domain of EF-Tu interacting with N-terminal domain of EF-Ts,

threonine was changed into alanine), 26 amino acids were conserved (Figure 6.3.). It can be inferred that dramatic increase in the growth temperature doesn't cause any effect on the interaction mechanism of EF-Tu and EF-Ts, even though EF-Ts sequence outside this domain is drastically different.

Some aminoacid residues in EF-Tu were shown to have regulatory functions, as determined in the previous studies on the *E. coli* EF-Tu. These residues are subjected to post-translational modifications. For example, during exponential growth, EF-Tu becomes monomethylated at Lys56; this is converted to Me2Lys (dimethyl-Lysine) when the bacteria enter the stationary phase (Kraal *et al.*, 1999). This highly conserved residue is not conserved in the *G. anatolicus* EF-Tu, neither in *G. kaustophilus*. However, the change is from a positively charged Lys to another positively charged residue, Arg. Another site of post-translational modification, the site of phosphorylation, Thr382 was also conserved in *G. anatolicus*. In addition to this, as *E. coli*, *T. thermophilus* EF-Tu was found to be phosphorylated in vivo (Kraal *et al.*, 1999). Phosphorylation of EF-Tu in Thr382 abolishes the binding of EF-Tu to aminoacyl tRNA (Alexander *et al.*, 1995). High degree of conservation of this residue may indicate its essential role in protein synthesis.

```
             ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45
G. anatoli   MAKAKFERTK PHVNIGTIGH VDHGKTTLTA AITTVLAKQG KAEARAYDQI
E. coli      MSKEKFERTK PHVNVGTIGH VDHGKTTLTA AITTVLAKTY GGAARAFDQI
Consensus    M K KFERTK PHVN GTIGH VDHGKTTLTA AITTVLAK       ARA DQI

             ....|....| ....|....| ....|....| ....|....| ....|....|
                 55        65         75         85         95
G. anatoli   DAAPEERERG ITISTAHVEY ETDARHYAHV DCPGHADYVK NMITGAAQMD
E. coli      DNAPEEKARG ITINTSHVEY DTPTRHYAHV DCPGHADYVK NMITGAAQMD
Consensus    D APEE  RG ITI T HVEY  T  RHYAHV DCPGHADYVK NMITGAAQMD

             ....|....| ....|....| ....|....| ....|....| ....|....|
                105        115        125        135        145
G. anatoli   GAILVVSAAD GPMPQTREHI LLSRQVGVPY IVVFLNKCDM VDDEELLELV
E. coli      GAILVVAATD GPMPQTREHI LLGRQVGVPY IIVFLNKCDM VDDEELLELV
Consensus    GAILVV A D GPMPQTREHI LL RQVGVPY I VFLNKCDM VDDEELLELV

             ....|....| ....|....| ....|....| ....|....| ....|....|
                155        165        175        185        195
G. anatoli   EMEVRDLLSE YDFPGDEVPV IKGSALKALE GDPQWEEKII ELMNAVDEYI
E. coli      EMEVRELLSQ YDFPGDDTPI VRGSALKALE GDAEWEAKIL ELAGFLDSYI
Consensus    EMEVR LLS  YDFPGD  P     GSALKALE GD  WE KI  EL    D YI
```

Figure 6.3. Amino acid sequence comparison and EF-Tu interactions sites. Conserved residues were shown in blue. Non-conserved residues were shown in red.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                   205        215        225        235        245
G. anatoli      PTPQREVDKP FMMPVEDVFS ITGRGTVATG RVERGTLKVG DPVEIIGLSD
E. coli         PEPERAIDKP FLLPIEDVFS ISGRGTVVTG RVERGIIKVG EEVEIVGIK-
Consensus       P P R  DKP F   P EDVFS I GRGTV TG RVERG  KVG   VEI G

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   255        265        275        285        295
G. anatoli      EPKTTTVTGV EMFRKLLDQA EAGDNIGALL RGVSRDEVER GQVLAKPGSI
E. coli         ETQKSTCTGV EMFRKLLDEG RAGENVGVLL RGIKREEIER GQVLAKPGTI
Consensus       E    T TGV EMFRKLLD    AG N G LL RG  R E ER GQVLAKPG I

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   305        315        325        335        345
G. anatoli      TPHTKFKAQV YVLTKEEGGR HTPFFSNYRP QFYFRTTDVT GIITLPEGVE
E. coli         KPHTKFESEV YILSKDEGGR HTPFFKGYRP QFYFRTTDVT GTIELPEGVE
Consensus        PHTKF    V Y L K EGGR HTPFF   YRP QFYFRTTDVT G I LPEGVE

                ....|....| ....|....| ....|....| ....|....| ....|
                   355        365        375        385        395
G. anatoli      MVMPGDNVEM TVELIAPIAI EEGTKFSIRE GGRTVGAGSV SEIIE
E. coli         MVMPGDNIKM VVTLIHPIAM DDGLRFAIRE GGRTVGAGVV AKVLS
Consensus       MVMPGDN  M  V LI PIA    G  F IRE GGRTVGAG V
```

Figure 6.3. Amino acid sequence comparison and EF-Tu interactions sites. Conserved residues were shown in blue. Non-conserved residues were shown in red.

6.2. Number of *tuf* genes

Most gram-positive bacteria carry one *tuf* gene on their chromosomes, whereas gram-negative bacteria contain usually two almost identical genes, *tuf*A and *tuf*B for EF-Tu (Krasny *et al.*, 1998). For example, in *E.coli*, it was found that there are two *tuf* genes and they are almost identical but for 13 bases (Jaskunas *et al.*, 1975). Also in *Thermus thermophilus HB8* it was determined that there are two *tuf* genes; these genes are different from each other in 10 nucleotide positions (Satoh *et al.*, 1991). Given that *Geobacillus anatolicus* is a gram-negative bacterium (Caglayan, 2005, MSci. Thesis), and that the complete genome sequence of *G. kaustophilus* reveals a single copy of *tuf* gene, it may be hypothesized that *G. anatolicus* may contain one *tuf* gene. However, it remains to be investigated whether there is one *tuf* gene in *Geobacillus anatolicus* or not.

6.3. Cloning of the *tuf* Gene

In this study, the vector pCR® T7/NT-TOPO® was used to clone *tuf* gene. This vector contains multiple restriction enzyme sites which could be used for PCR cloning reactions. Also there is an enterokinase cleavage site in the vector to discard the 6xHis-tag. However, during this study, it was found that the commercially available enterokinases cleaves *E.coli* EF-Tu protein unspecifically. This made it impossible to use this vector as it is commercially supplied. According to the literature, 6xHis-tag at the C-terminal position of EF-Tu doesn't interfere with the EF-Tu function in *E. coli* (Boon *et al.*, 1992) Because of this, enterokinase cleavage site was removed from the vector and a new vector was constructed named pT7D3 which expresses EF-Tu with a His-tag.

6.4. Expression of the Recombinant Elongation Factor Tu

The expression vector, pT7D3Tu contains a strong T7 RNA polymerase promoter site. The vector was cloned into a special *E. coli* strain, DE3 which expresses T7 RNA polymerase. DE3 is the designation for the lambda DE3 lysogen carrying the gene for T7 RNA polymerase under the control of the *lac*UV5 promoter, inducable by IPTG. Because there is a similar *tuf* gene in *E.coli*, to avoid recombination, a recombination deficient *Escherichia coli* JM109 (DE3) strain was used in the expression experiments.

6.5. Purification of the Recombinant His-tagged Elongation Factor Tu

Metal-ion chelating chromotography was used for the purification of the recombinant 6xHis-tagged elongation factor Tu. By adding 18 nucleotides to the TuR2 primer, 6xHis-tag was inserted into the recombinant *Geobacillus anatolicus* elongation factor Tu. This insertion allowed EF-Tu to be purified by Ni^{2+}-affinity chromatography.

6.6. Mobility-Shift Assay for EF-Tu.EF-Ts complex formation

Because *Geobacillus anatolicus* EF-Ts was available from a previous study (Sahin, 2006, MSci. Thesis), interaction of EF-Tu with EF-Ts could be examined. There is a significant difference between the mobilities of *Geobacillus anatolicus* EF-Tu and *E. coli*

EF-Tu, not explainable with the presence of 6 additional His residues in the recombinant EF-Tu from *Geobacillus anatolicus*. In fact, *Geobacillus anatolicus* EF-Tu has a higher mobility towards the positive electrode.

The mobilities of free *Geobacillus anatolicus* EF-Tu and *Geobacillus anatolicus* EF-Tu.EF-Ts complex were almost identical, in contrast to heterologous complex between *E. coli* EF-Tu and *Geobacillus anatolicus* EF-Ts. This may indicate a significant structural change due to complex formation. In addition to this, titrations suggests that there is a 1:1 stoichiometry between *Geobacillus anatolicus* EF-Tu and EF-Ts.

6.7. Ternary Complex Formation between G. *anatolicus* EF-Tu, GTP and Phe-tRNA^{Phe}

G. *anatolicus* EF-Tu forms a complex with GTP and Phe-tRNA^{Phe} as indicated by mobility shift assay. However, some EF-Tu remained free. During complex formation, samples were incubated at 37 °C for GDP exchange to GTP. Because *Geobacillus anatolicus* Tu is thermophilic, some GDP may not be exchanged to GTP because of lower exchange rate at 37 °C. To overcome this, the last sample was incubated at 37 °C for additional 20 min. No significant change was observed. Some of the recombinant Tu may become inactivated for tRNA binding during the purification process even though fully active in EF-Ts binding. This observation remains to be investigated.

6.8. Thermostability

G. *anatolicus* EF-Tu was fully stable up to 60 °C. After 10 min at 70°C, most of the EF-Tu becomes inactive. Although G. anatolicus is taken from a spring at 98°C, this is unexpected. However, during the purification of the bacteria to homogenious cultures, agar plates were used, because of its convenience for selecting colonies. Because agar melts above 65 °C, G. *anatolicus* was grown at 65 °C throughout this study. Although G. *anatolicus* is a hyperthermophilic organism originally living at 98 °C, a possible adaptation to this temperature might have occurred.

6.9. Conclusion and Future Perspective

Because G. *anatolicus* EF-Ts is also available, G. *anatolicus* EF-Tu can further be studied functionally in its interactions with EF-Ts. Kinetic parameters of the EF-Tu.EF-Ts interaction should be measured.

As a future perspective, this study can be considered as a part of construction of a *in vitro* thermophilic translation system. If such a system can be constructed, this may provide information about thermophilic protein translational machinary. EF-G and ribosomes remain to be purified for this purpose. A thermophilic in vitro translation system may have biotechnological applications as well.

Because G. *anatolicus* is a hyperthermophilic organism, structural and functional investigation of EF-Tu will gain insight into thermophilic proteins.

7. REFERENCES

Abel, K., MD. Yoder, R. Hilgenfeld and F. Jurnak, 1996, "An alpha to beta conformational switch in EF-Tu", Vol. 4, pp. 1153-1159.

Alexander, C., N. Bilgin, C. Lindschau, J.R. Mesters, B. Kraal, R. Hilgenfeld, V.A. Erdmann and C. Lippmann, 1995, "Phosphorylation of EF-Tu prevents ternary complex formation", *J. Biol. Chem.* Vol. 270, pp. 14541-14547.

Andersen, G.R., P. Nissen and J. Nyborg, 2003, "Elongation factors in protein biosynthesis", *TRENDS in Biochemical Sciences,* Vol. 28, pp. 434-441.

Baldauf, S. L., J. D. Palmer and W. F. Doolittle, 1996, "The root of universal tree and the origin of eukaryotes based on elongation factor phylogeny", *Proc. Natl. Acad. Sci.*, Vol. 93, pp. 7749-7754.

Berchtold, H., L. Reshetnikova, CO. Reiser, NK. Schirmer, M. Sprinzl and R. Hilgenfeld, 1993, "Crystal structure of active elongation factor Tu reveals major domain rearrangements", *Nature*, Vol. 365, pp. 126-132.

Boon, K., E. Vijgenboom, L. V. Madsen, A. Talens, B. Kraal and L. Bosch, 1992, "Isolation and functional analysis of histidine-tagged elongation factor Tu", *Eur. J. Biochem.*, Vol. 210, pp. 177-183.

Bradford, M. M., 1976, "Bradford Assay", *Analytical Biochemistry*, Vol. 72, pp. 248-254.

Clark, B.F.C., S. Thirup, M. Kjeldgaard and J. Nyborg, 1999, "Structural information for explaining the molecular mechanism of protein biosynthesis", *FEBS Letters*, Vol. 452, pp. 41-46.

Daniel, R. M. and D. A. Cowan, 2000, "Biomolecular Stability and Life at High Temperatures", *Cell. Mol. Life Sci.*, Vol. 57, pp. 250-264.

Frank, J., 2003, “Toward an understanding of the structural basis of translation”, *Genome Biology*, Vol. 4, pp. 237.1-237.7.

Fukuchi, S. and K. Nishikawa, 2001, “Protein Surface Amino Acid Compositions Distinctively Differ Between Thermophilic and Mesophilic Bacteria”, *Journal of Molecular Biology*, Vol. 309, pp. 835-843.

Gao, N., AV. Zavialov, W. Li, J. Sengupta, M. Valle, RP. Gursky, M. Ehrenberg and J. Frank, “Mechanism for the disassembly of the posttermination complex inferred from cryo-EM studies”, *Mol. Cell.*, Vol. 18, pp. 663-674.

Gouy, M. and R. Grantham, 1980, “Polypeptide Elongation and tRNA Cycling in *Escherichia Coli*: A Dynamic Approach, *FEBS Letters*, Vol.115, pp. 151-155.

Hopfield, J. J., 1974, “Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity”, *Proc. Natl. Acad. Sci. USA*., Vol. 71, pp. 4135-4139.

Huber, H. and O. K. Stetter, 1998, “Hyperthermophiles and Their Possible Potential in Biotechnology”, *Journal of Biotechnology*, Vol. 64, pp. 39-52.

Jaenicke, R. and G. Böhm, 1998, “The Stability of Proteins in Extreme Environments”, *Current Opinion in Structural Biology*, Vol. 8, pp. 738-748.

Jaskunas S.R., L. Lindahl and M. Nomura, 1975, “Identification of two copies of the gene for the elongation factor EF- Tu in E. coli.”, *Nature*, Vol.257, pp. 458-462.

Kawashima, T., C. Berthet-Colominas, M. Wulff, S. Cusack and R. Leberman, 1996, “The structure of the *Escherichia coli* EFTu·EF-Ts complex at 2·5 Å resolution”, *Nature*, Vol. 379, pp. 511-518.

Kisselev, L.L. and R.H. Buckingham, 2000, “Translational termination comes of age”, *Trends Biochem. Sci.*, Vol. 25, pp. 561-566.

Kjeldgaard, M., P. Nissen, S. Thirup and J. Nyborg, 1993, "The crystal structure of elongation factor EF-Tu from Thermus aquaticus in the GTP conformation", *Structure*, Vol. 1, pp. 35-50.

Kjeldgaard, M., J. Nyborg and BF. Clark, 1996, "The GTP binding motif: variations on a theme", *FASEB J.*, Vol. 12, pp. 1347-1368.

Krab, IM. and A. Parmeggiani, 1998, "EF-Tu, a GTPase odyssey", *Biochim. Biophys. Acta*, Vol. 1443, pp. 1-22.

Kraal, B., C. Lippman, and C. Kleanthous, 1999, "Translational regulation by modifications of the elongation factor Tu", *Folia Microbiol. (Praha)*, Vol. 44, pp. 131-141.

Krasny, L., J. R. Mesters, L. N. Tieleman, B. Kral, V. Fucik, R. Hilgenfeld and J. Jonak, 1998, "Structure and Expression of Elongation Factor Tu from *Bacillus stearothermophilus*", *J. Mol. Biol.*, Vol 283, pp. 371-381.

Marintchev, A. and G. Wagner, 2005, "Translation initiation: structures,mechanisms and evolution", *Quarterly Reviews of Biophysics,* Vol. 37, pp. 1-88.

Moazed, D. and HF. Noller, 1989, "Intermediate states in the movement of transfer RNA in the ribosome", *Nature*, Vol. 342, pp. 142-148.

Nazina, T. N., T. P. Tourova, A. B. Poltaraus, E. V. Novikova, A. A. Grigoryan, A. E. Ivanova, A. M. Lysenko, V. V. Petrunyaka, G. A. Osipov, S. S. Belyaev and M. V. Ivanov, 2001, "Taxonomic Study of Aerobic Thermophilic *Bacilli*: Descriptions of *Geobacillus subterraneus* gen. nov., sp. nov. and *Geobacillus uzenensis* sp. nov. from Petroleum Reservoirs and Transfer of *Bacillus stearothermophilus*, *Bacillus thermocatenulatus*, *Bacillus thermoleovorans*, *Bacillus kaustophilus*, *Bacillus thermoglucosidasius* and *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus, G. thermocatenulatus*, *G. thermoleovorans*, *G.*

kaustophilus, G. thermoglucosidasius and *G. thermodenitrificans*", *International Journal of Systematic and Evolutionary Microbiology*, Vol. 51, pp. 433-446.

Nilsson, J. and P. Nissen, 2005, "Elongation factors on the ribosome", *Current Opinion in Structural Biology*, Vol. 15, pp. 349-355.

Nissen, P., J. Hansen, N. Ban, PB. Moore and TA. Steitz, 2000, "The structural basis of ribosome activity in peptide bond synthesis", *Science*, Vol. 289, pp. 920-930.

Polekhina, G., S. Thirup, M. Kjeldgaard, P. Nissen, C. Lippmann and J. Nyborg, 1996, "Helix unwinding in the effector region of elongation factor EF-Tu-GDP", *Structure*, Vol. 4, pp. 1141-1151.

Ramakrishnan, V., 2002, "Ribosome structure and the Mechanism of Translation", *Cell*, Vol. 108, pp. 557-572.

Richarme, G., 1998, "Protein-Disulfide Isomerase Activity of Elongation Factor EF-Tu", *Biochemical and Biophysical Research Communications*, Vol. 252, pp. 156-161.

Rodnina, MV. and W. Wintermeyer, 2003, "Peptide bond formation on the ribosome: structure and mechanism", *Current Opinion in Structural Biology*, Vol. 13, pp.334-340.

Sanderova, H., M. Hulkova, P. Malon, M. Kepkova, and J. Jonak, 2004, "Thermostability of multidomain proteins: Elongation factors EF-Tu from *Escherichia coli* and *Bacillus stearothermophilus* and their chimeric forms", *Protein Science*, Vol. 13, pp. 89-99.

Satoh M., T. Tanaka, A. Kushiro, T. Hakoshima and K. Tomita, 1991, "Molecular cloning, nucleotide sequence and expression of the *tuf*B gene encoding elongation factor Tu from *Thermus thermophilus HB8*", *FEBS Lett.*, Vol. 288, pp. 98-100.

Scandurra, R., V. Consalvi, R. Chiaraluce, L. Politi and P. C. Engel, 1998, “Protein Thermostability in Extremophiles”, *Biochimie*, Vol. 80, pp. 933-941.

Takami, H., Y. Takaki, GJ. Chee, S. Nishi, S. Shimamura, H. Suzuki, S. Matsui and I. Uchiyama, 2004, “Thermoadaptation trait revealed by the genome sequence of thermophilic Geobacillus kaustophilus”, Vol. 32, pp. 6292-6303.

Uysal, H., S. Bakkal, D. Erturk and N. Bilgin, 2001, “*Geobacillus anatolicus* Strain SB 16S Ribosomal RNA Gene, Partial”, (Unpublished) Bacterial GenBank Accession Number : AF411064.

Valle, M., A. Zavialov, W. Li, SM. Stagg, J. Sengupta, RC. Nielsen, P. Nissen, SC. Harvey, M. Ehrenberg and J. Frank, 2003, “Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy”, *Nat. Struct. Biol.*, Vol. 10, pp. 899-906.

Vieille, C. and G. Zeikus, 2001, “Hyperthermophilic Enzymes: Source, Uses, and Molecular Mechanisms for Thermostability”, *Microbiology and Molecular Biology Reviews*, Vol. 65, pp. 1-43.

Young, C.C. and R.W. Bernlohr, 1991, “Elongation Factor Tu is Methylated in Response to Nutrient Deprivation in *Escherichia coli*”, *Journal of Bacteriology*, Vol. 173, pp. 3096-3100.