CLONING, EXPRESSION AND SEQUENCING OF CITRATE SYNTHASE FROM *THERMOPLASMA VOLCANIUM*

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

ÇAĞLAR ÇEKİÇ

IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF BIOTECHNOLOGY

JANUARY 2004

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Canan ÖZGEN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Doç. Dr. Dilek SANİN
Head of Institute

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Mahinur AKKAYA	Prof. Dr. Semra KOCABIYIK
Co-Supervisor	Supervisor

Examining Committee Members

Prof. Dr. Semra KOCABIYIK _____

Prof. Dr. Emel ARINÇ _____

Prof. Dr. Haluk Hamamcı _____

Doç. Dr. Fatih İZGÜ _____

Doç. Dr. Can Bilgin _____

ABSTRACT

CLONING, EXPRESSION AND SEQUENCING OF CITRATE SYNTHASE FROM *THERMOPLASMA VOLCANIUM*

Çekiç, Çağlar

M.S., Department of Biotechnology

Supervisor: Prof. Dr. Semra Kocabıyık

Co-Supervisor: Prof. Dr. Mahinur Akkaya

January, 115 pages

In this study first time, we have cloned and sequenced the citrate synthase gene from a thermoacidophilic archaeon *Thermoplasma (Tp.) volcanium* (Optimum growth temperature of *Tp.volcanium* is 60^{o}C and optimum pH is 2.0.). For cloning we have followed a PCR based approach. Amplification of citrate synthase gene from chromosomal DNA of *Tp.volcanium* yielded a product of 1476 bp containing an open reading frame of 1161 bp comprising the structural gene. After ligation of the PCR amplicon to pDrive vector through AU complementation, recombinant plasmids were transferred into *E.coli* TG-1 competent cells. Out of three recombinants, *E.coli* pDriveCS-31 was selected for further characterization by restriction mapping and DNA sequencing. Southern Blotting and Hybridization using the membrane blot of pDriveCS-31 plasmid and DIG-labeled PCR amplified citrate synthase gene probe, also confirmed the cloning of *Tp.volcanium* citrate synthase gene in *E.coli*. Clustal W Version 1.82 was used for alignment of aminoacid sequence of *Tp.volcanium* citrate synthase with that of other archaeal, bacterial and eukaryotic citrate synthases. The

highest sequence similarity (87%) was found between *Tp.volcanium and Tp.acidophilum* enzymes. Despite low sequence homology (18%) with the pig enzyme, of the 11 residues implicated in catalytic activity of the pig citrate synthase 9 were conserved in the *Tp.volcanium* enzyme.

Heterologous expression of this citrate synthase gene in *E.coli* has been achieved under the control of its promoter sequences. The recombinant enzyme (386 aa) has been purified to homogeneity by affinity chromatography on Reactive Red 120 column. The subunit molecular size was estimated as 43 kDa. The purified enzyme followed classical Michaelis-Menten kinetics. The K_m values of 5.15 μM and 5.60 μM, and V_{max} values of 1.74 μmoles/ml/min and 1.60 μmoles/ml/min were calculated from Lineweaver-Burk plots for acetyl-CoA and oxaloacetate, respectively. The recombinant enzyme was thermostable and retained about 80% of the activity at 85^{o}C after 1 hour.

Keywords: Citrate Synthase, Cloning, Sequencing, Expression, Thermostability, Thermoactivity, *Thermoplasma volcanium.*

ÖZ

THERMOPLASMA VOLCANIUM`UN SİTRAT SENTAZ GENİNİN KLONLANMASI, EKSPRESYONU VE NÜKLEİK ASİT DİZİLİMİNİN BELİRLENMESİ

Çekiç, Çağlar
Master, Biyoteknoloji Bölümü
Tez yöneticisi: Prof. Dr. Semra Kocabıyık
Ortak Tez Yöneticisi: Prof. Dr. Mahinur Akkaya
January, 115 sayfa

Bu çalışmada ilk defa, termoasidofilik bir arkea olan *Thermoplasma (Tp.) volcanium*`dan (Optimum gelişme sıcaklığı 60$^{o}$C, ve optimum pH ise 2.0 dir.) sitrat sentaz geni klonlanarak nükleik asit dizilimi belirlenmiştir. Klonlama için PCR temelli bir yaklaşım izlenmiştir. *Tp.volcanium*`un kromozomal DNA`sından sitrat sentaz geninin amplifikasyonu sonucu 1161 bp büyüklüğünde yapısal gen içeren 1476 bp bir DNA bölgesi elde edilmiştir. PCR amplikonun, AU eşleşmesi ile pDrive vektöre bağlanmasından sonra oluşan rekombinant plazmitler *E.coli* TG-1 competan hücrelerine aktarılmıştır. Üç rekombinant arasından arasından seçilen *E.coli* pDriveCS-31, restriksiyon haritalaması ve DNA dizi belirlenmesi yolu ile karekterize edilmek üzere seçilmiştir. pDriveCS-31 plazmitinin membran blotu ile DIG ile markalanmış PCR ile çoğaltılmış sitrat sentaz gen probunun kullanıldığı Southern Blotlama ve Hibridizasyon, *Tp.volcanium* sitrat sentaz geninin *E.*coli`de klonlandığını bir kez daha doğrulamıştır. Clustal W Version 1.82 programı kullanılarak *Tp.volcanium* sitrat sentazın amino asit dizisi, diğer arkea, bakteri ve

ökaryotik sitrat sentaz dizileri ile eşleştirilmiştir. En yüksek dizi benzerliği (%87) *Tp.volcanium* ve *Tp.acidophilum* enzimleri arasında bulunmuştur. Domuz enzimi ile dizi homolojisinin (18%) düşük olmasına karşın, bu enzimin katalitik aktiveteden sorumlu 11 amino asitinden 9 tanesi *Tp.volcanium* enziminde korunmuştur.

Sitrat sentaz geninin *E.coli*`de ekspresyonu kendi promotörünün kontrolü altında gerçekleştirilmiştir. Rekombinant enzim (386 aa) Reaktive Red 120 kolonu kullanılarak Affinite kromotografi ile homojen bir şekilde saflaştırılmıştır. Herbir alt ünitenin görünür moleküler ağırlığı yaklaşık 43 kDa olarak belirlenmiştir. Saflaştırılmış enzim klasik Michaelis-Menten kinetiğine uymaktadır. Asetil-CoA ve okzaloasetat için çizilen Lineweaver-Burk eğrilerinden sırası ile K_m değerleri 5.15 µM ve 5.60 µM olarak, V_{max} değerleri ise 1.74 µmoles/ml/min ve 1.60 µmoles/ml/min olarak hesaplanmıştır. Recombinant enzim termostabil olup, 85^{o}C de 1 saat süresince aktivitesinin yaklaşık %80`i korunmuştur.

Anahtar Kelimeler: Sitrat sentaz , Klonlama, Ekspresyon, Termostabilite, Termoaktivite, *Thermoplasma volcanium*

Dedicated to my family

ACKNOWLEDGEMENTS

I am very grateful to my supervisor Prof. Dr. Semra Kocabıyık for her endless support, help and patience in this study. Also I am very proud to work in her laboratory where she provides us with continuous guidance and encouragement as well as very professional working environment.

I am also very grateful to Dr. F.İnci Özdemir for her helpful advises, support and friendship. I would like to thank Seda Özdoğan for thoughtfulness of her about sharing lab works, help and very close friendship as well. I also thank my lab mate Hatice Özel.

I would like to thank Graduate School of Natural and Applied Sciences METU for the financial support regarding this thesis project. (Project Number: BAP-2003-07-02-00-41)

Finally, I would like to express my best regards to my family, particularly my parents, Cemalinur and Zafer Çekiç for their endless love, great support, patience and encouragement during the whole period of my study.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

1 Introduction

For the maintenance of life, some enzymatic pathways play key roles. Citric acid cycle is one of those pathways, which mediates not only energy production in aerobic conditions but also production of very important biological compounds such as some kinds of amino acids, fatty acids and sterols, purines and pyrimidines. Some of the enzymes of the pathway show variety of structure, regulation pattern and catalytic properties that can be used to investigate the classification of the source organism. This is especially demonstrated for the first enzyme of the cycle: Citrate synthase (CS) (Danson *et al*., 1985). CS catalyses the first reaction of the cycle: the claisen condensations of the acetyl-coenzyme A (CoA) and oxaloacetate to form citryl-CoA, and hydrolysis of citryl-CoA to yield citric acid and CoA.

$$OAA + AcCoA \longrightarrow \text{citryl-CoA} + H_2O \longrightarrow \text{citrate} + \text{CoA-SH} + H^+$$

As the enzyme presents in organisms belonging to all three kingdoms of life; Eukarya, Bacteria and Archaea (Woese *et al*., 1990), it possess a great deal of evolutionary importance to find out taxonomic status of the organisms and to relate this status with the environmental conditions of the source organism, by comparative studies. Beside the other type of organisms CS is found in all phenotypes of archaea: halophiles, methanogens and

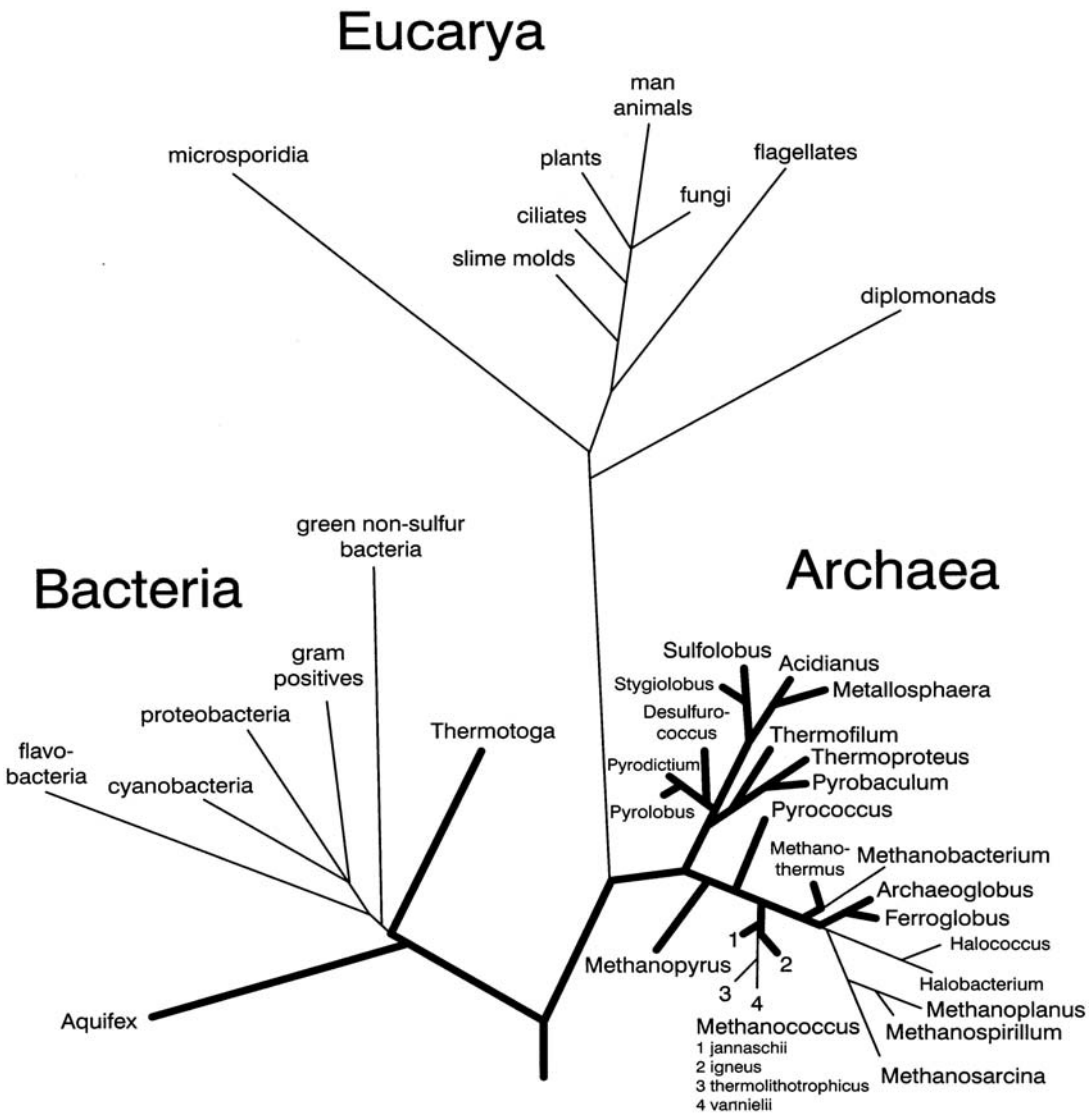


Figure 1.1 Thermophiles within the phylogenetic tree from Woese *et al*.

thermophiles (Figure 1). In this respect, while the taxonomic status and adaptive environmental conditions of the organisms change they start to possess different physiological organizations best mirrored by the mechanistic properties of the enzymes (Szilagyi *et al*., 2000). Understanding of these properties can provide the future design of more stable and catalytically more active enzymes, which can be utilized in many industrial processes. Because the archaea is the kingdom of organisms living in extreme conditions such as high salinity, high temperature, characterizing and cloning the gene of the key enzymes like CS from those organisms is the basic necessity to compare them with their moderate counterparts. To extend the border of this comparative approach we have chosen an acido-thermophilic archaeon, *Thermoplasma volcanium* as source organism for the gene encoding for CS.

1.1 General Physical Properties of the Citrate Synthase:

The enzyme, CS, is present in the organisms as either homodimeric or homohexameric form consisting of polypeptide subunits M. Each monomer has a molecular mass of about 42,000-50,000 Da. In addition, a monomer consists of a large and a small domain. The hexameric form is proposed to operate as a trimer of the basic dimer (Sutherland *et al*., 1990). However, catalytic regulations of these two forms are different. Gram-negative eubacteria generally have homo-hexameric form which is allosterically regulated by NADH and in the facultative anaerobes of this group additionally inhibited by 2-oxoglutarate. Whereas, Gram-positive eubacteria and eukaryotes possess dimeric forms, which are isosterically inhibited by ATP but insensitive to NADH and 2-oxoglutarate (Danson *et al*., 1985). Nevertheless, *in vivo* mutagenesis studies for the enzyme from the Gram negative eubacteria, *E.coli*, revealed that if dimeric form is expressed instead of wild-type hexameric one, the dimeric mutant enzyme is also isosterically regulated by ATP but insensitive to NADH and 2-oxoglutarate. Therefore, the

association manner of the subunits of the active enzyme could affect the regulation pattern for the enzyme (Henneke *et al*., 1989).

Archaeal CSs (Danson *et al*., 1985) are similar to those from Gram-positive eubacteria and eukaryotes. They are all small type and sensitive to ATP as an inhibitor but not to NADH and 2-oxoglutarate. Also it has been found out that the affinity to ATP in archaeal CSs is directly related with the affinity of the enzyme towards Acetyl-CoA. Thus, the form of the enzyme and its regulation pattern are corresponding features for the taxonomical level of the organisms.

1.2 Structural Properties:

The crystal structures exist for CS from a psychrophilic Antartic bacterium *Arthrobacter* strain DS2-3R (optimum growth temperature=31^{o}C) (Bell *et al*., 2002), pig (37^{o}C) (Remington *et al*., 1982), *Thermoplasma acidophilum* (55^{o}C) (Russell *et al*., 1994) (Figure 2), *Sulfolobus solfataricus* (85^{o}C) (Bell *et al*., 2002), and *Pyrococcus furiosus* (100^{o}C) (Russell *et al*., 1997). The most extensive structural (Remington *et al*., 1982), kinetic (Remington 1992) and catalytic data is available for the mesophilic CS enzyme from pig heart citrate synthase. As indicated earlier the enzyme is multimeric and it is composed of identical subunits as monomer. For pig heart citrate synthase, which is a dimeric enzyme, each monomer is almost entirely α-helical. Each peptide for the monomer has two domains: a large domain that consists of about 15 helices and a small domain with five helices. The enzyme is not active when it is in monomeric form as both monomers contribute for active site residues. In other words, for the formation of the catalytic core both subunits are necessary. For this reason, CS is one of the enzymes for its stability and activity subunit interactions are crucial (Kocabıyık *et al*., 1998, Kocabıyık *et al*., 2000). The subunit interface of the enzyme consists of four pair of antiparallel helices: FF', GG', LL', MM' where each pair of the helices is provided by each monomer.

CS from pig heart has been viewed in two forms: an open form and a closed form. This suggests that the enzyme undergoes very large conformational changes, which is equal to approximately 19º rigid body rotation of the small domain with respect to the large one. This change requires very large number of changes in the packing of the side chains in the interior of the small domains (Wiegand *et al*., 1984). It is proposed that the change in the conformation is the serial re-organization of adjacent helices instead of a single rigid body rotation (Lesk *et al*., 1984). As will be discussed later these conformational changes and both forms are necessarily related with the catalytical action of the enzyme. Moreover these general structural principles are conserved along with CSs whose structures determined.

Likewise, archaeal and Gram-positive bacteria CSs have homo-dimeric structures with each monomer consisting of a large and small domain. In addition, they are almost entirely α-helical; containing 16 α-helices. The large domain contains 11 helices (C-M and S) whereas small domain has 5 helices (N-R) in its structure. Moreover, in large domain there is a three antiparallel-stranded β-sheet structure that faces each outer surface of the dimeric enzyme. The dimer interface of all the CSs is made up of two parts and comprises residues solely in the large domain; the main part is the eight α-helical sandwich of four antiparallel pairs of helices (F, G, M and L) with the second being the additional interaction of N- and C-terminal regions. When the enzyme undergoes dimerization again same 8 antiparallel α helices FF`, GG`, LL `, MM` interact to form subunit interface (Russell *et al*., 1994).

The pig CS is different from other four CSs in terms of the topology of the C-terminal region. In the other four, the C-terminal arm of one monomer, elapsing the two together and results in more extensive interactions, including those with the N-terminus (Russell *et al*., 1997). Sequence identities between the various CSs for which 3D-structures have been

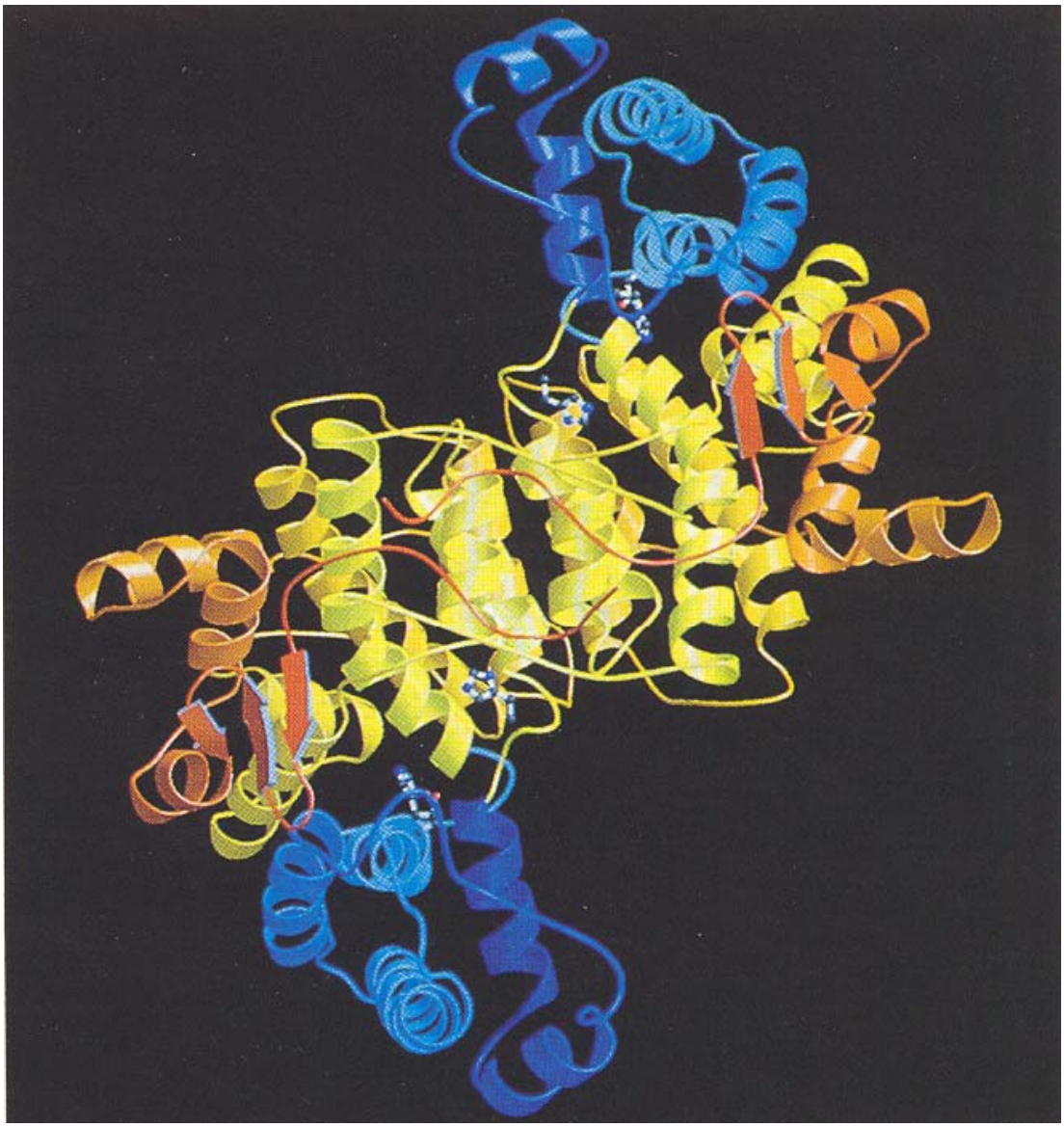


Figure 1.2. Structure of homodimeric *Tp.acidophilum* citrate synthase in open form. The large domain is shown in yellow and small domain in blue. (Russell *et al*.,1994)

determined range from 20% (eucaryotic vs. bacterial or archaeal) to 60% (*S.solfataricus* CS and *T.acidophilum* CS). For the whole dimers, in general the small domains tend to be more highly conserved, e.g. sequence similarities between *Pf*CS and *Ta*CS large and small domains are 39% and 53%, respectively. This result may reflect the fact that large domain is involved in thermal stability and majority of the substrate-binding and catalytic residues are from small domain (Bell *et al*., 2002). As will be discussed in comparative studies, *Pf*CS contains two additional small helices at C- and N-terminal ends (Russell *et al*., 1997).

1.3 Mechanism of Catalysis

As indicated earlier citrate synthase catalyses a claisen condensation of two carbon atoms to form citryl-CoA as intermediate and hydrolysis of this compound yields citrate and CoA. It is one of the few enzymes that do not require any metal cofactors to form covalent bonds between two carbon atoms.

Early studies (Srere, 1966) indicated that oxaloacetate binding to enzyme changed the UV absorption spectrum of it and also binding of OAA increases the affinity of the enzyme towards Acetyl-CoA. These studies suggested that upon binding of the substrate to the enzyme there occurs a conformational change in the enzyme, which triggers the catalytic action of it. The X-ray structures of citrate synthases fall into two conformational classes, open and closed forms. Open and closed refer to the accessibility of the active site to bulk solvent. The open conformation is believed to allow association of substrates and dissociation of products. All chemical transformations including both condensation and hydrolysis are believed to take place in closed forms in which the active site is buried deep within the protein (Remington *et al*., 1982, Russell *et al*., 1994). Therefore, in the presence of OAA and carboxymethyl-CoA (a transition state analog) the

crystals in closed form were obtained; however, in the presence of CoA, Acetyl-CoA and its analogs both open and closed forms of the crystals of the enzyme were obtained (Remington *et al*., 1982, Remington 1992). As a result, the enzyme demonstrates Emil Fischer's "induced fit model" for its catalytic activity.

Catalytic strategies are the same for pig CS and *Ta*CS, as reported by Remington (1992) and Kurtz *et al* (2000). These include carbonyl polarization, the generation of a nucleophilic intermediate, transition state from methyl group of AcCoA, and the sequence of macromolecular conformational changes between several open and closed forms that are integral steps of the catalytic cycle.

Carbonyl polarization is a general strategy used by many enzymes that catalyze reactions involving nucleophile attack on an electrophilic carbonyl carbon. By increasing the positive charge at the reaction center, these enzymes increase the reactivity for condensation with the nucleophile derived from acetyl-CoA in the case of citrate synthase. The nucleophile in pig and *Thermoplasma* CS reaction is derived from the methyl group of AcCoA. Both enzymes catalyze proton transfer from the methyl group to generate an enolate/carbonion intermediate or transition state. The reaction begins with the enzymatic recognition of OAA. For the recognition most susceptible amino acid residues are Arginine 329, 401 and 421, as well as histidine 238 and 320, where the numbers are indicating the positions of the amino acids in primary structure of pig CS. Remington (1992) suggests that histidines are to be charged providing five positively charged residues around the OAA such that third carbon of OAA could be prepared for a nucleophilic attack (carbonyl polarization). *Ta*CS is able to polarize the OAA carbonyl as well as pig CS. The second step is the association of the acetyl-CoA with the enzyme. The key amino acids for this event are assumed to be Arg-164', Arg-324 and Arg-64 which form ligands for three phosphate on CoA. Furthermore there could be some H-bonding with adenine ring of the CoA

and non-specific hydrophobic contacts with the nonpolar tail of the CoA. Therefore, CoA acts only as complement for the conformational change rather than involving the catalytic mechanism of the enzyme. When both substrates are bound to the enzyme the conformational shift of the enzyme from open form to closed form buries the substrates from the solvent and starts the catalysis.

The reaction mechanism (Figure 1.3) for the pig CS is described, in brief, as follows: the catalytic reaction begins with the deprotonation of the methyl group of acetyl-CoA by Asp-375 and protonation of third carbon of OAA by His-320. Deprotonation of methyl group yields enolate so this step is called enolization. The enolization of the acetyl-CoA makes the oxygenated

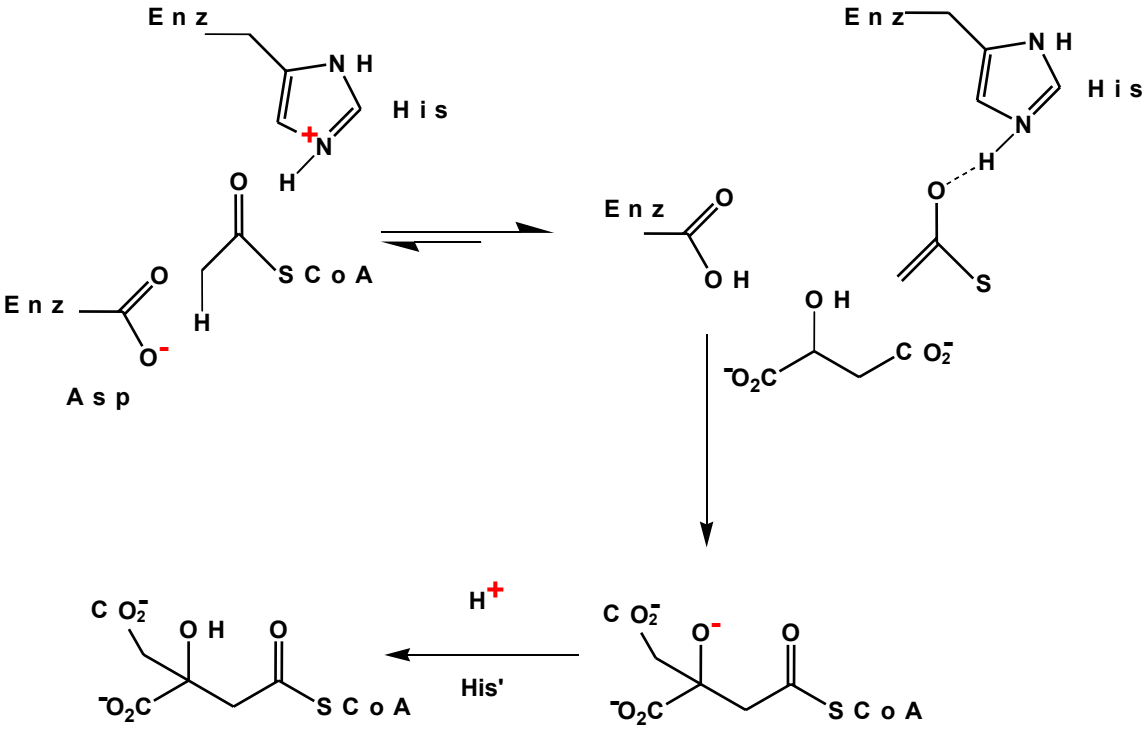


Figure 1.3 Catalytic mechanism of Citrate Synthase from Remington, 1992**.**

carbon atom carrying more partial negative charge and more susceptible for a nucleophilic attack. Reversibly, destabilization of OAA by the addition of a proton makes this substrate a target for a nucleophilic attack. Remington (1992) asserts that the forming enolate is subsequently converted to enol by protonation from its oxygen by His-274. However, Mulholland and Richards (1998) suggests that enolate seems to be the more nucleophilic intermediate than enol. When two substrates are destabilized, the enolate or enol intermediate condenses with the OAA *via* a nucleophilic attack. If enol is the intermediate it is supposed that His-320 takes its proton back from the oxygen. For enolate being the intermediate, there is no requirement for the His-274 to take action in catalysis except for stabilizing the substrates. As a result, citryl-CoA forms as an intermediate which then be hydrolyzed to citrate and CoA. The most unknown step of the reaction is this last hydrolysis of citryl-CoA. However, it is assumed that Asp-375 and an activated water molecule take role in this step (Remington 1992, Mulholland *et al*., 1998). *In vitro* mutagenesis experiments (Alter *et al*., 1990) demonstrated that replacement of His-274 decreases the enzyme activity 1000 times whereas replacement of Asp-375 by any amino acids other than glutamic acid results in no enzyme activity. Therefore, at least the inevitable necessity for this amino acid became certain. These three catalytically important amino acids, His274, His 320 and Asp375 in pig heart citrate synthase have structural counterparts; His222, His262 and Asp317 in *T.acidophilum*, His223, His262 and Asp312 in *P.furiosus*, and His 218, His 258 and Asp313 in *S.solfataricus*, respectively (Russell *et al*., 1994, Russell *et al*., 1997, Bell *et al*., 2002).

Crystallographic studies (Bell *et al*., 2002, Remington *et al*., 1982, Russell *et al*., 1994, Russell *et al*., 1997) revealed that one conformation, i.e. closed conformation of the enzyme catalyses the overall reaction after substrate binding although each step of the condensation reaction very different and there are only small movements of atoms, especially the ones belonging to substrates (Remington, 1992).

1.4 Thermo-stability: Comparative Studies on Structure and Sequence Identity

Comparative studies on the structure and sequence identity of the enzymes provide an excellent source to understand the evolutionary relations of the organisms, catalytic mechanism of the given reaction, and the basis of the stability of the enzymes in extreme environmental conditions. Sometimes, evolutionary positions of the organisms remain to be identified by some molecular approaches rather then general physiological and structural characteristics. These comparative studies become major scopes as molecular tools to make the taxonomic status of the source organism obvious.

Comprehending the catalytical mechanisms of the enzymes and their molecular basis for stability and activity in different environmental conditions is the most important tool for the future enzyme design. To make an enzyme more thermo-active and thermo-resistant many structural and sequential information is needed about thermo-resistant and thermo-active counterparts of the enzyme as well as mesophilic ones. Hence, comparison of a number of families containing homologous proteins from hyperthermophilic, thermophilic and mesophilic organisms is a powerful method to understand the molecular sequence/structural basis of protein thermostability. Furthermore, such comparisons may be useful for the rational design of thermostable proteins and in interpreting directed-evolution experiments (Kumar & Nussinow, 2001). Traditionally, such studies have been hampered by insufficient sequence and structural data. However, this situation is rapidly changing, owing to the availability of complete genome sequences of thermophilic and hyperthermophilic and structural genomic efforts. Such comprehensive analysis of sequence and structural parameters, revealed some general strategies adopted by thermophilic proteins to achieve stability and activity at high temperatures (Sailágyi *et al*., 2000, Böhm *et al*., 1998, Kumar & Nussinow, 2001). Moreover, recent studies demonstrate that

proteins from extreme thermophiles (T_{opt} ≈100^{o}C) show different stabilization patterns in comparison with those from moderately thermophilic organisms (T_{opt} ≈45-80^{o}C) (Sailágyi *et al*., 2000). According to these studies, the main difference between mesophiles and thermopiles is increasing number of ion pairs. In thermophilic ones there is a significant increase in the weak ion pairs while in extremely thermophilic ones additional strong ion pairs are observed. Complex ion-pairs are especially likely to enhance stability through a cooperative strengthening mechanism. Also, at higher temperatures, salt bridges and hydrophobic interactions might make positive contributions to protein stability. Especially, in extremely thermophilic ones there is a significant decrease in the numbers and sizes of the cavities as compared to mesophilic ones. Instead, in the thermophilic proteins an important increase in the polarity of exposed surface area, which is buried upon oligomerization, is observed. In thermophilic organisms there is an apparent increase in the fraction of α helices while in extremely thermophilic proteins it is replaced by an increase in ß strands. In addition, both thermophilic and extremely thermophilic proteins possess a decreased number of irregular regions compared to mesophilic proteins. Nevertheless, there are no generalized differences with respect to hydrogen bonding and potential number of H-bond donor and acceptors. When the amino acid compositions are taken into consideration there are also some significant differences between three types of proteins. In thermophilic proteins there is a replacement of arginine for lysine, which might impose an expected stabilizing effect. However, in the extremely thermophilic proteins the requirement for favorable electrostatic interactions leads to increase in the lysine content besides other charged residues including arginine, glutamic acid and aspargine. Among four heat labile aminoacids: aspargine, cysteine, glycine, and methionine; methionine and aspargine contents in extremely thermophilic organisms are considerably low which can be explained by the chemical instability of these residues at high temperatures. But serine content appears to be low in thermophilic proteins as well as their mesophilic homologues. Recent statistical analysis, based on a non-redundant data set comprising all high quality structures of

thermophilic proteins and their mesophilic homologues from Protein Data Bank demonstrated the expected decrease in glycine content observed but an unexpected increase in the moderately thermophilic proteins (Szilagyi *et al*., 2000).

Although these general comparative studies are very useful to find out generalized rule for the thermostability, it suffers from availability of low amount of 3-D structural information about the proteins and unique properties of some proteins that is not shared by the other protein structures. For example, some proteins are active in their monomeric forms but some others such as citrate synthase are multimeric proteins and are not active without association of the monomeric subunits. Thus, citrate synthase displays an outstanding character to be used as a model enzyme for analysis of structural basis of thermostability since subunit interactions are crucial for the maintenance of thermo-stability as well as thermo-activity. Comparison of the crystal structures of CSs from psychrophilic (*Arthrobacter* strain DS2-3R), mesophilic (pig heart), thermophilic (*Tp.acidophilum, S.solfataricus*) and hyperthermophilic (*P.furiosus*) have allowed a detailed structural analysis to be performed to investigate the structural mechanisms underlying protein thermostability in this enzyme (Remington *et al*., 1982, Remington 1992, Russell *et al*., 1994, Russell *et al*., 1997, Connaris *et al*., 1998, Kocabıyık & Erduran, 1998, Kocabıyık & Erduran, 2000, Russell & Taylor, 1995, Bell *et al*., 2002). One of the identified structural differences is the general compactness of the thermostable enzymes. Thermostable CSs achieve this by shortening of their loop regions, thus having a smaller accessible surface area and volume, and optimized packaging of the side chains in the interior of the proteins and reducing the numbers and the sizes of internal cavities. The loops present in pig citrate synthase may act as weak points during thermal unfolding owing to their relatively high flexibility. Because the loops in archaeal CSs are shorter they may afford this enzyme some form of resistance to thermal unfolding. Thermolabile residues such as Asn, Gln and Lys are generally avoided in thermostable enzymes. The archaeal CSs have

a lower number of these residues compared with pig CS as revealed by sequence analysis. The effect of their replacement is 2-fold: first elimination of potential points of deamination or oxidation and, second, introduction of new stabilizing interactions (Russell *et al*., 1997).

The second strategy employed by the CS to withstand high temperature without losing its activity is the more rigid assembly of the monomer subunits. Thermophilic *Thermoplasma acidophilum* CS accomplishes this stronger intersubunit association by increasing hydrophobicity at helices: FF', GG', LL', and MM' forming subunit interface. According to the site directed mutagenesis experiment results Glycine-209 (present in the helix M) substitution with more hydrophobic Alanine residue an obvious increase was observed in both thermal activity and thermostability in *Ta*CS (Kocabıyık and Erduran 1998, Kocabıyık and Erduran 2000). In extremely thermophilic *Pyrococcus* instead, an increase in ion-pair networks at the subunit interface is observed. In addition, in *P.furiosus*, C-terminals of each monomer wrap around each other. In *Sulfolobus*, which has growth optimum between that *Tp.acidophilum* and *P.furiosus*, shows increased hydrophobic and ionic interactions as compared to *Ta*CS, in the central helices (G and M) of the eight-helical sandwich. These results may be indicative of the increasing strength of the hydrophobic interaction with the temperature, at least to temperatures approaching 100^{o}C, and this may be compensated by the more extensive ionic interactions in the hyperthermophilic protein (Connaris *et al*., 1998).

The third and the last important structural modification in CSs which are more thermoresistant is the reduction in the content of thermolabile aminoacid residues: Asn, Gln, and Cys and Asp (Connaris *et al*., 1998, Russell and Taylor, 1995). As discussed hyperthermophilic proteins still contain significant numbers of thermolabile amino acids but there is evidence that these residues are protected from thermal destruction by folded conformation of the proteins (Connaris *et al*., 1998)

The comparative studies also show that although changes occur in the structure of the enzyme while thermostability of CS increases, the over all structure of the enzyme is conserved demonstrating the importance of structure-function relations. Moreover, these studies reveal that most of the catalytically important aminoacid residues proposed for the pig heart CS are conserved along with the all sequenced and structurally defined CSs although there are very low sequential homologies between CSs from organisms spanning a wide–range of growth temperatures. This can infer the importance of these residues in catalytic action of the enzyme.

1.5 Recombinant Archaeal Citrate Synthases

Thermoplasma acidophilum Citrate Synthase (*Ta*CS): has been cloned and expressed by Sutherland *et al.* (1990) in *E.coli*. In this study *Thermoplasma acidophilum* genomic DNA was digested with *Bam*HI/*Pst*I. The digestion fragments hybridized with radiolabelled probe that was designed according to the N-terminal amino-acid sequence of the enzyme, were selected by Southern blot analysis. Only the signal producing 1.6 kb fragment was identified and purified, and than it was ligated onto pUC19 vector. Transformant colonies were screened with the 48-mer probe and of these one producing positive signal was designated as pTaCS1. The sequencing showed that this clone encoded first 212 aminoacids of the *Ta*CS and *Pst*I site was at 3' end of the gene. A *Sac*I site was determined 160bp downstream of the initiation codon. 3' portion of the gene was identified on Southern blots of *Kpn*I/*Sac*I digested genomic DNA by probing with 16-mer radiolabelled oligonucleotide corresponding the region of the cloned fragment between *Sac*I and *Pst*I sites. This 2.7-kb fragment was cloned into Bluescript vector, recombinant plasmids were identified using 16-mer probe. Sequencing of one positive clone, designated pTaCS2, confirmed that it contained the portion of *Ta*CS gene downstream of the known *Sac*I site. After two subclonings the complementary restriction sites were generated in

the cloned fragments. Then a construct designated pTaCS19 containing the complete *Ta*CS gene was assembled. *Ta*CS cloned gene codes for a protein of 384 aminoacids. Subunit molecular sizes of the purified native and recombinant *Ta*CS proteins were found to be 43.000 (±2000) Da. The dimeric protein has Mr value of 85.000 Da. The recombinant enzyme had similar kinetic properties to that of purified native enzyme. K_m values of the native enzyme for OAA and Acetyl-CoA were 4,5μM (± 0.5) and 6,3μM (± 0.8), respectively. Similarly the plasmid encoded *Ta*CS had a Km values 3.9μM (± 0.8) and 7.7μM (± 1.0) for OAA and Acetyl-CoA, respectively. The coding region of *Ta*CS contains a G + C content of 51%, whereas the genome contains 46% G + C content. Sequence alignment demonstrated that *Ta*CS was more closely related to eubacterial sequence (identity 27,5%) than to the eukaryotic sequence (identity 19,7%).

Sulfolobus solfataricus Citrate Synthase (*Ss*CS) gene (Connaris *et al*., 1998) was cloned in *E.coli* by λEMBL3 page vector. For cloning by the sequence information of N-terminal aminoacid sequence and a conserved region of all known CSs degenerate PCR primers were designed and PCR product of 800bp was obtained, which was confirmed to be a partial CS gene by sequencing and alignment with other known CSs. These PCR fragments were radioactively labeled and used as probe to screen a λEMBL3 library constructed from partially digested *S.solfataricus* genomic DNA. Among the positive recombinants, one plaque designated λA2a was found to contain the complete *S.solfataricus* CS gene on an 11-kb fragment of which the sequence was confirmed. For expression, the gene was amplified from λ clone by PCR incorporating restriction sites to facilitate the cloning into the expression vector. The 1.13 kb gene fragments were then ligated into expression vector pREC7*Nde*I. This expression vector accommodates *rec*A promoter whose activity is induced by nalidixic acid to obtain high-level expression. The gene encodes for a protein of 378 amino acids, with an Mr of 42.679. The SDS/PAGE results for plasmid encoded *Ss*Cs revealed that the

recombinant enzyme has Mr value of 41.000 Da. The calculated G + C content of the coding sequence was 38%, which is almost the same as G + C content of *Sulfolobus* genome. The Km values of native enzyme for oxaloacetate and acetyl-CoA are 4,7μM (± 1.2) and 2.1μM (± 0.1), respectively, which are very similar to the values of the recombinant citrate synthase [5.1μM (± 1.4) and 2.3μM (± 0,2)]. Moreover both native and recombinant enzymes were equally resistant to thermal inactivation as both with half-lives of 8 min at 95^{o}C.

Pyrococcus furiosus Citrate Synthase (*Pf*CS) gene was cloned (Muir *et al*., 1995) in two parts by using a PCR based approach. By the aid of previously determined aminoacid sequences of the enzyme and by choosing a highly conserved sequence in all citrate synthase sequences redundant sense and antisense primers were designed. By using these primers first part of the CS gene, a 760 bp fragment was amplified and cloned into pUC18 vector, and than sequenced. The cloned fragment was labeled and used as probe in Southern blots of digested chromosomal DNA, to determine the location of citrate synthase gene. The coding region was mapped to a 1.6 kb fragment of *Nde*I digested genomic DNA. Then the purified 1.6 kb fragments were ligated by T4 DNA ligase and used as template for inverse PCR. Thus, the region of known sequence including the C-terminus and upstream noncoding region of the gene were amplified. The product was cloned into pUC18 vector and the undetermined regions were sequenced. After the gene was cloned it was transferred to expression vector pKK223-3 with *tac* promoter, vector encoded ribosome binding site and transcription termination signal. To this end, reverse and forward PCR primers were designed to amplify only the coding region. To ensure insertion orientation, suitable restriction sites were introduced *via* engineered primers. The recombinant expression vector was then transformed into competent *E.coli* Jm105 and successfully expressed in this host system. A 1.3 kb region of the cloned gene sequence was found to code for a protein of 376 amino acid (Mr=40.800 Da). The subunit Mr of 42,600 Da was determined by SDS-

PAGE. The Mr of the native enzyme was determined to be 89.700 Da by gel filtration, which confirms the dimeric structure of *Pf*CS like all other archaeal citrate synthases.

Another cloned archaeal citrate synthase gene is from a halophilic, i.e. salt resistant source organism; *Haloferax volcanii* (Connaris *et al*., 1999). The gene encoding *Haloferax volcanii* citrate synthase (*Hv*CS) gene was amplified by PCR from a clone isolated from an λEMBL3 *H.volcanii* genomic library. The amplified fragment of 1,13kb was inserted into two type expression vectors, pET3a for cytoplasmic expression, and pET20b (+) and pET22b (+) for periplasmic expression. All constructs were used to transform *E.coli* BL21 cells for expression. Over expression of the recombinant CS gene was achieved by growing cells in the presence of carbenicillin and through the induction by adding IPTG. Cytoplasmic *Hv*CS was predominantly expressed as inclusion bodies. However, yields of soluble protein were increased when induction temperature was reduced from 37^{o}C to 25^{o}C, although soluble protein is still inactive. The attempts to express the CS gene using periplasmic vectors were unsuccessful. After the recombinant citrate synthase of pET 3a construct was purified it was partially characterized and Mr was found to be 42.000 Da by SDS-PAGE. The recombinant and native enzyme demonstrated similar kinetic characteristics. The Km values for native enzyme were calculated to be 26 μM and 125 μM for OAA and Acetyl-CoA, respectively. The Km values for recombinant citrate synthase were 27μM and 123μM for OAA and Ac-CoA, respectively. In addition, the recombinant enzyme has a V_{max} value of 36 U/mg (± 4.2) that is also very similar to that of the native enzyme (Connaris *et al*., 1999).

1.6 Application Potentials

Citrate synthase is also industrially important enzyme since catalyses the product of citric acid (2-hydroxy-1,2,3-propanetricarboxylic acid), which is

widely used organic acid in foods, beverages and pharmaceuticals. Hence, it has a wide buffering range in aqueous solutions. It is used in many industrial applications for chelation, pH adjustment and derivatization. Aqueous solutions of citric acid are an excellent buffer when partially neutralized since citric acid is a weak acid and it has three pKa's (pKa_1=3.15, pKa_2=4.77 and pKa_3=6.39). Applications include laundry detergents, shampoos, cosmetics, enhanced oil recovery, and chemical cleaning (Crolla and Kennedy, 2001).

The most common production of citric acid is by submerged fermentation using *Aspergillus niger* and a carbohydrate feedstock. Yeasts may also be potential sources for citric acid production. Particularly several Candida sp. have the ability to utilize hydrocarbons for production of citric acid (Kirimura et al., 1999, Crolla and Kennedy, 2001).

Another application for citric acid is bioleaching process. Different groups of microorganisms have been used to extract metals, such as magnesium, zinc, copper, aluminum and lead (II) from aqueous solutions and silicates (Castro *et al*., 2000, Jianlong *et al*., 2000). This is due to the fact that these organisms produce organic acids like citric acid. The solubilization of silicon (to extract Ca, Mg and Zn) is the result of the direct action of the acid or occurs through the complexation of the metals (e.g. aluminum-silicates).

1.7 *Thermoplasma volcanium*

The source organism in this study is *Tp.volcanium* which is a thermo-acidophilic archaeon with an optimum growth temperature of 60ºC and optimum pH of 2. The organism grows over the temperature range between 33ºC and 67ºC and pH range between 1.0 and 4.0 (Segerer *et al*., 1988). It possesses no cell wall so it is gram-negative. Cell membrane mainly composed of ether linked C_{40} biphytantyl diglycerol tetraethers along with small amounts of C_{20} phytantyl glycerol tetraethers instead of ether-linked

lipids, which is a characteristic of archaea. It can grow both anaerobic and aerobic conditions so it is a facultative aerobe. As excessive aeration inhibits growth the organism is also considered as microaerophilic. All *Thermoplasma* species can grow under strictly anaerobic conditions by sulfur respiration. It is an obligate heterotroph. Under optimal conditions the organism has a generation time of about 5 hours (Reysenbach, 2001).

It is isolated form submarine and continental solfataras at Vulcano, Italy (so named as volcanium: a Roman god of fire); continental solfataras and a tropical swam in Java, and continental solfataras in Iceland and the Yellowstone National Park.

The complete genome of the organism has been identified and released to genomic databases (Kawashima *et al*., 2000). The complete genome consists of about 1,584,799 bases with a G-C content of about 38%.

1.8 Scope and Aim

In this study we have aimed at cloning, expression and sequencing of the citrate synthase gene from thermoacidophilic archaeon *Thermoplasma volcanium* (*Tv*). Comparative analyses of structures of model enzymes with varying degrees of thermostabilities provide valuable insights into molecular basis of thermal resistance and mechanistic principles for engineering enzyme stability which is of great significance in various biotechnological applications. On the other hand, overexpression, purification and characterization of the model thermostable enzyme *Tv* citrate synthase, should have biochemical relevance as an example of extremozyme research.

CHAPTER 2

MATERIALS AND METHODS

2.1Materials

2.1.1 Bacterial Strains

Thermoplasma volcanium GSS1 type strain 4299 purchased from DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany was used throughout the study to isolate genomic DNA for the cloning of citrate synthase gene (*Tv*CS).

In transformation experiments the *E.coli* strain TG1 (*supE hsd*Δ5 *thi*Δ(*lac-proAB*) F[*traD*36 *proAB*$^{+}$ *lacI*^{q}*lac Z*ΔM15]) from our laboratory collection was employed.

E.coli pBSK(2CS) strain carrying the phagemid with the *Tp.acidophilum* cloned citrate synthase (*Ta*CS) gene which was constructed by Kocabıyık *et al.* (1996) was used to prepare *Ta*CS gene probe.

2.1.2 Plasmids, Molecular Size Markers

Plasmids used throughout the study are shown in Appendix A. Molecular size markers are given in Appendix B.

2.1.3 Chemicals, Enzymes, Kits

All chemicals used were analytical grade. Acrylamide, Ammonium Sulfate $(NH_4)_2SO_4$, Ammonium per-sulfate, Ampicillin, Calcium Chloride dehydrate ($CaCl_2.2H_2O$), Co-enzyme A, Comassie Blue R250, 5-5'-bis-nitrobenzoate (DTNB), Oxaloacetate (OAA), Sodium per-chlorate ($NaClO_4$), agarose (low gelling temperature), Ethidium Bromide, Glycerol, Phenol, N-Laurylsarcosine, Formamide, Lithium Chloride (LiCl), TEMED, Tween 20 were purchased from Sigma Chemical Co., St. Louis. Magnesium Chloride ($MgCl_2$), Potassium dihydrogen Phosphate (KH_2PO_4), Tris-base, Chloroform, Dodecyl Sulfate Sodium Salt (SDS), Ethylenediaminetetraacetic acid (EDTA), Glycine, Dimethyl Sulfoxide (DMSO), Bromophenol Blue ($C_{19}H_{10}BrO_5S$), Acetic Acid, Ether $(C_2H_5OH)_2O$, Methanol, Mercaptoethanol, Tryptone, Sodium Chloride (NaCl), Sodium Hydroxide (NaOH) were form Boehringer Mannheim, Germany. Agar was purchased from Acumedia, Baltimore, USA and Yeast Extract was from Difco, Detroit, USA and OXOID Basingstoke, England.

Taq DNA polymerase and RNAase were purchased from Sigma Chemical Co., St. Louis, Missouri, USA. Restriction endonucleases *Eco*RI, *Bam*HI, *Hin*dIII, *Sac*I, *Pvu*II, *Ava*II and their buffers were from MBI Fermentas AB, Vilnus, Lithuania and Sigma Chemical Co., St. Louis, Missouri, USA.

DNA Extraction Kit was from MBI Fermentas AB, Vilnus, Lithuania. QIAGEN PCR Cloning Kit was obtained from QIAGEN Inc. Valencia, USA.

2.1.4 Buffers and Solutions

The buffers and the solutions used and their preparations are given in the Appendix C.

2.2 Genomic DNA Isolation from *Thermoplasma volcanium*

Genomic DNA of *Tp.volcanium* was isolated essentially using the modified method of Sutherland *et al.* (1990). After cell lysis, $NaClO_4$ was added to obtain a final concentration of 1 M. Then, one volume chloroform:isoamylalcohol (24:1, v/v) was added and the mixture was swirled on ice-water bath for 15 minutes until it was homogenous. After centrifugation at 11 000 rpm for 15 min. at 4ºC, the aqueous phase was transferred into a new tube and the DNA was precipitated in cold absolute ethanol. The pellet was washed in 70% ethanol, air-dried and then, was dissolved in minimal volume of sterile double distilled water.

2.2.1 Purification of Chromosomal DNA of *Thermoplasma volcanium*

Isolated chromosomal DNA was purified as described by Sambrook and Russel (2001). After removal of RNA by RNAase treatment, two successive phenol: chloroform (1/1, v/v) extractions were performed. The contaminating phenol was removed by ether extraction. The purified chromosomal DNA was precipitated in absolute ethanol and washed in 70 % (v/v) ethanol. Resulting DNA pellet was air dried, dissolved in minimal volume of double distilled water and stored at -20^{o}C until use.

2.2.2 Spectrophotometric Analysis of DNA

The amount of purified DNA was quantified as described by Sambrook *et al.* (2001). A series of dilutions (1/200-1/500) of DNA samples were prepared in sterile double distilled water. The DNA concentrations of these samples were determined spectrophotometrically (Shimadzu Analytical Co. Kyoto, Japan) at 260 nm. The ratio between the readings at 260 nm and 280 nm an (OD_{260}/OD_{280}) was calculated to estimate the purity of the nucleic

acids in the samples. A pure solution of double stranded DNA solution at 50µg/ml should have an optical density of 1.0 at 260 nm and OD_{260}/OD_{280} ratio of 1.8 (Sambrook *et al.*, 2001).

2.3 Agarose Gel Electrophoresis

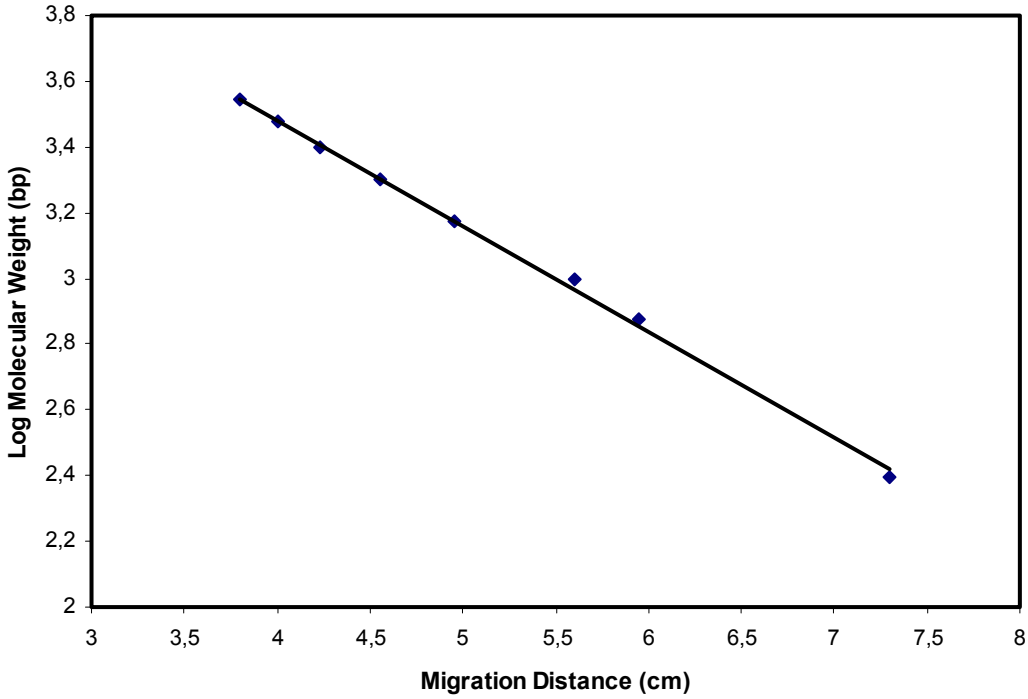
DNA fragments after restriction digestion, PCR amplicons and DNA samples following any manipulation were analyzed by agarose gel elecrophoresis with 1% (w/v) agarose in 1X TAE buffer. DNA samples of 10-20 µl, mixed with 1/10 volume of 6x tracking dye, were applied to gel which was supplemented by ethidium bromide (0.5 µg/ml). After electrophoresis the bands were visualized with a UV transilluminator (Vilber Lourmat, Marne La Vallée Cedex 1, France) and gel photographs were taken using a gel imaging and documentation system (Vilber Lourmat Gel Imaging and Analysis System, Marne La Vallée Cedex 1, France). The molecular sizes of DNA fragments were determined by referring to calibration curves, which were obtained by plotting the log molecular weights of known marker fragments against migration distance on the gel. Calibration curves generated for gene ruler 1 kb DNA Ladder (MBI Fermentas AB, Vilnius, Lithuania) and *Eco*RI/*Hind*III cut Lambda DNA (MBI Fermentas AB, Vilnius, Lithuania) are shown in the Figure 2.1.

2.4 Southern Blotting and Hybridization

2.4.1 Probe Preparation and Labeling

We have used *Tp.acidophilum* citrate synthase gene as probe for detection of the citrate synthase gene in the genomic DNA of *Tp.volcanium* by Southern hybridization. The *Tp.acidophilum* citrate synthase gene from the recombinant plasmid vector pCSEH19 was removed by double digestion

A)



B)

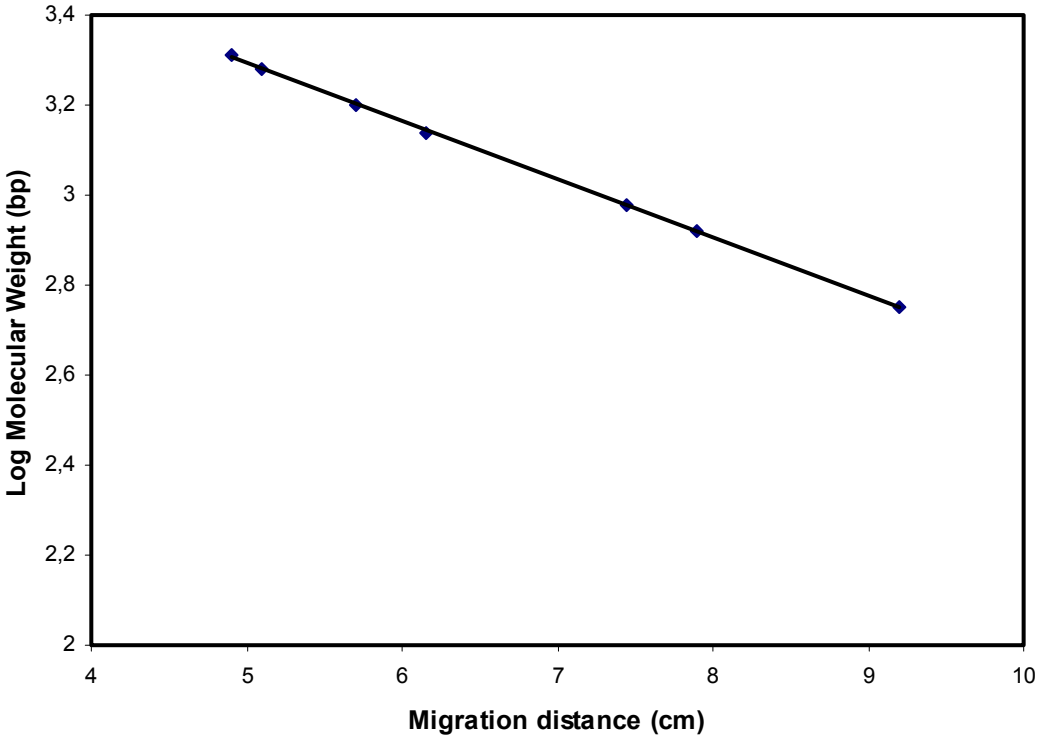


Figure 2.1 Calibration curves for molecular size markers A) Gene Ruler, 1 kb DNA Ladder and B) *Eco*RI and *Hind*III cut Lambda DNA, both are (MBI Fermentas AB, Vilnius, Lithuania)

with *Hin*dIII and *Eco*RI. The gene fragment then was isolated from the gel using MBI DNA Extraction Kit.

Another probe was the PCR amplified *Tv*CS gene fragments that was used to demonstrate the existence of the *Tv*CS gene in the recombinant pDrive Vector.

To label probe DNA a non-radioactive DIG system that uses digoxigenine was employed (DIG DNA Labeling and Detection Kit, Boehringer Mannheim Roche Diagnostics GmbH, Mannheim, Germany). Template DNA (0.5 µg- 3 µg) was denatured by heating and than chilled on ice. To denatured probe hexonucleotide mix, dNTP mixture and Klenow enzyme were added. The solution was incubated at 37^{o}C at least 60 min. Reaction was stopped by adding 20mM EDTA, pH 8.0. Ethanol precipitation was carried out in the presence of LiCl, as described before.

Standard molecular size markers Gene Ruler, 1 kb DNA Ladder and *Eco*RI and *Hind*III cut Lambda DNA (MBI Fermentas AB, Vilnius, Lithuania), were used for size determination in hybridization experiments (see Appendix B) and labeled, as described before.

2.4.2 DNA Capillary Transfer and Hybridization

Genomic DNA of *Thermoplasma volcanium* was digested with *Hin*dIII, *Bam*HI, and *Eco*RI (MBI Fermentas AB, Vilnus Lithuania) restriction enzyme separately and then samples were electrophoresed on a 1.0% agarose gel. The DNA molecules from gels were transferred onto a nylon membrane (Zeta-Probe, Bio-Rad, Richmond, California, USA) by capillary transfer technique using 0.4 N NaOH as the transfer solution, as shown in Figure2.2.

For characterization of recombinant plasmid, pDrive-CS31 was cut with

various restriction enzymes and their membrane blots were prepared as described above. Pre-hybridization, hybridization and post-hybridization washes were the same as described in the manufacturer's protocol (DIG-DNA Labeling and Detection kit, Boehringer Mannheim Roche Diagnostics GmbH, Mannheim, Germany). The nylon membrane blot was washed in 2X SSC buffer, dried and incubated in prewarmed hybridization buffer in nylon bags for 30 min by gentle agitation (Stuart Scientific, U.K). For hybridization, hybridization buffer was replaced by fresh hybridization buffer containing denatured digoxigenine labeled probe DNA. Hybridization was performed at 45^{o}C overnight.

2.4.3 Immunological Detection

After hybridization and stringency washes membrane was directly used for enzyme-linked immunoassay using a DIG DNA Labeling and Detection kit (Boehringer Mannheim Roche Diagnostics GmbH, Mannheim, Germany) following the manufacturer's instructions. The labeled probes were detected after hybridization to target nucleic acids using an antibody conjugate (anti-digoxigenine alkaline phosphatase conjugate). A subsequent enzyme catalyzed color reaction with 5-bromo-4-chloro-3-indolyl phosphate (X-phosphate) and nitro-blue (NTB) salt was used to visualize hybrid molecules.

2.5 PCR Amplification of Citrate Synthase Gene from *Tp.volcanium* Genomic DNA

2.5.1 Design of PCR Amplification Primers

An oligoprimer set for PCR was designed on the basis of predicted citrate synthase gene sequence of *Tp.volcanium* (Kawashima *et al*., 2000). The

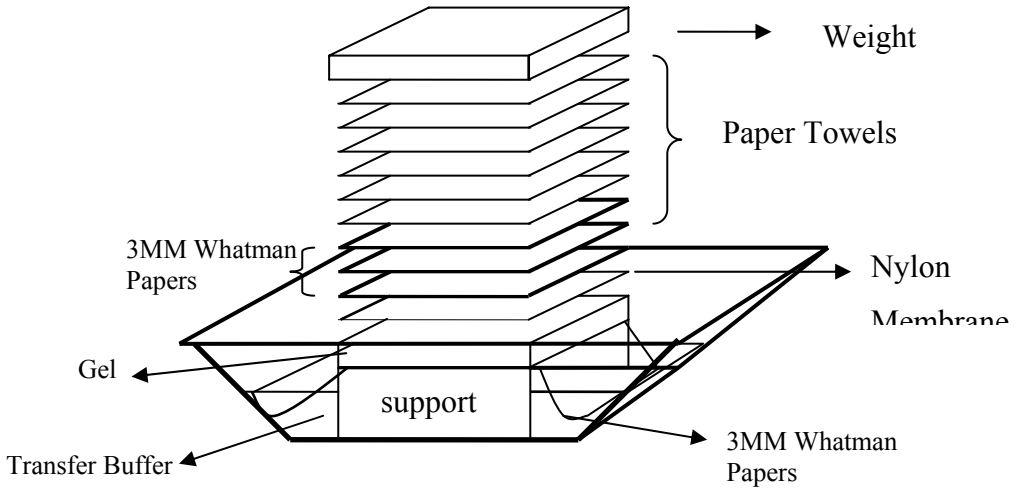


Figure 2.2 Schematic representation of Southern blot apparatus.

forward (FPcs) and reverse (RPcs) primers covered a genome region of 1476 bp extending from -241 position in upstream to +1235 position in downstream.

2.5.1 PCR Amplification

PCR amplification was carried out with the FPcs and RPcs, using *Tp. volcanium* genomic DNA as template. Reaction mixture (100μl) is composed of 200 μM of deoxyribonucleoside triphosphate (dNTP) mixture, 1x PCR buffer (750 mM Tris-HCl pH 8.8 at 25°C, 200 mM $(NH_4)_2SO_4$, 0.1 % Tween20), approximately 48 ng of the template DNA, 3 μM each of primers and 2.5 U of Taq DNA polymerase. The thermal cycling program (Techgene, Techne Inc. NJ. USA) started with denaturing at 94°C for 5 min followed by 30 cycles of amplification (denaturation at 94°C for 1 min, annealing at 55°C to 67°C for 2 min, and polymerization at 72°C for 3 min). An additional 10 min. at 72°C was allowed for completion of primer extension. PCR products (20μl) were analyzed by agarose gel electrophoresis.

2.6 Isolation of DNA Fragments from Agarose Gel

For the isolation of DNA fragments from agarose gel MBI Fermentas DNA Extraction Kit was used. The DNA fragments were run on the 1% Low Gelling Temperature Agarose gel (Sigma Chemical Co., St. Louis Missouri, USA) Gel. Isolation was performed according to the Manufacturer's Recommendations. The gel region containing the target DNA fragments was dissected after electrophoresis. After weighting the gel was solubilized in Gel Solubilization Buffer (30µl/10mg gel). One microliter of Silica Resin was added for every 10 mg of gel. The mixture was vortexed and then incubated at 50^{o}C for 15 min. During incubation the mixture was mixed for every 3 min. DNA-silica complex was isolated from the solution by centrifuge at 12,000 rpm for 30 seconds. A high salt wash of silica resin was performed with Wash Solution (30µl/10mg gel). After two low salt washes the silica resin was air-dried. For elution of DNA, the resin was suspended in TE buffer and centrifuged at 12.000 g for 30 seconds.

2.7 Cloning of PCR Amplified *Tp.volcanium* Citrate Synthase (*Tv*CS) Gene

2.7.1 Preparation of Competent *E.coli* cells

Competent *E.coli* TG1 cells were prepared according to the modified method of Chung *et al*. (1989). The *E.coli* TG1 cells were grown in 20 ml of LB medium through vigorous shaking (Heidolph UNIMAX1010, Heidolph Instruments GmbH, Kelheim, Germany). The growth was followed by measuring the optical absorbance of cell culture at 600 nm using the Shimadzu UV-160A double beam spectrophotometer (Shimadzu Analytical Co. Kyoto, Japan). When the number of cells were 10^{8} they were collected by centrifugation at 4000rpm for 10 min. The supernatant was discarded and, the pellet was dissolved in 1/10 TSS solution. Then the competent cells were distributed into eppendorf tubes as 100µl aliquots which were stored at -80^{o}C until use.

2.7.2 Ligation and Transformation

PCR amplified *Tv*CS gene was cloned using Qiagen PCR Cloning Kit. The PCR fragments were ligated to pDrive cloning vector, which was provided with the Kit, following the Manufacturer's Instructions. The ligation mixture (10µl), contained 1-4µl extracted PCR fragments, 1µl pDrive Cloning Vector and 5µl Ligation Master Mix (2X) and variable amount of distilled water. The mixture was then incubated at 4-16^{o}C for at least 30 min.

For transformation the competent cells stored at -80^{o}C were thawed on ice. An aliquot of ligation mixture was added into thawed competent *E.coli* TG1 cells and mixed gently. The mixture was incubated on ice for 30 min and then transferred to glass tubes containing LB broth. The cells were grown at 37^{o}C by vigorous shaking (Heidolph UNIMAX1010, Heidolph Instruments GmbH, Kelheim, Germany). Then a series of ten-fold dilutions of the cells were spread onto LB agar plates containing ampicillin, X-Gal and IPTG. The plates were incubated at 37^{o}C for 16-18 hours.

2.8 Isolation of Plasmid DNA from Recombinant Colonies

Cells from the white colonies were inoculated into 10 ml LB culture containing ampicillin and incubated at 37^{o}C overnight. Plasmid isolation was carried out by using Wizard® Plus SV Minipreps DNA Purification System (Promega Corporation, Maidson, WI, USA) according to the procedure provided by the manufacturer. The bacterial cells were precipitated by centrifugation at 4000 rpm for 15 min. The pellet was thoroughly resuspended in 250µl Cell Resuspension Solution. The cells were lysed by

adding 250µl Cell Lysis Solution and the proteins were digested with 10µl Alkaline Protease. After incubation for 5 min at room temperature 350µl Neutralization Solution was added and incubated for an additional 5 minutes.

The cells debris was pelleted by centrifugation at 13,000 rpm for 10 minutes and the supernatant was transferred to Spin Column, which was provided with the Kit and centrifuged at 13,000 rpm to bind the DNA to the column. The bound DNA was washed twice by Washing Solution and eluted with 100µl Nuclease Free water. Finally 100µl of plasmid sample, at a concentration of 1.5 mg/ml, was obtained from the starting material. Nearly 15µl of the sample was run on the agarose gel to check the efficiency of purification and the remaining sample was stored at -20^{o}C for further use.

2.9 Selection of True Recombinants by Restriction Analysis

The putative recombinant plasmids isolated form white colonies were analyzed by restriction enzyme digestion to confirm the insertion of *Tv*CS gene to the cloning plasmid vector. The restriction enzymes used in these experiments were *Bam*HI, *Eco*RI and *Hin*dIII, *Pvu*II, *Ava*II and *Sph*I. Single and double digestions with various restriction enzymes also allowed us to construct the restriction map of the cloned *Tv*CS gene experimentally and to find out its insertion orientation.

2.10 Sequence Determination of *Tv*CS Gene

The sequence of inserted PCR product containing *Tv*CS gene was determined by automated DNA sequencing by Yale University, New Haven, Connecticut, USA using the reverse and forward PCR primers.

2.11 Multiple Sequence Alignment and Processing Sequence Data

Clone-4 software was used to map the sites for restriction endonucleases in the putative and experimental *Tp.volcanium* citrate synthase gene sequence.

EMBL (European Molecular Biology Laboratories) sequence database was searched to obtain the genomic DNA sequence of *Tp.volcanium* and the position of putative *Tp.volcanium* citrate synthase gene sequence.

ClustalW version 1.82 (Thompson *et al*., 1994) was used for multiple sequence alignment and construction of phylogenetic tree, based on the citrate synthase amino acid sequence information from different organisms.

2.12 Purification of Recombinant *Tv*CS

Purification of recombinant *Tv*CS enzyme was carried out using Dye-ligand affinity column chromatography as described by Kocabıyık *et al*. (1996).

2.12.1 Preparation of Cell-free Extracts

The culture of *E.coli* TVCS C-35 cells was grown in 75 ml LB-broth containing 50μg/ml of ampicillin in 250ml flasks. The flasks were incubated for 18 hours, at 37^{o}C through continuous shaking at 160rev/min in an environmental shaker (Heidolph Incubator 1000, Heidolph Instruments GmbH and CoKG, Germany)

The bacterial cells were collected by centrifugation at 4000 rpm for 15 min and the pellet was resuspended in 1/10 volume of 20 mM Tris-HCl, 1mM EDTA (TE) buffer, pH 8.0. The cells were disrupted by sonication (Sonicator VC 100, Sonics and Materials Inc., Danbury, CT) and the cell debris was removed by centrifugation at 12,000 rpm for 45 min by Sigma 3K30 Centrifuge (Sigma Chemical Co., St. Louis, Missouri, USA). The supernatant was called cell-free extract.

To denature heat-labile *E.coli* proteins, the extract was incubated in

glass bottles at 60^{o}C for 15 min. Then, denatured host proteins were removed by centrifugation at 12,000 rpm for 1.5 hrs, and the supernatant was kept at +4^{o}C, until use.

2.12.3 Dye-ligand Affinity Chromatography on Reactive Red 120 Column

For the one-step purification, affinity chromatography of the recombinant citrate synthase on Reactive Red 120 column (Dye Matrex Red A, Sigma Chemical Co. St. Louis, MO) was performed. After heat treatment, cell extract was applied to the affinity column, which is equilibrated with 10 column volumes of TE buffer. The unbound proteins were removed by passing TE buffer until the spectrophotometric readings of fractions at A_{280} were leveled off. Specifically bound citrate synthase was eluted by washing the column with 5mM OAA-1mM Acetyl-CoA solution. Then the fractions collected by the aid of a Biorad Model 2110 Fraction Collector (Bio-Rad laboratories, USA) were checked for their protein contents by monitoring the absorbance at 280nm and for citrate synthase activities. Chromatography was performed at 4^{o}C.

2.13 Concentrating Enzyme Samples

The fractions with detectable enzyme activity were pooled, washed and then concentrated about three fold by ultrafiltration using 5000 cut-off ultrafiltration device (Vivascience, Sartorious Group, Stonehouse, Gloucestershire, UK). The concentrated purified enzyme was stored at +4^{o}C.

2.14 Determination of Protein Content

Protein amounts of cell-free extract before heat treatment (BH sample), cell free extract after heat treatment (AH sample) and purified enzymes were determined by the method of Whitaker and Granum (1980). The

absorbencies of the samples were measured at 235 nm and 280 nm. The concentrations of proteins were determined by the formula of $C_{(mg/ml)} = (A_{235}-A_{280}/2.51) \times$ Dilution Factor.

2.15 Enzyme Assay for Citrate Synthase

The citrate synthase activity was determined as described by Srere *et al.* (1963). This method is based on measurement of the formation rate of 5-thio-2-nitrobenzoate (TNB) as a result of the reaction between 5-5'-dithio-bis-nitrobenzoate (DTNB) and free –SH group of AcetylS-CoA formed after the citrate synthase reaction. Absorbance was measured at 421nm in a Shimadzu UV-visible 1601A double-beam spectrophotometer with a temperature controlled cell holder (Shimadzu Corp., Scientific Instruments Division, Kyoto, Japan).

Standard reaction mixture (1ml total volume) contained 0.2 mM OAA, 0.15 mM acetyl-CoA, 0.2 mM DTNB and appropriate volume of enzyme solution in 20 mM Tris-HCl-1mM EDTA buffer (pH8.0). For the assay, mixture was preincubated at 55^{o}C for 5 min. and reaction was initiated by the addition of OAA. Absorbance at 412 nm was measured continuously at 55^{o}C. 1U of enzyme activity was defined 1μmol CoA-SH produced per minute under standard assay conditions.

2.15.1 Partial Characterization of Recombinant *Tv*CS Enzyme

2.15.1 Determination of Kinetic Parameters

To determine kinetic parameters *i.e.* V_{max} and K_m values the enzyme activities at different concentration of the substrates Acetyl-CoA and OAA were measured. To achieve this, concentration of one substrate was kept constant while concentration of the other substrate was changed. In these experiments, OAA concentration was kept constant at 200µM while acetyl-CoA concentration varied between 0.75µM and 26.5µM .On the other hand,

when acetyl-CoA concentration was fixed at 0.15mM, OAA concentration was changed between 2μM and 30μM. In all assays the enzyme concentration was 1.93 U/ml.

The kinetic parameters (K_m and V_{max}) were determined from the Lineweaver-Burk Plot (1/V *vs*. 1/S) where V is the initial velocities (μmoles CoA-SH/ml/min) and S is the substrate concentration (μM).

2.15.2 Thermostability Analysis

Thermal stability of *Tv*CS enzyme was measured by incubating appropriate volumes of the enzyme at temperatures: 30°C, 40°C, 50°C, 60°C, 70°C, 80°C, 83°C, 85°C, 87°C, 90°C and 95°C for 60 min. Aliquots of about 25μl were removed at regular interval and quickly chilled on ice. The remaining enzyme activity was measured following the standard assay.

2.12.3 SDS-polyacrylamide Gel Electrophoresis

SDS-polyacrylamide gel electrophoresis was performed essentially as described by Laemmli *et al*. (1970) to check the enzyme purity and to determine the subunit molecular weight of the recombinant enzyme. For the separation, 10% (w/v) polyacrylamide gel with 0.5% SDS was used. Volume of each sample was adjusted so that 10μg protein was applied to each well. An electric field of 100 V for 3hrs was applied to separate proteins through the gel. Then, the gel was stained with Coomassie Blue-R solution, which was followed by destaining in Coomassie Gel Destain Solution. For the molecular weight determination standard protein markers (Sigma Chemical Co., St. Louis, Missouri, USA) were used (Figure 2.3).

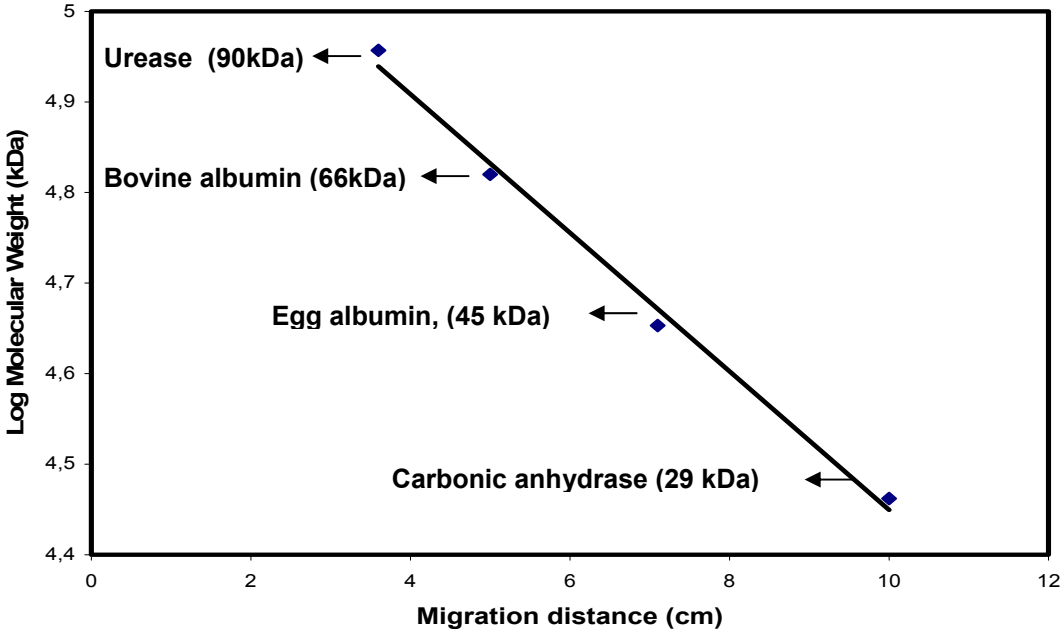


Figure 2.3 Calibration curve for protein molecular size marker.

CHAPTER 3

RESULTS

3.1 Isolation and Purification of *Tp.volcanium* Genomic DNA

Tp.volcanium total genomic DNA was isolated and purified as described in the Materials and Methods. Purified DNA samples were visualized as single bands on agarose gel after electrophoresis. Figure 3.1 shows the DNA samples before and after purification procedure. Purified DNA samples were analyzed spectrophotometrically. The absorbance readings at 260 nm and 280 nm were recorded to calculate the quantity and purity of DNA samples as described in the Section 2.2.2. The concentrations of the samples varied between 2.8 and 3.0mg/ml. OD_{260}/OD_{280} was between 1.8 and 1.87, indicating a sufficient purity for DNA samples to be used in gene manipulation studies. The eligibility of the purified DNA samples for gene manipulation experiments was also checked by restriction enzyme (*Hin*dIII) digestion. (Figure 3.2)

3.2 Amplification of *Tp.volcanium* Citrate Synthase Gene

*Tv*CS gene was amplified using FPcs and RPcs from the genomic DNA of *Tp.volcanium* by PCR. PCR reactions were carried out at highly stringent hybridization temperatures (60^{o}C, 65^{o}C, 67^{o}C) and every time a single unique band was observed on agarose gel. Amplification has yielded a PCR

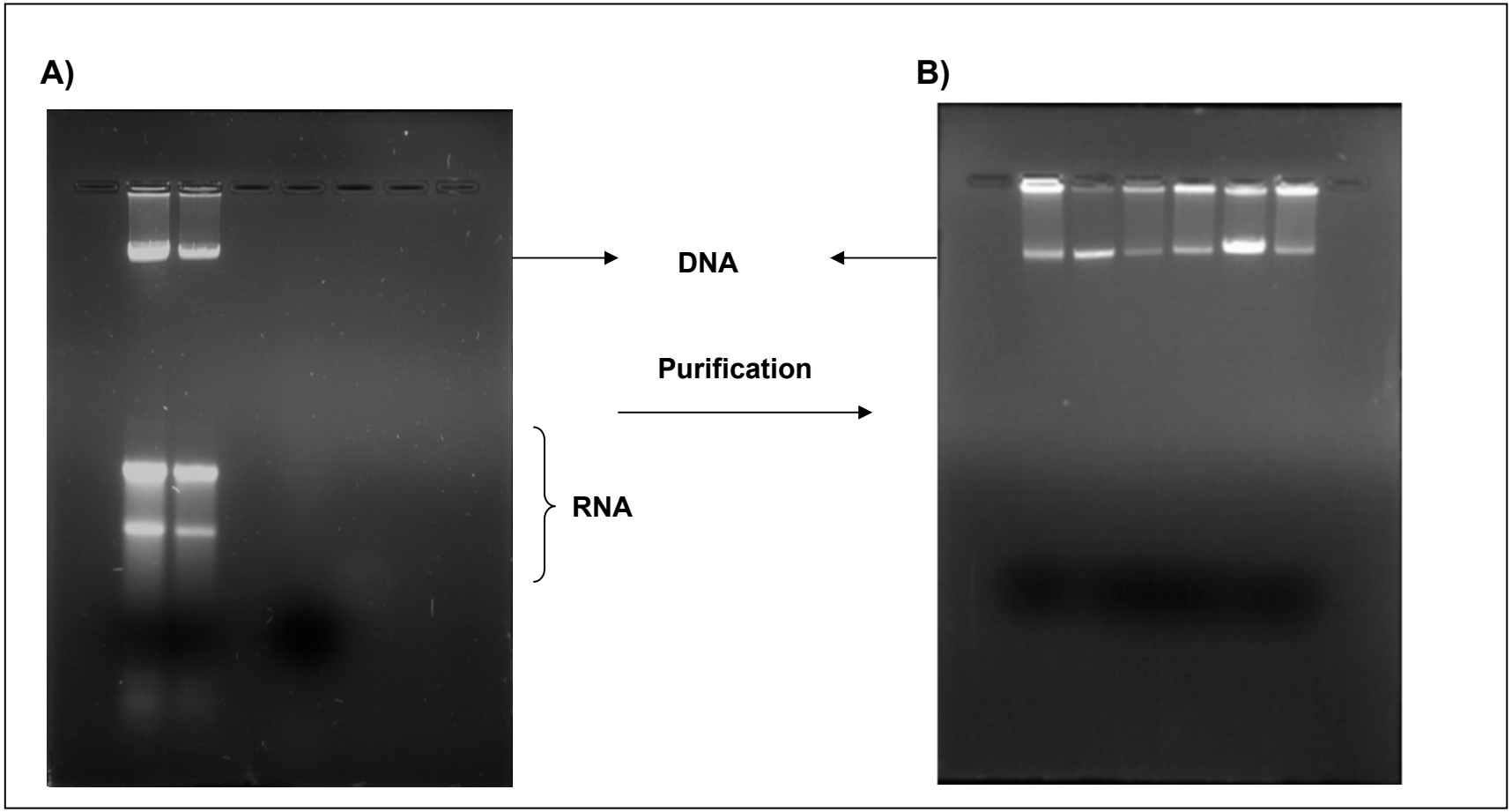


Figure 3.1 Total genomic DNA of *Tp.volcanium* before (A) and after purification (B).

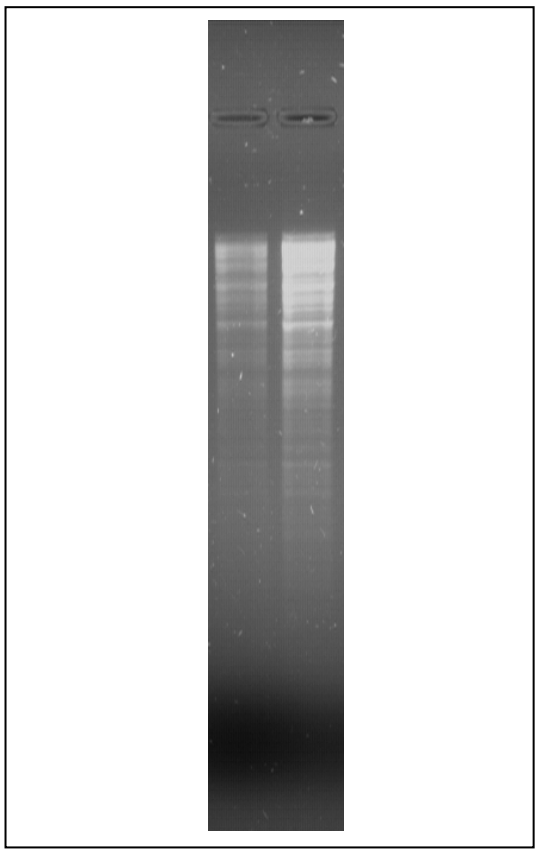


Figure 3.2. Purified *Tp.volcanium* genomic DNA digested by *Hin*dIII.

product of an expected size of about 1476 bp (Figure 3.3). This fragment contains a 1161bp open reading frame corresponding to *Tv*CS structural gene, 247 bp upstream and 78 bp downstream untranslated regions.

3.3 Cloning of PCR Fragments Containing *Tp.volcanium* Citrate Synthase Gene

The PCR amplified *Tv*CS gene (1476bp) was ligated to pDrive cloning vector. For cloning a Qiagen PCR cloning kit was used. The cloning process is shown in Figure 3.4. The cloning provides high-specifity UA hybridization and efficient cloning of PCR products generated by Taq Polymerase and other non-proofreading polymerases. These enzymes add a non-template

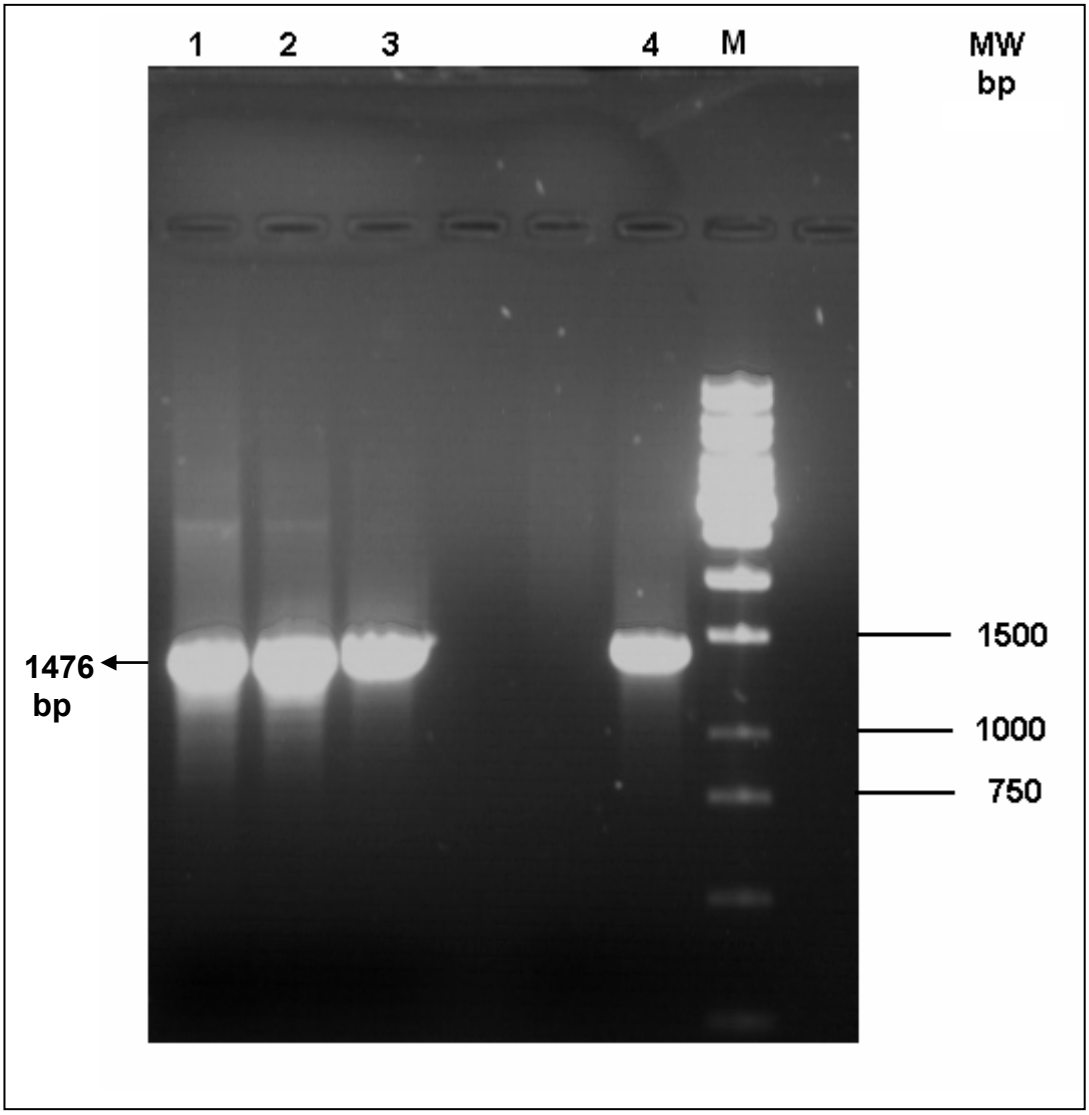


Figure 3.3 PCR amplification results at different annealing temperatures: Lane 1, at 55^{o}C; lane 2 at 60^{o}C; lane 3, at 65^{o}C and lane 4, at 67^{o}C. Amplifications yielded single unique bands of 1476 bp size. M: Molecular size marker: Gene Ruler, 1 kb DNA Ladder.

specified adenine (A) residues to 3' termini of both strands of amplicons, which can be exploited in cloning. In the Kit, pDrive vector is supplied which has a U overhang at each 3'-end, so hybridizes with high specifity to the A overhang of PCR products. In addition, the vector allows ampicillin and kanamycin selection as well as blue/white screening of recombinant colonies. Competent *E.coli* TG-1 cells were used as recipients. The transformation efficiency varied between $4x10^7$-10^8 transformants/µgDNA. The recombinant clones were selected through blue/white screening. A total of 12 white colonies were picked up and plasmid DNAs were isolated from 10 of these colonies as described in the Materials and Methods (Figure3.5)

3.4 Characterization of Recombinant Plasmid Containing Citrate Synthase Gene

The presence of insert in the putative recombinant plasmids, which were isolated form white colonies restriction digestion analysis, was performed. The recombinant plasmids were linearized with *Bam*HI digestion, which yielded a single fragment of 5327bp corresponding to that of pDrive vector (3851 bp) plus insert (1476bp). (Figure3.6)

There are two *Eco*RI sites in the multiple cloning sites (MCSs) of the vector flanking the insertion sites, which facilitate the excision of the intact insert. Similarly, *Bam*HI-*Sac*I (each located in either of the flanking MCSs) double digestion, like *Eco*RI single digestion of the recombinant plasmids yielded two bands: the larger one is 3835 bp linear pDrive cloning vector and the smaller one is 1492 bp *Tv*CS gene.

Out of 10 plasmids analyzed by single and double digestions, 3 (C31, C34, C35) have been identified as true recombinants (Figure 3.6-3.7). Among these recombinants, pDriveCS-31 was selected to be further characterized by restriction analysis.

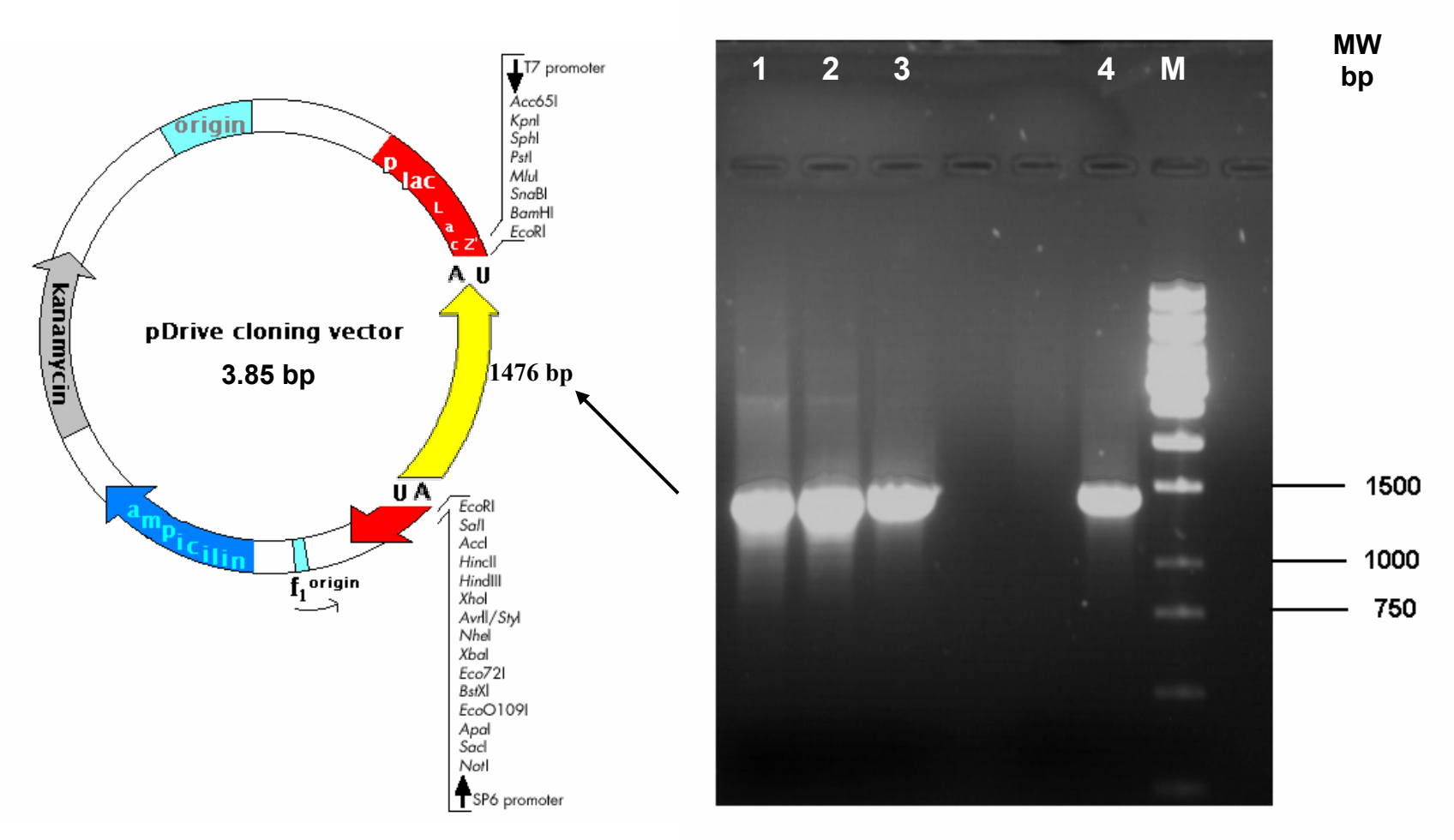


Figure 3.4 Cloning process of *Tv*CS gene into pDrive Cloning Vector. In lanes 1,2,3, and 4, PCR amplified fragments (1476 bp) are shown. The PCR amplified fragment at 65°C (shown in yellow), in the lane 3, ligated into pDrive Vector (3851 bp) by AU complementation. M: Molecular size marker: Gene Ruler, 1 kb DNA Ladder.

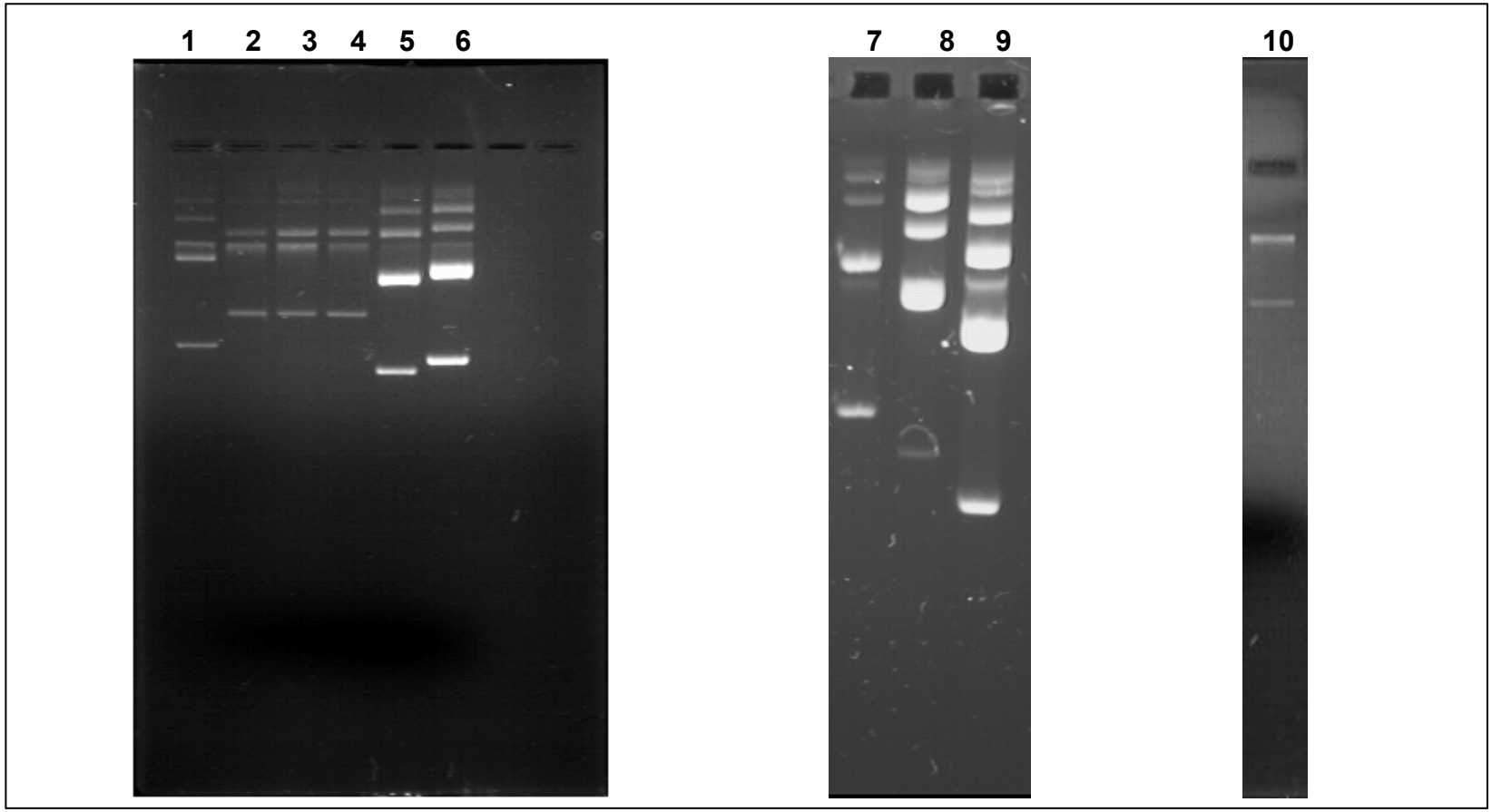


Figure 3.5 Putative recombinant plasmids isolated from white colonies: lane 1, pDriveCS-39; lane 2; pDriveCS-38; lane 3, pDriveCS-36; lane 4, pDriveCS-34; lane 5, pDriveCS-33; lane 6, pDriveCS-32; lane 7, pDriveCS-35; .lane 8, pDriveCS-37; lane 9, pDriveCS-310; lane 10, pDriveCS-31.

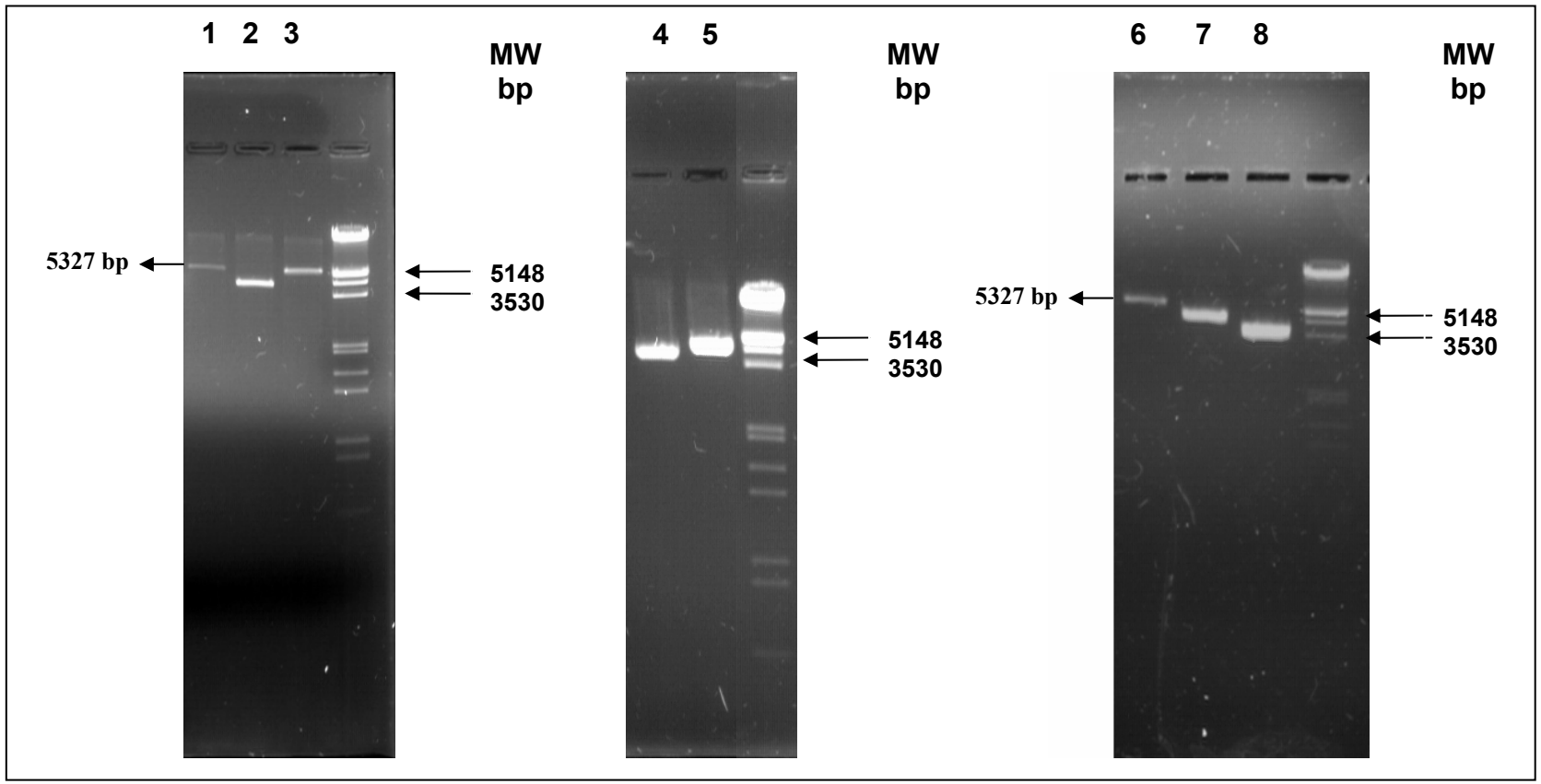


Figure 3.6 Digestions of putative recombinant plasmids by *Bam*HI to find out true recombinants. *Bam*HI digestion was expected to linearize plasmids. Lane1, pDriveCS-31; lane 2, pDriveCS-32, lane3, pDriveCS-34; lane4, pDriveCS-38; lane5, pDriveCS-39; lane6, pDriveCS-35; lane7, pDriveCS-37 and lane8, pDriveCS-310. A linearized recombinant vector would be 5,327 bp long including 1476 bp amplified *Tv*CS gene and 3,851 bp pDrive vector. Thus, pDriveCS-31, pDriveCS-34, and pDriveCS-35, proved to be true recombinants.

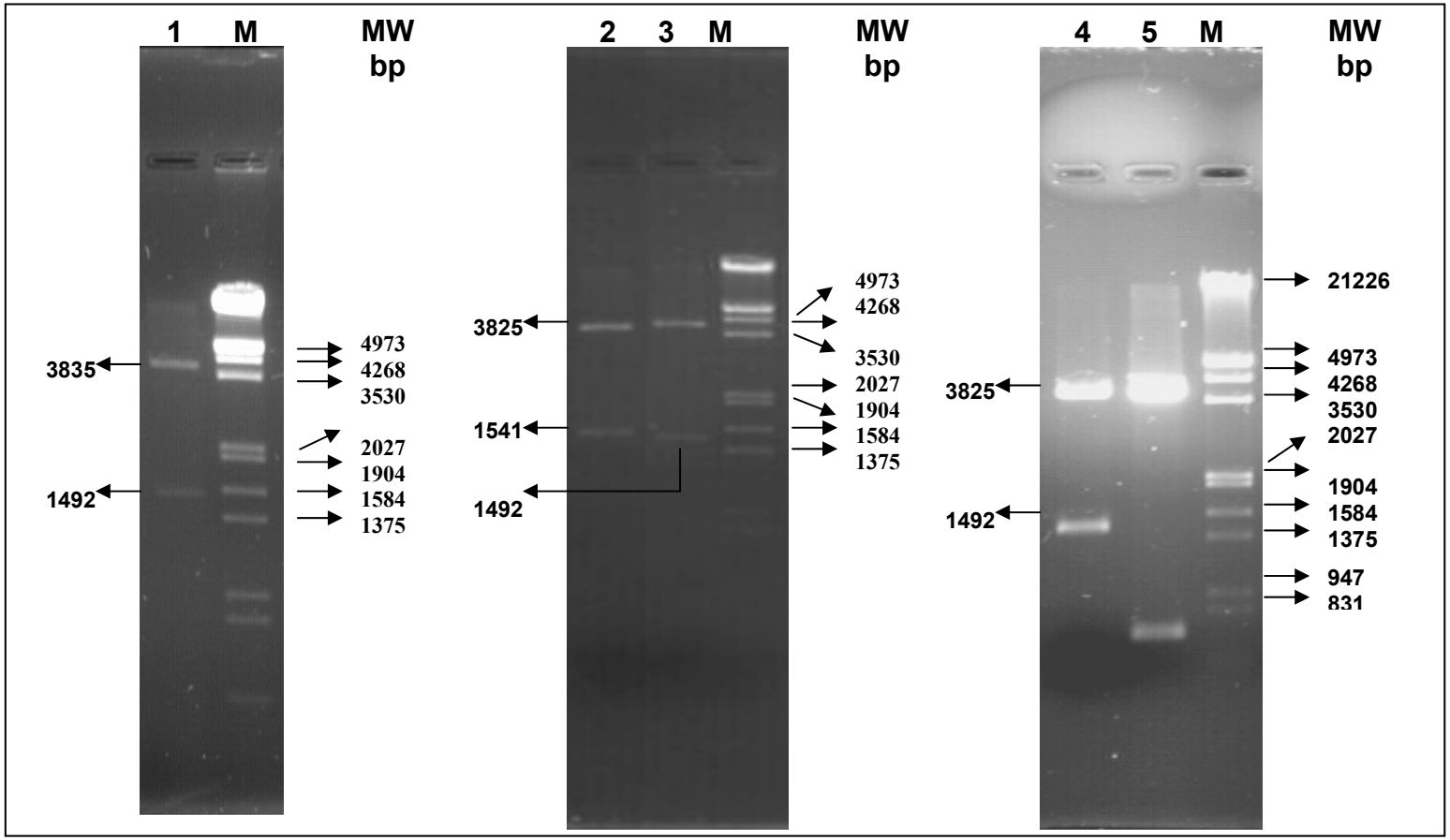


Figure 3.7 Double and single digestions of selected recombinant plasmids. Lane 1, *Bam*HI-*Sac*I digested pDriveCS-34, Lane 2, *Bam*HI-*Sac*I digested pDriveCS-31, Lane 3, *Eco*RI digested pDriveCS-31, Lane 4, *Eco*RI digested pDriveCS-35, Lane 5, *Eco*RI digested pDriveCS-37. M: Marker; *Eco*RI/*Hin*dIII digested λ DNA.

3.5 Southern Blotting/Hybridization Using *Tp.acidophilum* Citrate Synthase Gene Probe

3.5.1 Cross-hybridization between *Tp.volcanium* Genomic DNA Digests and *Ta*CS Gene Probe

To search for homology between *Tp.acidophilum* and *Tp.volcanium* citrate synthase genes genomic DNA blot of *Tp.volcanium* was hybridized with Dig labeled *Ta*CS gene probe. To do this, recombinant plasmid pBSK(2CS) containing *Ta*CS gene was isolated and double digested with *Hin*dIII and *Eco*RI (Figure 3.8). This digestion produced two fragments, a 900 bp *Ta*CS gene fragment, and 3824 bp vector plus remaining gene sequences. *Ta*CS gene fragment was purified from low melting agarose gel (Figure 3.9) and then DIG-labeled through random priming to be used as probe in the hybridization as described in the Materials and Methods. Total genomic DNA of *Tp.volcanium* was digested with *Eco*RI, *Hin*dIII and *Bam*HI. Membrane blots of these fragments when probed with 900 bp DIG labeled TaCS gene fragment, unique bands corresponding to 5530 bp *Eco*RI, 4830 bp *Bam*HI and 5130 bp *Hin*dIII fragments were observed. This result indicates a considerable homology between *Ta*CS and *Tv*CS genes, and the existence of the putative CS gene as a single copy in the *Tp.volcanium* genome (Figure 3.10).

3.5.2 Cross-hybridization between Cloned *Tv*CS Gene and *Ta*CS Gene Probe

Southern Blotting/Hybridization technique was used to confirm that the amplified region of the genome by PCR included the *Tv*CS gene. The membrane blot of PCR amplified putative *Tv*CS gene products was hybridized with DIG-labeled *Ta*CS gene probe. Unique strong hybridization

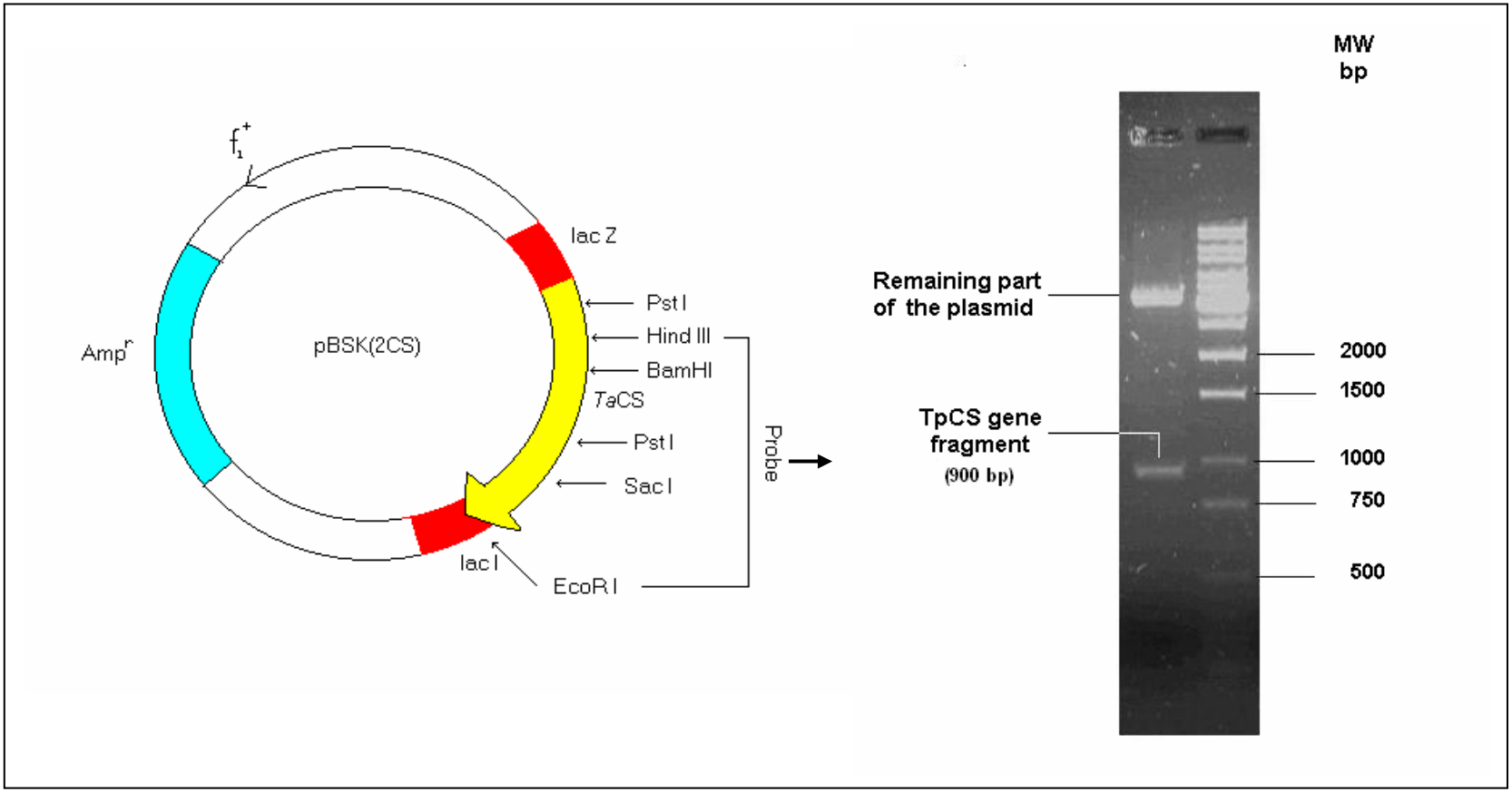


Figure 3.8 The *Ta*CS gene fragment (900bp) was recovered from pBSK(2CS) by double restriction digestion with *Hin*dIII and *Eco*RI and then isolation from the gel. Isolated gene fragment was DIG-labeled and used as probe in Southern hybridizations to investigate homology between *Tv*CS gene and TaCS gene. Molecular size marker: Gene Ruler, 1 kb DNA Ladder.

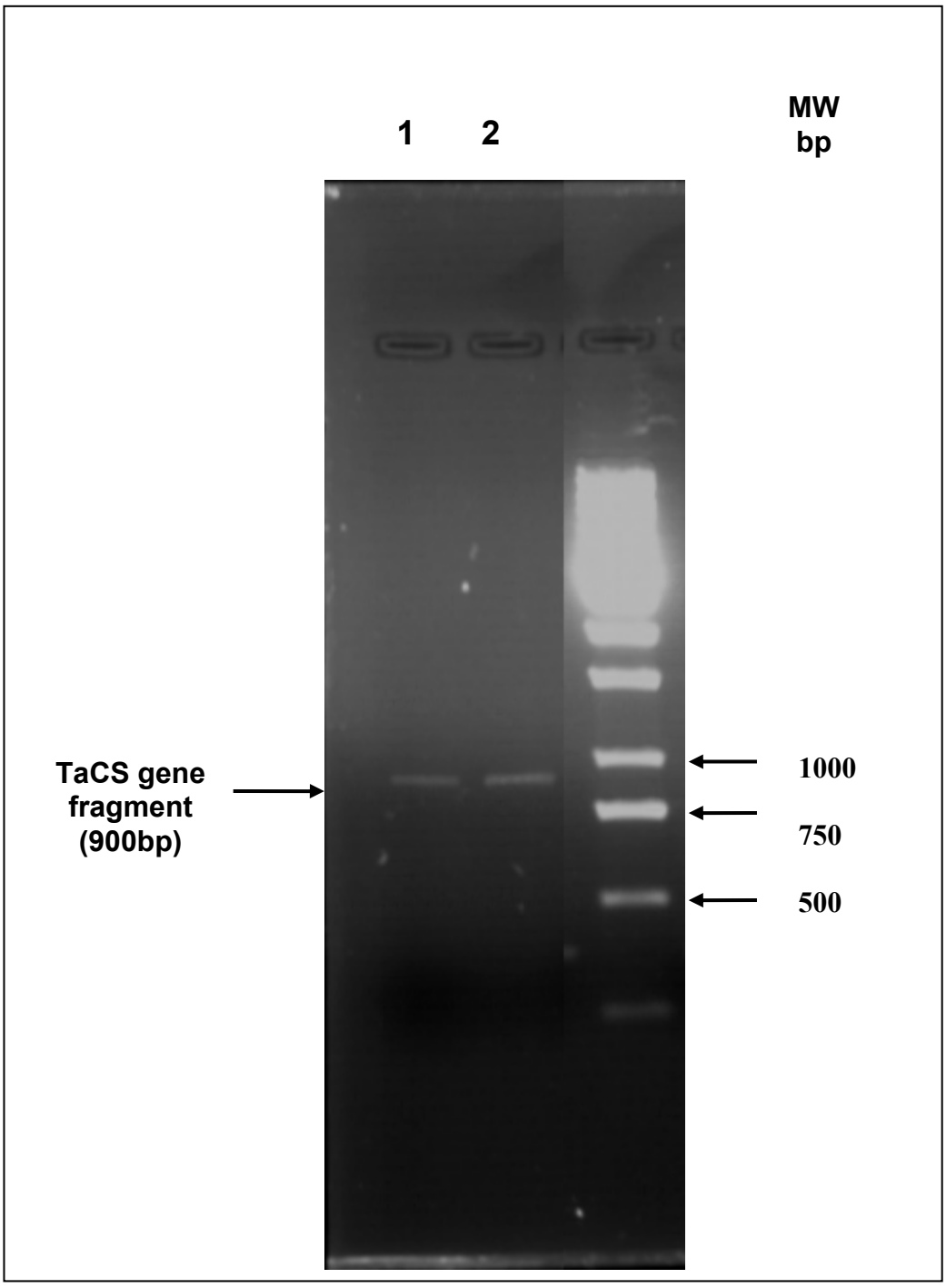


Figure 3.9 *Ta*CS gene fragment was removed from the plasmid (pBSK(2CS)) by double digestion with *Hin*dIII and *Eco*RI. Then it was extracted from low melting agarose gel as described in Materials and Methods. Lane 1 and 2 show the isolated *Ta*CS gene. Molecular size marker: Gene Ruler, 1 kb DNA Ladder.

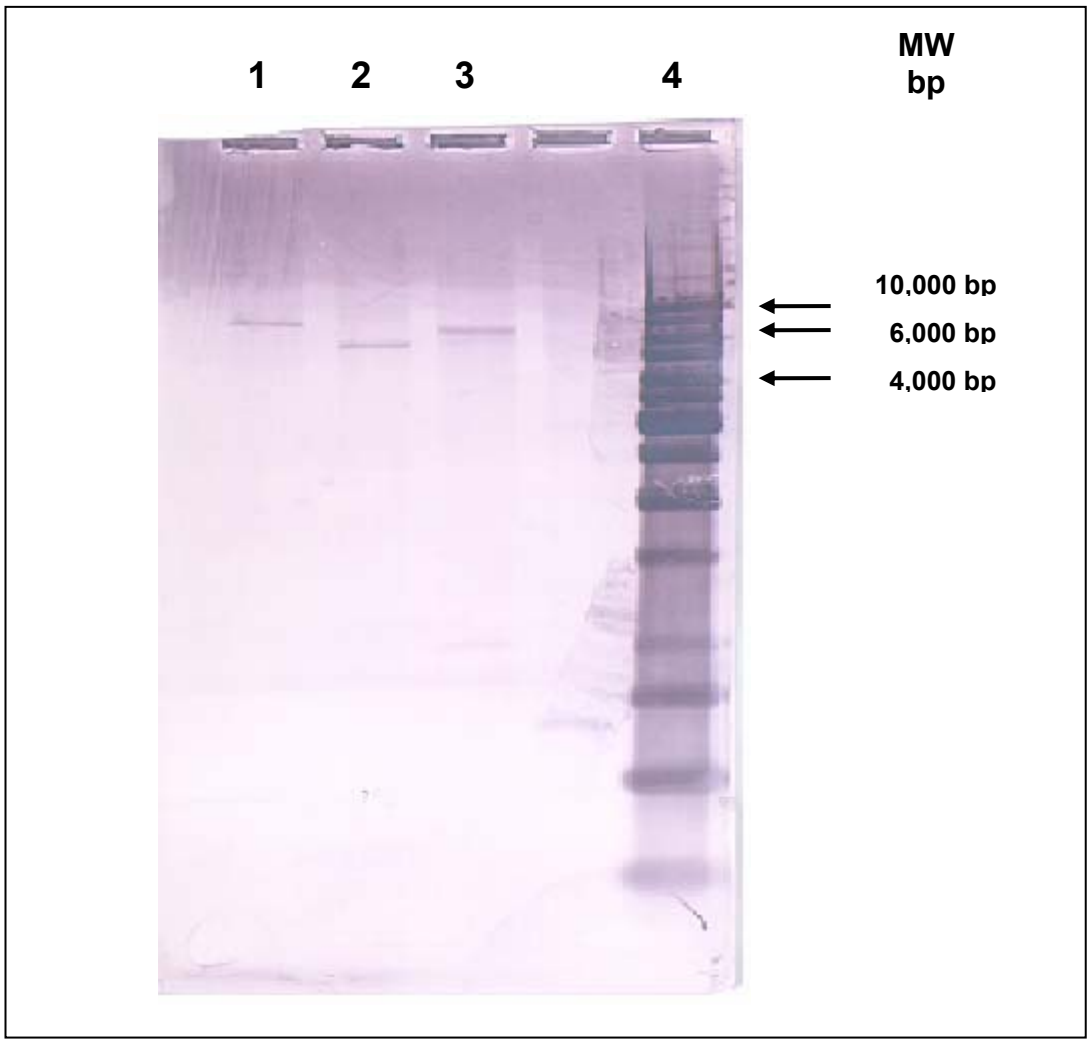


Figure 3.10 Southern Blot/Hybridization results using *Tp.volcanium* genomic DNA blot and *Tp.acidophilum* citrate synthase gene probe. In each lane, hybridization signal at a specific position was observed. Hybridizations with *Eco*RI (lane 1); *Bam*HI (lane 2), and *Hin*dIII (lane 3) yielded signal corresponding to about 5530 bp, 4830 bp and 5130 bp length fragments respectively. Size Marker: Gene Ruler, 1 kb DNA Ladder.

signals were obtained at the expected positions of about 1476 bp. This result further confirmed the high level homology between *Ta*CS and *Tv*CS genes as well as specific amplification of *Tv*CS gene sequences, by PCR successfully (Figure 3.11).

3.6 Southern Blot Analysis to Confirm Cloning

Cloning of PCR amplified *Tv*CS gene into pDrive cloning vector was confirmed by another Southern Blot/Hybridization analysis using the 1476 bp PCR product as probe. In order to prepare the probe for hybridization, PCR products obtained at an annealing temperature of 65^{o}C (Figure3.3) were isolated from low melting agarose gel (Figure3.12) and DIG-labeled via random priming as described in the Materials and Methods. Then, recombinant plasmid (pDriveCS-31) was digested with *Bam*HI, *Eco*RI and *Hin*dIII, and their membrane blots were prepared which were hybridized with DIG-labeled PCR probe under fairly high stringent conditions (hybridization at 45^{o}C). Strong hybridization signals were obtained at expected positions consistent with the restriction map of *Tv*CS gene shown in Figure 3.14 and Table 3.1. The PCR probes were also hybridized with the PCR fragments, which was a positive control (Figure 3.13). This result clearly showed that the target product had been inserted into the recombinant vectors isolated from the suspected recombinant colonies, *i.e.* pDriveCS-31, pDriveCS-34, and pDriveCS-35.

3.7 Restriction Enzyme Mapping of Recombinant Plasmid Containing *Tp.volcanium* Citrate Synthase Gene

A number of single and double digestions of the recombinant plasmid pDriveCS-31 were carried out with the restriction enzymes such as *Bam*HI, *Hin*dIII, *Eco*RI, *Pvu*II, *Ava*II and *Sph*I. Restriction map of *Tv*CS gene was

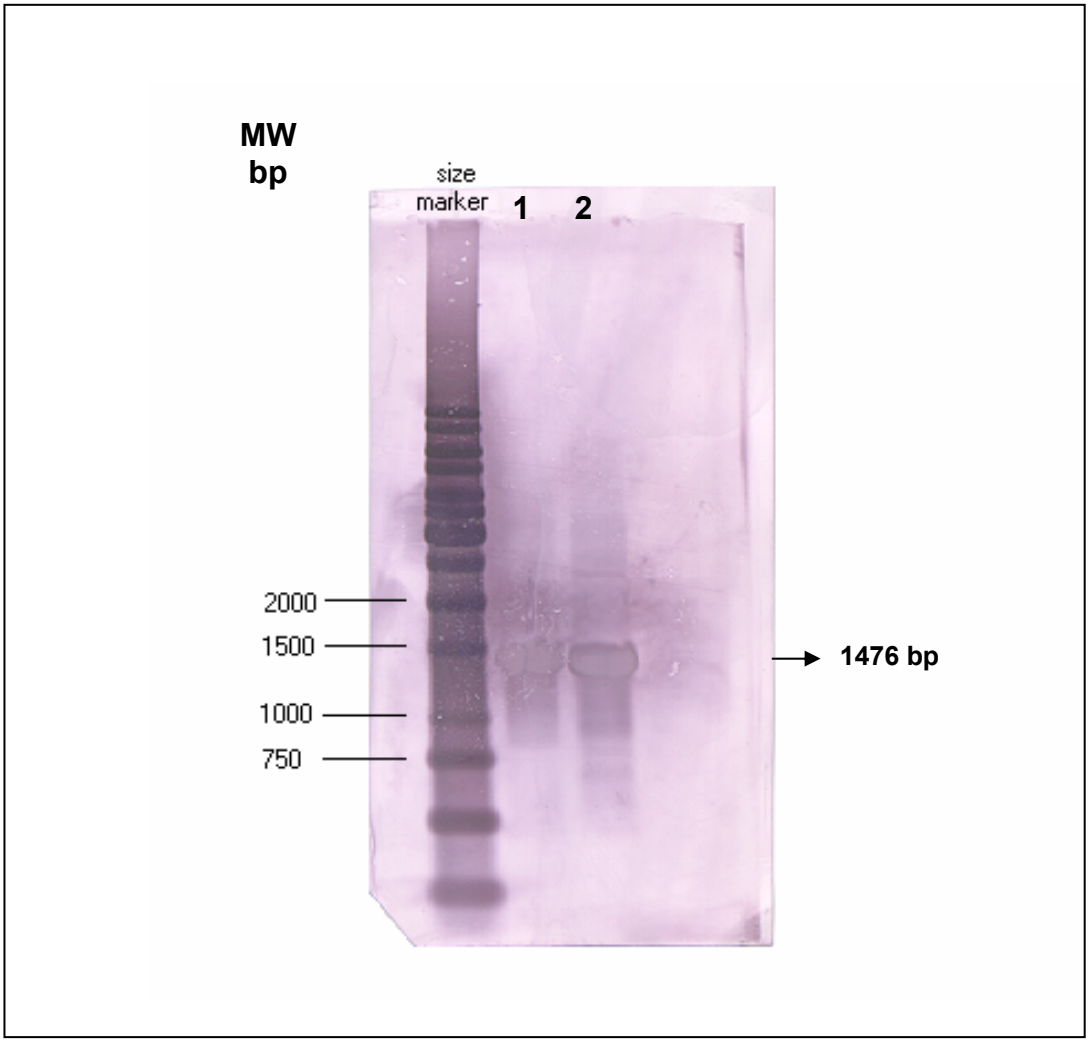


Figure 3.11 Result of Southern Blot/Hybridization using the blots of amplified PCR fragments (lane 1 at annealing temperature 65^{o}C and lane 2 at annealing temperature 67^{o}C) and *Ta*CS gene probe which was DIG labeled. The expected hybridization signals were obtained for both PCR fragments at the position corresponding to about 1476bp. Size Marker: Gene Ruler, 1 kb DNA Ladder.

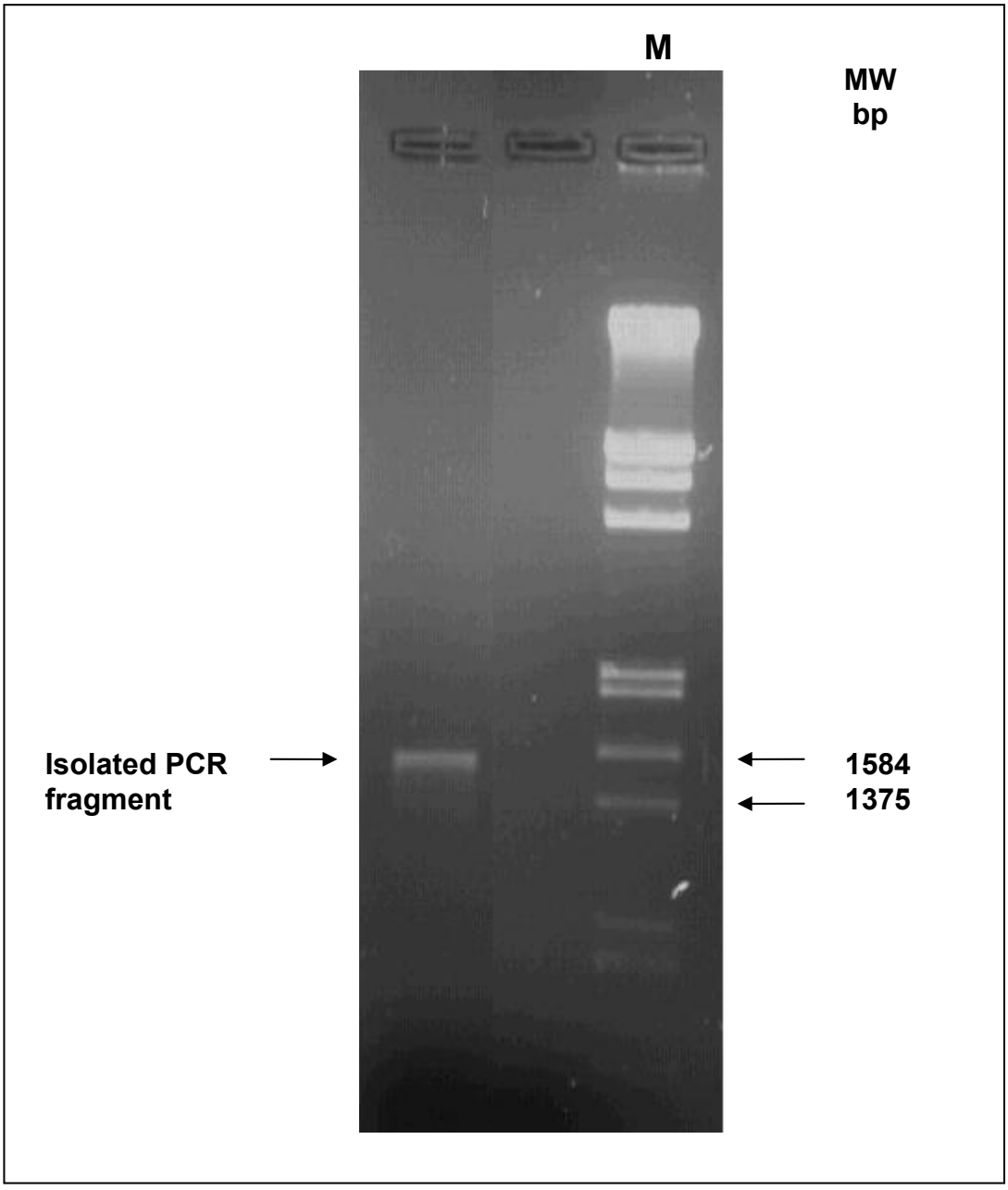


Figure 3.12 PCR fragment isolated from low melting agarose gel for probe preparation. Isolated PCR fragment was labeled through random priming and then hybridized with blots of pDriveCS-31 to confirm the presence of PCR amplified CS fragment. M: Molecular size marker: *Eco*RI/*Hin*dIII digested λ DNA.

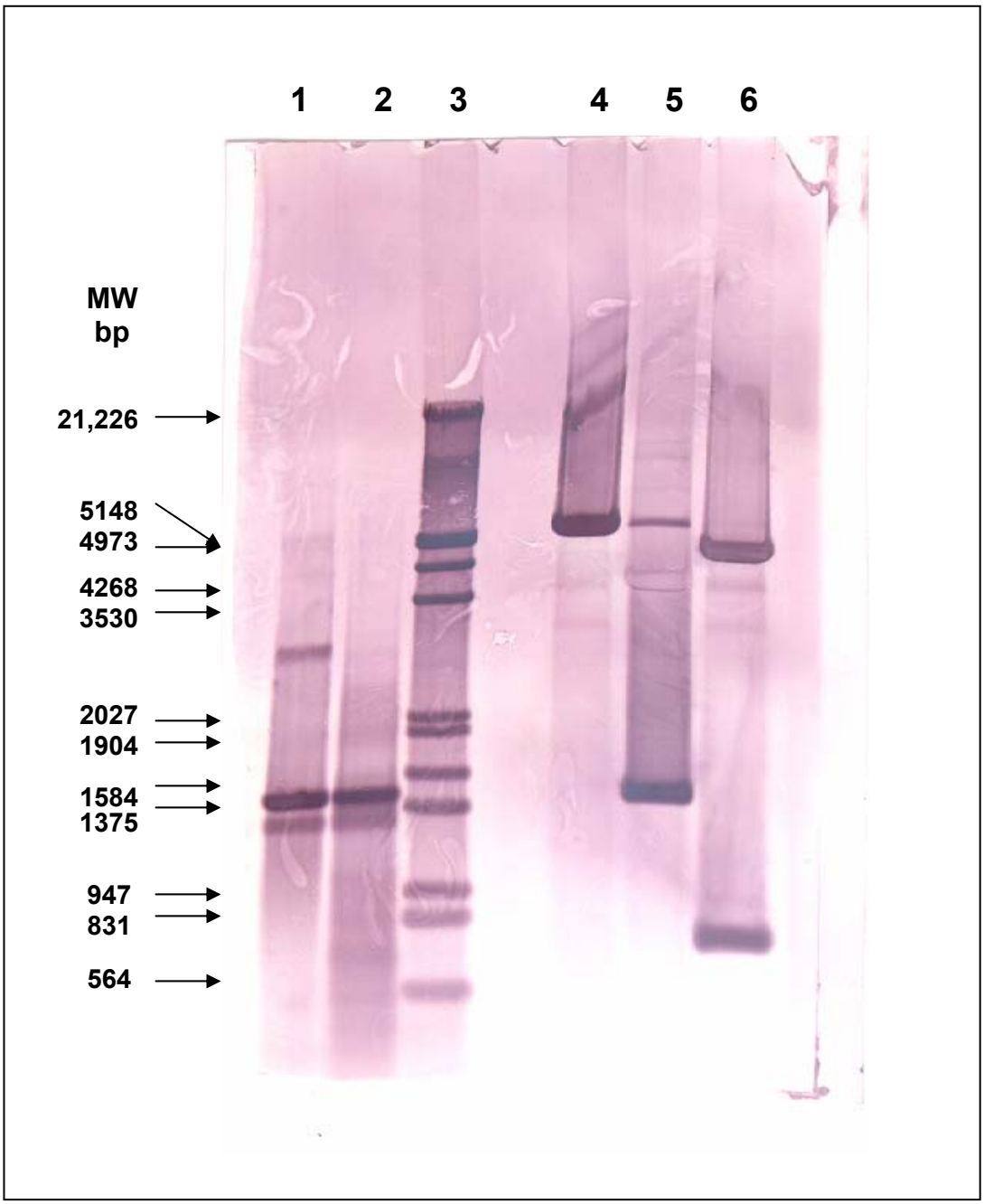


Figure 3.13 Southern blot/hybridization of recombinant plasmid (pDriveCS-31) with the DIG-labeled PCR amplified *Tv*CS gene fragment. Lane 1 and 2, positive control (PCR fragment); lane 3, molecular size marker (*Eco*RI/*Hin*dIII digested λ DNA); lane 4, *Bam*HI digested pDriveCS-31; lane 5; *Eco*RI digested pDriveCS-31 and lane 6, *Hin*dIII digested pDriveCS-31.

generated based on the nucleotide sequence of the putative gene available in the Genomic Database, using Clone 4 program as described in the Materials and Methods (Figure 3.14 and Table 3.1)

Restriction digestion pattern of the recombinant plasmid pDriveCS-31 on agarose gel is shown in Figure 15. The digestion profile is also shown schematically in the Figure 16. There is a single *Bam*HI site within one of the Vector's Multiple Cloning Site (MCS) and another site within the cloned PCR fragment at position 1471. Therefore *Bam*HI digestion gave rise to a 1472 bp large fragment and about 5 bp invisible small fragment. There is no *Eco*RI site in the insert, but two sites are present in the pDrive vector, each is being located in one of the flanking MCS sites. So, as a result of *Eco*RI digestion, insert was excised as a 1492 bp fragment. There is a single site for *Hin*dIII within *Tv*CS gene at the position 722 and another one in one of the MCS site. Thus, *Hin*dIII digestion produced two fragments of 740 bp (insert fragment) and 4587 bp (vector plus remaining gene fragment). There are 2 sites for *Ava*I inside the CS gene (at positions 1148 and 1420) and two more within the Vector (at positions 1598 and 1820). Therefore, *Ava*I digestion gave rise to 4 fragments of 2431bp, 2346bp, 272bp, and 222bp. There is a single *Pvu*I sites within *Tv*CS gene at position 685, but two more on the Vector itself (positions 52 and 515). Digestion with *Pvu*I produced three bands corresponding to the fragment sizes of 1054 bp, 885 bp and 3388bp. Single *Sph*I site is available within the Vector at position 278. Therefore, double digestion with *Pvu*I and *Sph*I yielded 4 fragments of 226 bp, 828 bp, 885 bp and 3388 bp.

Judging from the restriction digestion profile of the recombinant plasmid it is possible to conclude that *Tv*CS gene is ligated to the cloning vector in opposite direction with respect to the +1 position (i.e., Lac Z promoter/operator) of pDrive Vector.

The experimental data for restriction mapping is in well correlation with

Table 3.1 Cut site table for various restriction enzymes.

NONCUTTER ENZYMES	CUTTER ENZYMES	CUT SITES
AvrII	**HincII**	285-821
AccI	**BglII**	445
EcoRI	**Sau3A**	446-1053-1171-1472
EcoRV	**PstI**	550-829-922
KpnI	**AvaII**	1148-1420
MluI	**PvuII**	685
NheI	**ApaI**	902
NotI	**StyI**	911-1056-1341
SacI	**SauI**	1248
SalI	**BamHI**	1471
SmaI	**HindIII**	722
SnaBI	**BstXI**	1135
SphI	**SnaBI**	1311
XmaI	**ApaI**	902

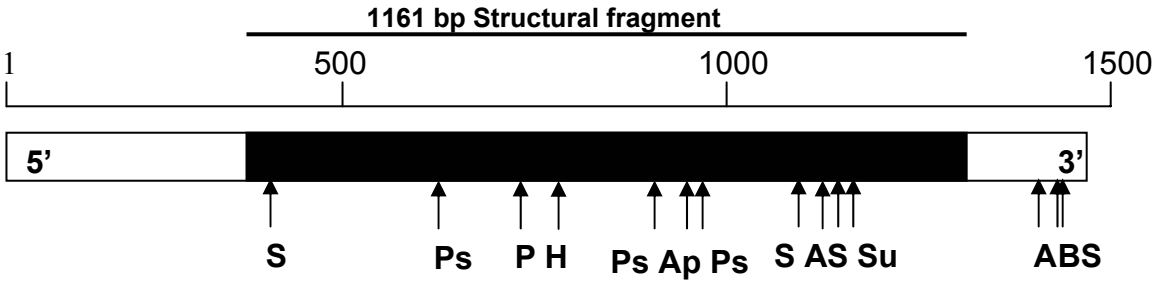


Figure 3.14 Restriction map of the PCR fragment (1476 bp). The area shown in dark corresponds to the coding region. A, *Ava*II; Ap, *Apa*I; B, *Bam*HI; H, *Hin*dIII; P, *Pvu*II; Ps, *Pst*I; Su, *Sau*I; S; *Sau*3AI.

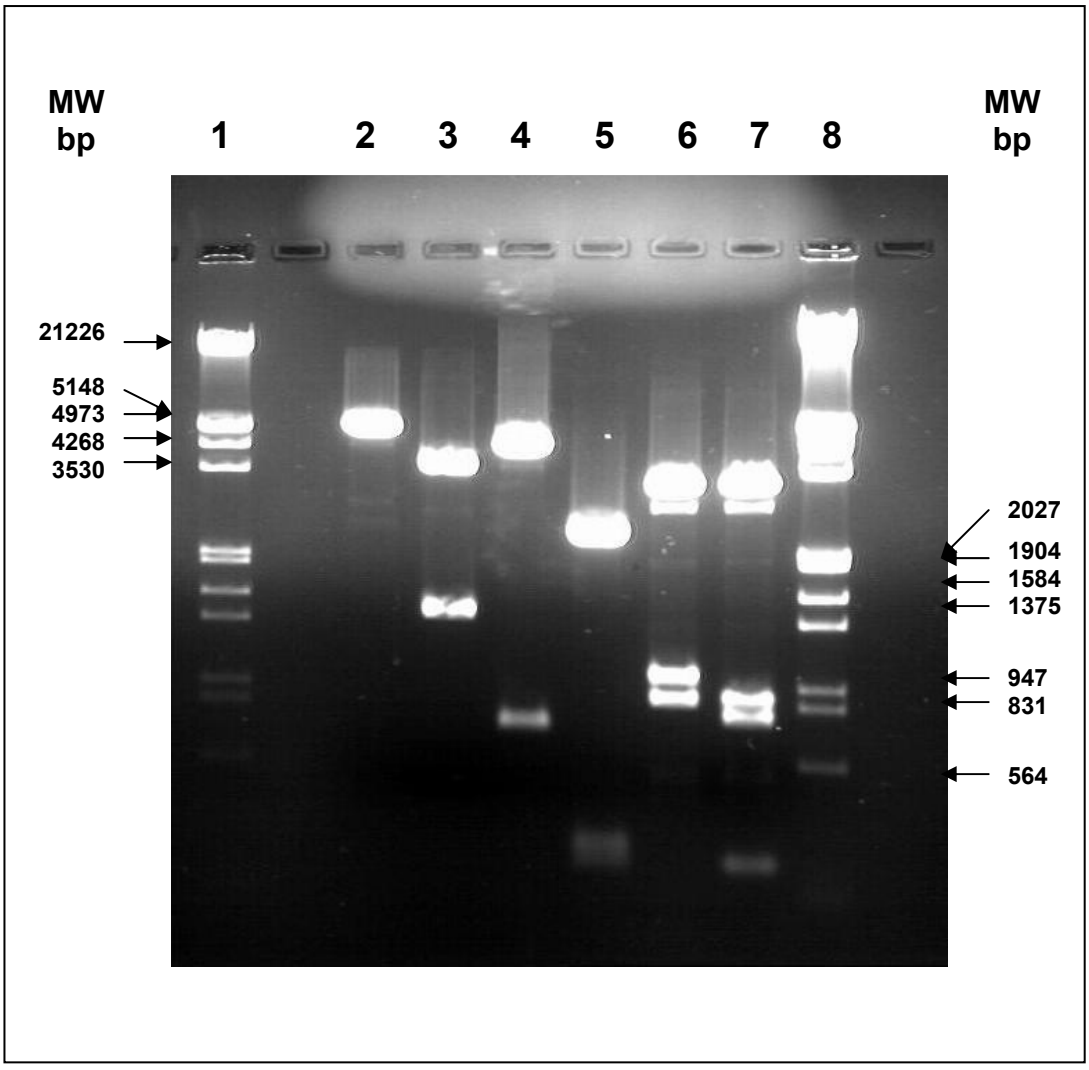


Figure 3.15 Restriction digestion profile of recombinant pDriveCS-31. Lane 1 and lane 8: molecular size markers (*Eco*RI/*Hin*dIII digested Lambda DNA), lane 2: *Bam*HI digestion, lane3: *Eco*RI, lane 4: *Hin*dIII, lane 5: *Ava*II, lane 6: *Pvu*II and lane 7: *Sph*I/*Pvu*II digestion.

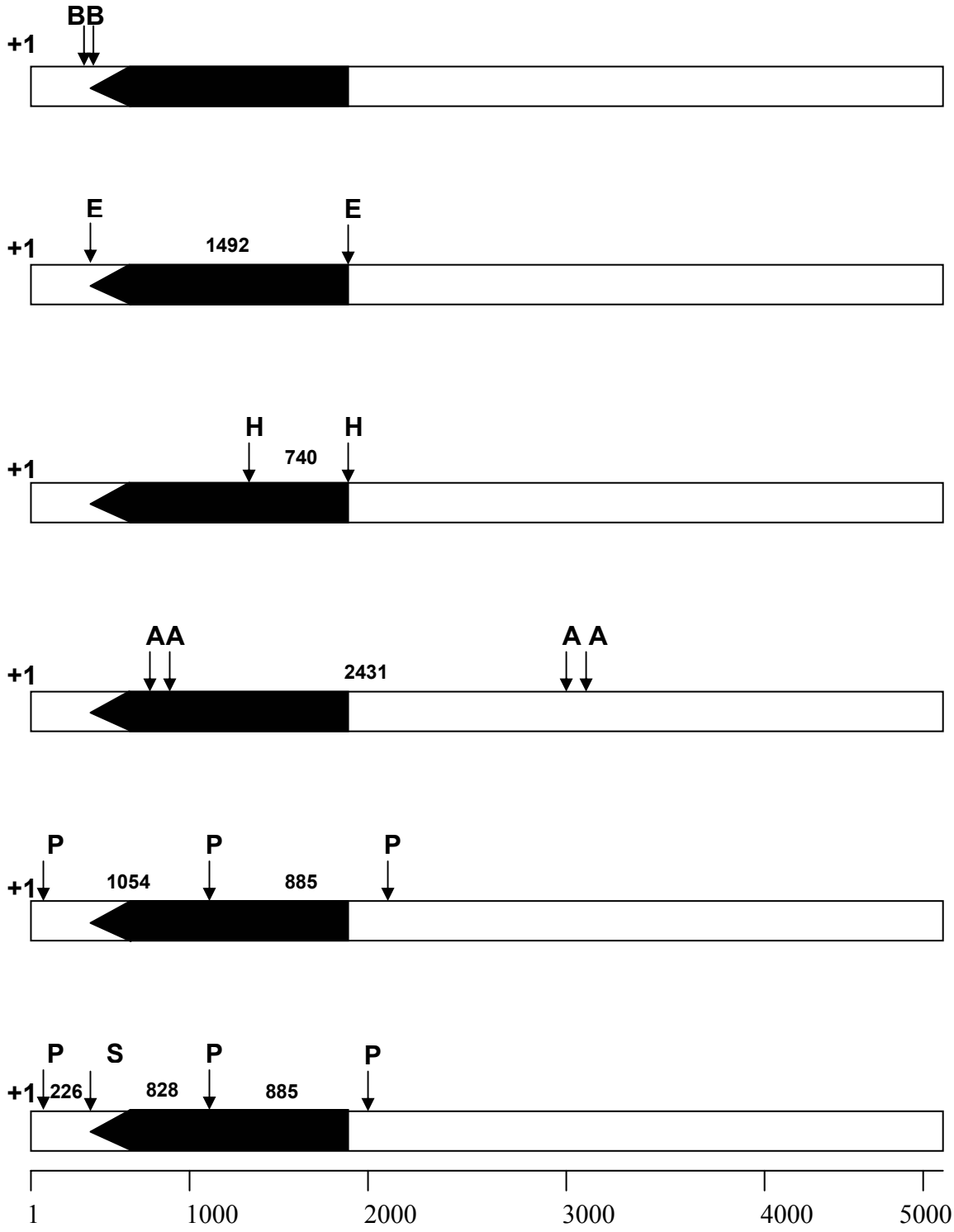


Figure 3.16 Restriction map of the recombinant pDrive vector containing the *Tp.volcanium* citrate synthase gene. The cloned *Tv*CS gene is marked by the dark bar, vector sequences are shown as open bars and the numbers refer to the base pairs between the various restriction sites. A, *Ava*II; B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; P, *Pst*I and S, *Sph*I.

the result derived from the mapping based on sequence analysis using software. This result clearly showed that we have cloned *Tv*CS gene in pDrive Vector.

3.8 Sequence Analysis of Citrate Synthase Gene and Its Promoter

The sequence of *Tv*CS gene was determined by automated sequencing Yale University, New Haven, Connecticut, USA. A 1362 bp region of 1476 bp cloned fragment was sequenced using the reverse and forward PCR primers with no conflict (Figure 3.17).

Sequence analysis of the amplified region revealed a 1161 bp open reading frame. The coding region has a G+C content of 42.38%. This is higher than the G+C content of genomic DNA which is about 38% (Segerer *et al*., 1988). The coding region (1161bp) encodes a protein of 386 amino acids and estimated molecular size of M_r= 43000 which is in agreement with the molecular mass determined by SDS/PAGE. When our sequence data was compared to that of putative *Tv*CS gene sequence available in the Archaic Data Base there are two nucleotide mismatches: one is G to A replacement at the position of +99 and the other is T to C replacement at the position of +198 in the sequence of cloned *Tv*CS gene. However, these two mismatches do not cause any change in the amino acid sequence.

At the -11 to -5 upstream region of the ATG start codon there is a putative 16S rRNA binding site 5'-GAGGTG-3'. This sequence shows high complementarity to the Shine-Dalgarno consensus sequence 5'GGAGGTGA-3' (Kawashima *et al*., 2002). Same sequence at similar position was also found to exist in the *Tp.acidophilum* citrate synthase gene (Sutherland *el al*., 1990). In *Tp.volcanium* genome a putative region for initiation of transcription was identified as 5'-AAAATTTATATA-3' (BoxA). In *Tv*CS gene, at -113 to -102 upstream position a related sequence of 5'-AATTTTAAATA-3' was

```
cagtgtaagaacgtatgtatgtattaagattgatttacggttcacacatctttcgcagcgcaataaattgaaacatg

atatgataatatttattcaattttaaatacatatttcacgtttagcgtatgcataatacaacgatatatgtagcg

                                                  1
aaccttaaaataacgtgaaacgatattagaaatgtagaggtgttttt ATGCCGGAAACAACTGAAGAAATTAGCA 28
                                                  M  P  E  T  T  E  E  I  S

AGGGTTTGGAAGATGTCAACATTAAATGGACAAGGCTAACAACTATAGATGGAAATAAAGGTATCCTTAGGTATG 103
K  G  L  E  D  V  N  I  K  W  T  R  L  T  T  I  D  G  N  K  G  I  L  R  Y

GGGGATACTCCGTAGAGGATATAATCAATAACGGTGCACAGGCAGAGGAGATTCAGTACCTTTTCTTATATGGAG 178
G  G  Y  S  V  E  D  I  I  N  N  G  A  Q  A  E  E  I  Q  Y  L  F  L  Y  G

AATTACCTAATGCACAGGAATTAAAGATCTTTAAAGAAAATGTGCAAAAAGGATATAACATTCCAGATTTCGTAA 253
E  L  P  N  A  Q  E  L  K  I  F  K  E  N  V  Q  K  G  Y  N  I  P  D  F  V

TAAATTCAATTAGGCAGCTTCCAAGAGAATCAGATGCAGTTGCAATGCAGATGGCTGCAGTTGCGTCTATGGCAG 328
I  N  S  I  R  Q  L  P  R  E  S  D  A  V  A  M  Q  M  A  A  V  A  S  M  A

CATCTGAAATAAAGTTCAAGTGGAATAAGGATACGGACAGGGATGTTGCTGCACAGATGATAGGCAGCATGTCCG 403
A  S  E  I  K  F  K  W  N  K  D  T  D  R  D  V  A  A  Q  M  I  G  S  M  S

CTATAACTGCAAATGTCTATAGGCATATCCTCGGTATGCCAGCTGAGAGGCCAAAACCGTCTGACAGCTATGCTG 478
A  I  T  A  N  V  Y  R  H  I  L  G  M  P  A  E  R  P  K  P  S  D  S  Y  A

AAAGCTTCCTTAAAGCTGCCTTTGGCAGGAATGTTACAAAGGAGGAAGTAGATGCAATAAACACTGCGCTAATAC 553
E  S  F  L  K  A  A  F  G  R  N  V  T  K  E  E  V  D  A  I  N  T  A  L  I

TTTACACAGACCACGAGGTTCCTGCGTCAACAACTGCAGGACTTGTTGCTGTATCAACATTGAGTGATATTTACT 628
L  Y  T  D  H  E  V  P  A  S  T  T  A  G  L  V  A  V  S  T  L  S  D  I  Y

CTGGCATAACTGCTGCTCTTGCTGCACTCAAGGGCCCGCTCCATGGTGGTGCTGCAGAAGCCGCAATAGCGCAGT 703
S  G  I  T  A  A  L  A  A  L  K  G  P  L  H  G  G  A  A  E  A  A  I  A  Q

TTGATGAAATAAAAGAGCCTTCAAACGTAGAGAAATGGTTCAATGACAACATAATAAACGGCAAGAAGAGGCTCA 778
F  D  E  I  K  E  P  S  N  V  E  K  W  F  N  D  N  I  I  N  G  K  K  R  L

TGGGATTCGGGCACAGAGTATATAAGACATATGATCCAAGGGCAAAGATATTCAAGGGAATAGCTGAGAATCTCT 853
M  G  F  G  H  R  V  Y  K  T  Y  D  P  R  A  K  I  F  K  G  I  A  E  N  L

CCAAGAATAATGCAGAGGTAAAGAAAATATACGATATAGCCACAAAGTTGGAGGACCTCGGTGTTAAACAGTTTG 928
S  K  N  N  A  E  V  K  K  I  Y  D  I  A  T  K  L  E  D  L  G  V  K  Q  F

GATCAAAAGGCATATACCCGAACACAGACTACTTCTCTGGCATAGTTTACATGTCCGTAGGTTTCCCCCTTAGGA 1003
G  S  K  G  I  Y  P  N  T  D  Y  F  S  G  I  V  Y  M  S  V  G  F  P  L  R

ACAACATATACACTGCTCTGTTTGCCCTATCAAGGGTTACTGGCTGGGAAGCACACTTCATAGAGTACGTAGAAG 1078
N  N  I  Y  T  A  L  F  A  L  S  R  V  T  G  W  E  A  H  F  I  E  Y  V  E

AACAGCAGAGATTAATAAGGCCAAGGGCTGTCTATGTTGGCCCTGCTGAGAGAAAATTTGTAAAACTACCTGACA 1153
E  Q  Q  R  L  I  R  P  R  A  V  Y  V  G  P  A  E  R  K  F  V  K  L  P  D

    1160
GAAAATAAaata
R  K
```

Figure 3.17 Nucleotide sequence and deduced amino acid sequence of the gene encoding citrate synthase (CS) from *Tp.volcanium*. The open reading frame is shown in bold. The nucleotides underlined with dots represent a putative ribosome binding site. The nucleotides underlined represent the putative promoter region.

found. This sequence as a part of promoter may play an important role in initiation of transcription. In *Tp.acidophilum* citrate synthase gene similar sequence of 5'-TTTAAT-3' was found at the position of -98 to-103. Also at -55 to -52 upstream region of *Tv*CS gene another putative promoter sequence of 5'-ATGT-3' (Box B) was observed. Same sequence was observed at upstream region of *Tp.acidophilum* citrate synthase gene promoter at position -75 to -72 (Sutherland *et al*., 1990).

According to the amino acid composition of *Tv*CS (Table 3.2) no Cys residues and very low number of Trp and His residues were observed. Total number of strong basic amino acid residues (Lys, Arg) and strong acidic amino acid residues (Asp, Glu) were found to be 47 and 48, respectively. The total numbers of hydrophobic residues (Ala, Ile, Leu, Phe, Trp, and Val) were 143, which is about the three fold of strong basic and acidic residues. The enzyme contains 91 polar amino acid residues (Asn, Gln, Ser, Thr, Tyr). At pH=7.0 the theoretical Isoelectric point was calculated to be 6.901. The codon usage profile in the *Tv*CS is given in Table 3.3.

3.8.1 Sequence Alignments of *Tp. volcanium* Citrate Synthase

A structure based sequence alignment of the *Tp.volcanium* citrate synthase (*Tv*CS) with CSs from *Tp.acidophilum* and Pig is shown in Figure 3.18. Although low sequence homology (18% for *Tv*CS and 19% for *Ta*CS) with the pig enzyme, of the 11 residues implicated in the catalytic activity of the pig enzyme, 9 are conserved in the archaeal citrate synthases. From the alignment data it is also evident that both archaeal citrate synthases are shorter at the N-terminus than pig enzyme, which could infer the enhanced packaging of these enzymes to withstand the high temperatures.

Sequence alignment of *Tv*CS with the other archaeal citrate synthases was also performed to compare sequence characteristics and to find out the

Table 3.2 Aminoacid composition of *Tp.volcainum* citrate synthase. The data is from the sequence given in Fig.3.13. Values are given as residues/molecule. Total number of amino acids = 386, Mr = 43,007.

Amino acid	Content
Ala	41
Arg	18
Asn	22
Asp	19
Cys	0
Gln	11
Glu	29
Gly	27
His	5
Ile	32
Leu	26
Lys	29
Met	9
Phe	16
Pro	16
Ser	21
Thr	19
Trp	4
Tyr	18
Val	24

Table 3.3 Codon usage of the citrate synthase gene of *Tp.volcanium*. Each triplet codon is preceeded by the single-letter code for the corresponding amino acid and followed by the frequency of its occurence in the *Tp.volcanium* citrate synthase gene.

A	GCA	18	Q	CAA	1	L	CUU	8	S	UCG	0
A	GCC	4	Q	CAG	10	L	UUA	4	S	UCU	5
A	GCG	4	E	GAA	16	L	UUG	3	.	UAA	1
A	GCU	15	E	GAG	13	K	AAA	14	.	UAG	0
R	AGA	5	G	GGA	8	K	AAG	15	.	UGA	0
R	AGG	13	G	GGC	9	M	AUG	9	T	ACA	10
R	CGA	0	G	GGG	2	F	UUC	10	T	ACC	0
R	CGC	0	G	GGU	8	F	UUU	6	T	ACG	1
R	CGG	0	H	CAC	3	P	CCA	6	T	ACU	8
R	CGU	0	H	CAU	2	P	CCC	1	W	UGG	4
N	AAC	10	I	AUA	22	P	CCG	4	Y	UAC	10
N	AAU	12	I	AUC	4	P	CCU	5	Y	UAU	8
D	GAC	7	I	AUU	6	S	AGC	4	V	GUA	10
D	GAU	12	L	CUA	4	S	AGU	1	V	GUC	3
C	UGC	0	L	CUC	6	S	UCA	7	V	GUG	1
C	UGU	0	L	CUG	1	S	UCC	4	V	GUU	10

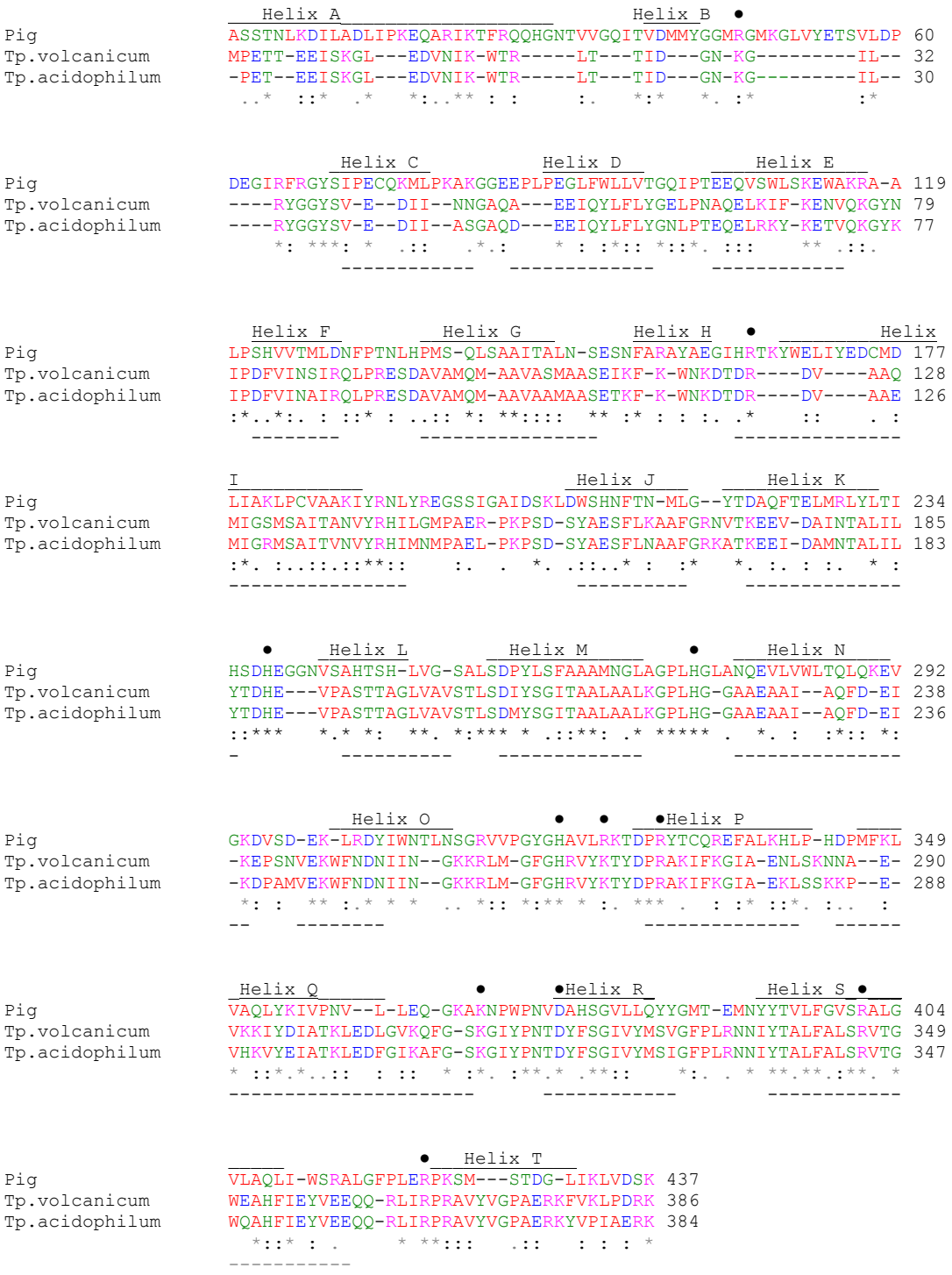


Figure 3.18 Clustal type structure-based sequence alignment of Pig heart (PigCS), *Tp.acidophilum* (*Ta*CS) and *Tp.volcanium* (*Ta*CS) citrate synthases. Helices are indicated by horizontal lines. (•) Amino acids implicated in catalysis in pig heart citrate synthase (Remington 1992), (*) residues conserved along all citrate synthases, (.) and (:) conserved substitutions of amino acids residues.

phylogenetic position of *Tp.volcanium* with respect to other archaeal species (Figure 3.19). Highest sequence similarity was found between *Tv*CS and *Ta*CS. The next closest species to *Tp.volcanium* were *S.solfataricus, Pyrococcus* and *S.tokodaii* with the alignment scores of 57, 40 and 39, respectively (Table 3.4). According to these results, *Pyrococcus furiosus, Sulfolobus* and *Thermoplasma* species form one phylogenetic group and diverge at one point from methanogens and halogens (Figure 3.20). When the similarity with respect to small and large domains of *Tp.acidophilum* CS is concerned small domain falls between the residues 225 and 235. There are more residues conserved in large domain than small domain.

One more aminoacid sequence comparison was performed by alignment of *Tv*CS sequences with eukaryotic and bacterial citrate synthase sequences (Figure 3.21). Among the 11 catalytic residues determined in PCS, 6 were conserved along the all lineages. As expected, the eukaryotic species (Human, pig, *A.niger, S.cerevisiae and N.crassa*), Gram (-) species (*E.coli, P.aeroginosa*), archaeal species (*Tp.volcanium, Tp.acidophilum, S.solfataricus* and *P.furiosus*) and Gram (+) bacteria (*B.subtilis*) formed four distinct groups according to the phylogenetic tree constructed by the comparison of citrate synthase sequences from these organisms (Figure 3.22 and Table 3.4). An interesting observation was the unexpected similarity (39%) between *B.subtilis* and *H.volcanii* citrate synthases that fall into same similarity group.

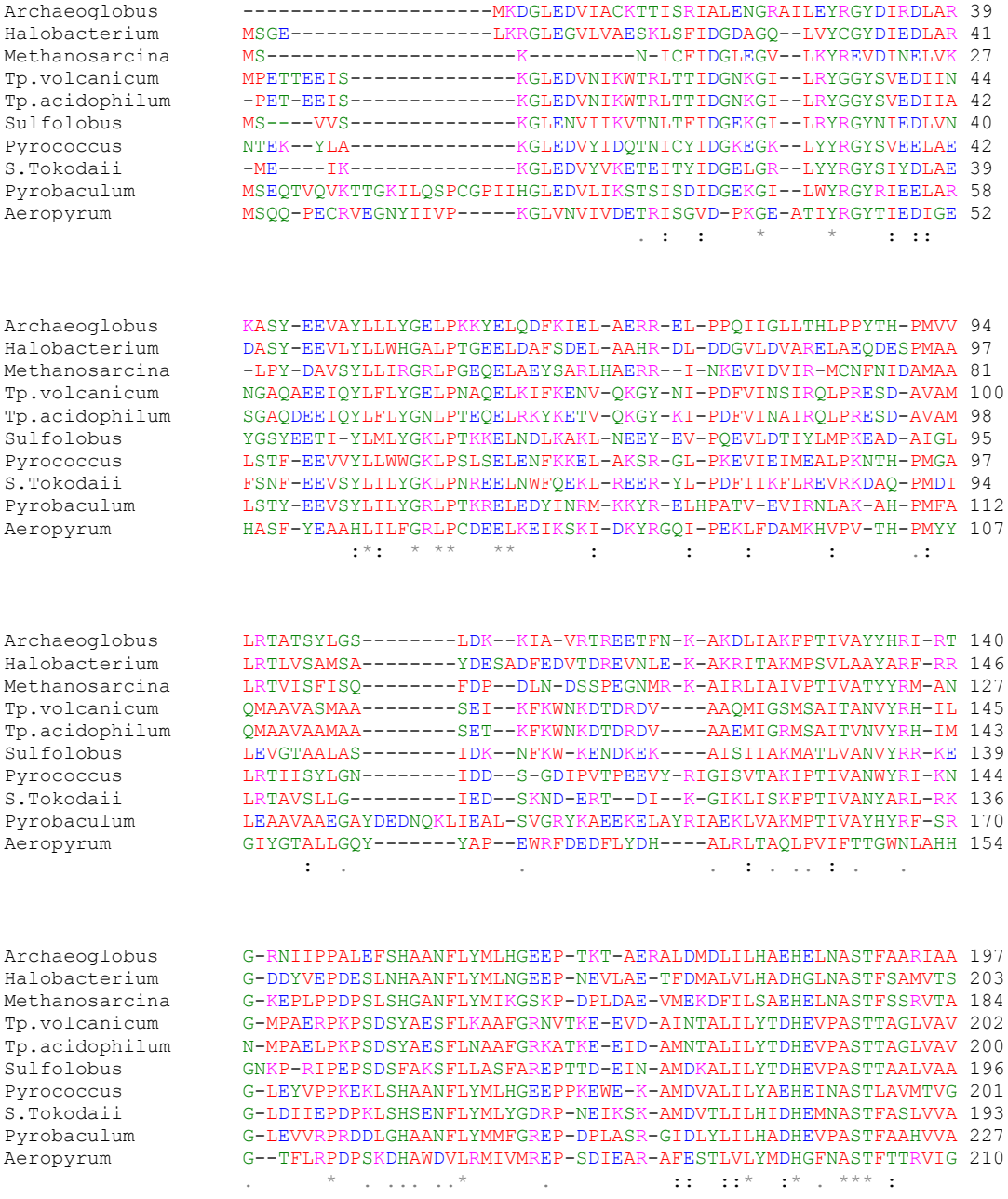


Figure 3.19 Clustal type multiple sequence alignments of amino acid sequences for archaeal citrate synthases. Source of the data: *Archaeaoglobus fulgidus, Aeropyrum pernix, Methanosarcina acetivorns*, *Sulfolobus tokodaii and Pyrobaculum aeropilum* are putative citrate synthases from EMBL sequence database; *Tp.acidophilum* (Sutherland *et al.,* 1990), *Pyrococcus furiosus* (Muir *et al.,* 1995), *Halobacterium volcanii* (Connaris *et al.,* 1998), *Sulfolobus solfataricus* (Connaris and West *et al.,* 1998). (*) conserved residues along all citrate synthases, (.) and (:) conserved substitutions of amino acids residues.

```
Archaeoglobus       STLADIYACVVAATGTLMGPLHGGAAQEVMRML-REVA-SPRR-------AEEYVK-RKI 247
Halobacterium       STLSDLYSAVTSAIGTLSGSLHGGANANVMRML-KDVDDSDMD-------PTEWVK-DAL 254
Methanosarcina      STMSDLYSAVVSGLCTLKGPLHGGARAEVMTML-DEVA-SPEN-------AEKFVL-EKI 234
Tp.volcanicum       STLSDIYSGITAALAALKGPLHGGAAEAAIAQF-DEIK-EPSN-------VEKWFNDNII 253
Tp.acidophilum      STLSDMYSGITAALAALKGPLHGGAAEAAIAQF-DEIK-DPAM-------VEKWFNDNII 251
Sulfolobus          STLSDMYSSLTAALAALKGPLHGGAAEEAFKQF-IEIG-DPNR-------VQNWFNDKVV 247
Pyrococcus          STLSDYYSAILAGIGALKGPIHGGAVEEAIKQF-MEIG-SPEK-------VEEWFF-KAL 251
S.Tokodaii          STFSDLYSSIVAGISALKGPLHGGANYEALKMF-KEIG-SPEK-------VNDYIL-NRL 243
Pyrobaculum         STLSDLYSSVAAAIAALKGPLHGGANEMAVRNY-LEIG-TPAK-------AKEIVEAATK 278
Aeropyrum           STLSDLYSAVAGGIAALKGPLHGGANEKAMEMFLDAMKKAEEKGVPLYDYIEEYIK-EKL 269
                    **::* *: : ..  :* *.:****   ..      :              . .

Archaeoglobus       EAGERIMGFGHRVYRGVMDPRAELLRYLAKRL-AAEG-STK-WFEISEAIAKA-AY--K- 300
Halobacterium       DRGERVAGFVHRVYN-VKDPRAKILGAKSEALGEAAG-DMK-WYEMSVAIEEY-IG--E- 307
Methanosarcina      SQREKIMGFGHRVYK-TYDPRGVIFKQLSKKLAESKG-DMH-WYTTAEAVENVVVR--EL 289
Tp.volcanicum       NGKKRLMGFGHRVYK-TYDPRAKIFKGIAENLSKNNA-EVKKIYDIATKLEDLGVK--QF 309
Tp.acidophilum      NGKKRLMGFGHRVYK-TYDPRAKIFKGIAEKLSSKKP-EVHKVYEIATKLEDFGIK--AF 307
Sulfolobus          NQKNRLMGFGHRVYK-TYDPRAKIFKKLALTLIERNA-DARRYFEIAQKLEELGIK--QF 303
Pyrococcus          QQKRKIMGAGHRVYK-TYDPRARIFKKYASKLGDKK------LFEIAERLERL-VEE-YL 302
S.Tokodaii          SNKQRIMGFGHRVYK-TYDPRARILKQYAKLLAEKEGGEIYTLYQIAEKVEEIGIK--YL 300
Pyrobaculum         PGGPKLMGVGHRVYK-AYDPRAKIFKEFSRDYVAKFG-DPQNLFAIASAIEQEVLSHPYF 336
Aeropyrum           ARKEKIMGFGHRIYK-LHDPRTDVAAKFVAKL--KDG-EF--WMKVLKKAEEVMWR---- 319
                        :: *  **:*.   ***  :

Archaeoglobus       -YK-K--LLPNVDFYSASVYANLGI-PD--DLFVNIFAMGRISGWTAHIIEQYEN-NRLI 352
Halobacterium       -EK-G--LAPNVDFYSASTYYQMGI-PI--DLYTPIFAVSRAGGWIAHVLEQYED-NRLI 359
Methanosarcina      VEKRGKPIYPNVDFYSGVIYKYMEI-PP--QLATSIFAIGRVSGWIAHCFDQYER-KKII 345
Tp.volcanicum       GSK-G--IYPNTDYFSGIVYMSVGF-PLRNNIYTALFALSRVTGWEAHFIEYVEEQQRLI 365
Tp.acidophilum      GSK-G--IYPNTDYFSGIVYMSIGF-PLRNNIYTALFALSRVTGWQAHFIEYVEEQQRLI 363
Sulfolobus          SSK-G--IYPNTDFYSGIVFYALGF-PV--YMFTALFALSRTLGWLAHIIEYVEEQHRLI 357
Pyrococcus          SKK-G--ISINVDYWSGLVFYGMKI-PI--ELYTTIFAMGRIAGWTAHLAEYVSH-NRII 355
S.Tokodaii          GPK-G--IYPNVDFFSSIVFYSLGFEP---DFFPAVFASARVVGWVAHIMEYIKD-NKII 353
Pyrobaculum         QQR-K--LYPNVDFWSGIAFYYMGI-PY--EYFTPIFAMSRVVGWVAHVLEYWEN-NRIF 389
Aeropyrum           -EK-K--IPANIDLYTAVLYYQLGI-PI--PMYTPIFAMGRVVGWSAHYIEQVLN-NKLI 371
                      :    :  * * ::.  :  : : *        :** .*  ** **  :     ::::

Archaeoglobus       RPRAEYVGEKEKKFIPL-SKR--- 372
Halobacterium       RPRARYTGEKDLDFTPV-DER--- 379
Methanosarcina      RPRA---------FM-L-DEC--- 355
Tp.volcanicum       RPRAVYVGPAERKFVKLPDRK--- 386
Tp.acidophilum      RPRAVYVGPAERKYVPIAERK--- 384
Sulfolobus          RPRALYVGPEYQEYVSI-DKR--- 377
Pyrococcus          RPRLQYVGEIGKKYLPI-ELRR-- 376
S.Tokodaii          RPKAYYKGEIGKKYIPI-DSR--- 373
Pyrobaculum         RPRACYIGPHDLQYIPL-EQR--- 409
Aeropyrum           RPTEKYVGPTGLKYQPI-EERCRK 394
                    **             :   : .
```

Figure 3.19 (continued)

Table 3.4 Sequence identities between archaeabacterial citrate synthases

Citrate synthase source	Percent Identity with citrate synthase from									
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
(1) *Tp.volcanium*	100	87	32	40	33	31	57	42	39	29
(2) *Tp.acidophilum*		100	32	40	32	34	58	43	39	30
(3) *M.acetivorans*			100	43	40	43	36	41	40	30
(4) *P.furiosus*				100	45	40	44	51	47	35
(5) *A.fulgidus*					100	48	40	46	44	34
(6) *H.volcanii*						100	36	41	43	34
(7) *S.solfataricus*							100	46	44	32
(8) *S.tokodaii*								100	46	34
(9*) Pyrobaculum*									100	34
(10 *)A.pernix*										100

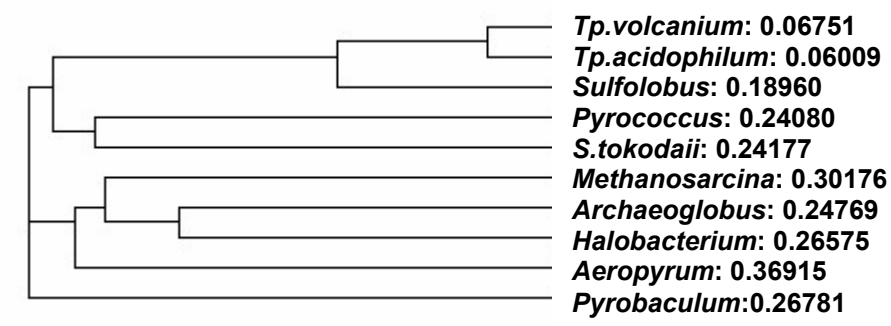


Figure 3.20 Phylogenetic tree constructed by citrate synthase amino acid sequence information of different archaea. The tree was constructed by as described in the Materials and Methods. The numbers indicate the calculated distances of organisms to the closest branching point.

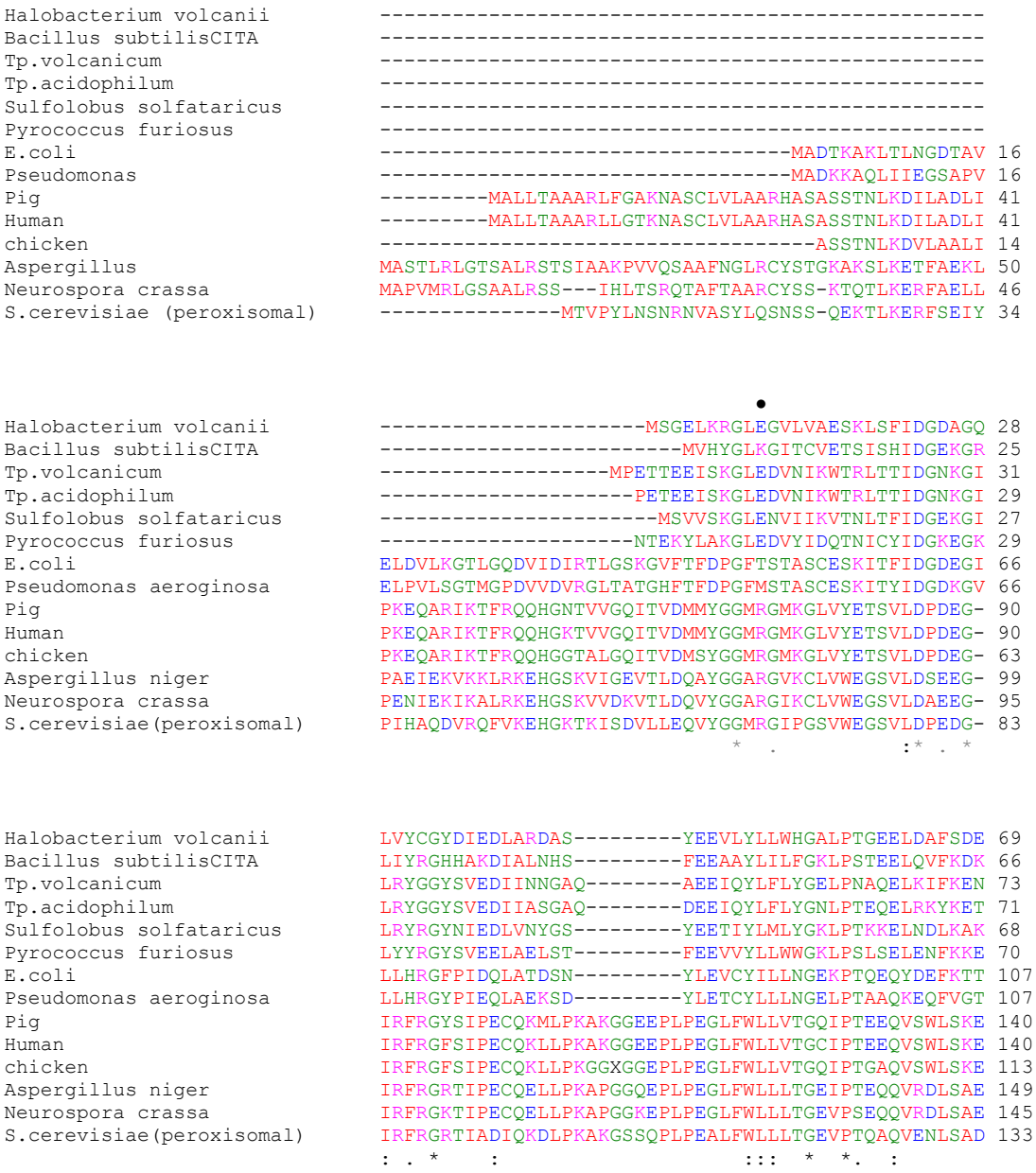


Figure 3.21 Clustal type multiple sequence alignments of amino acid sequences for archaeal citrate synthases. Source of the data: *E.coli* (Ner *et al*., 1983), *S.cerevisiae* (Rosenkrantz *et al*., 1986), Pig (Evans *et al*., 1988), *P.aeroginosa* (Donald *et al*., 1989), *Tp.acidophilum* (Sutherland *et al.,* 1990), Chicken (Saitoh *et al*., 1993), *B.subtilis* (Jin and Sonenshein, 1994), *N.crassa* (Ferea *et al*., 1994), *Pyrococcus furiosus* (Muir *et al.,* 1995), *Halobacterium volcanii* (Connaris *et al.,* 1998), *Sulfolobus solfataricus* (Connaris and Shauna *et al.,* 1998), Human (Goldenthal *et al*., 1998), and *A.niger* (Kirimura *et al*., 2000). (•) Amino acids implicated in catalysis in pig heart citrate synthase (Remington 1992) (*) conserved residues along all citrate synthases, (.) and (:) conserved substitutions of amino acids residues.

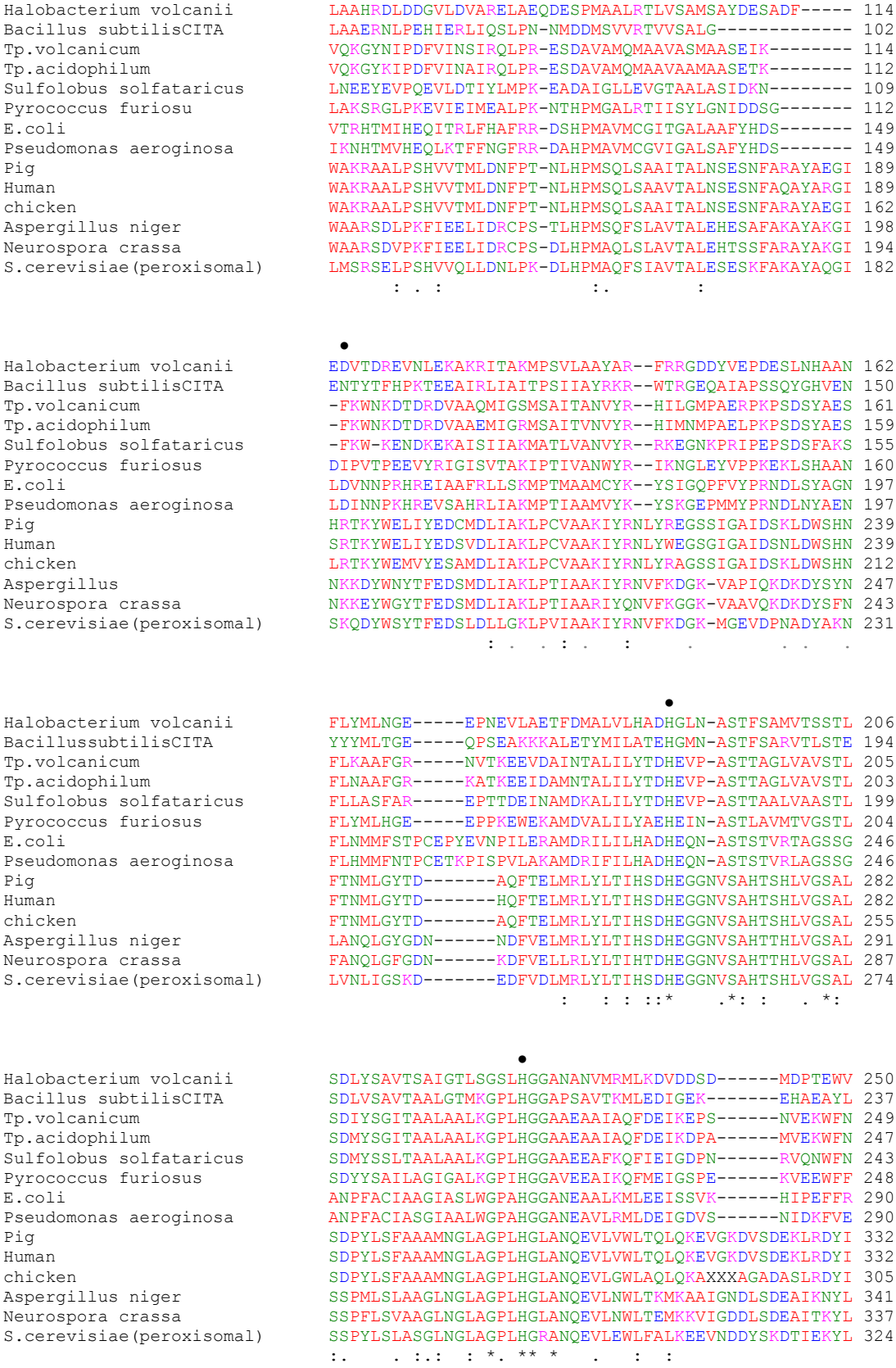


Figure 3.21 (continued)

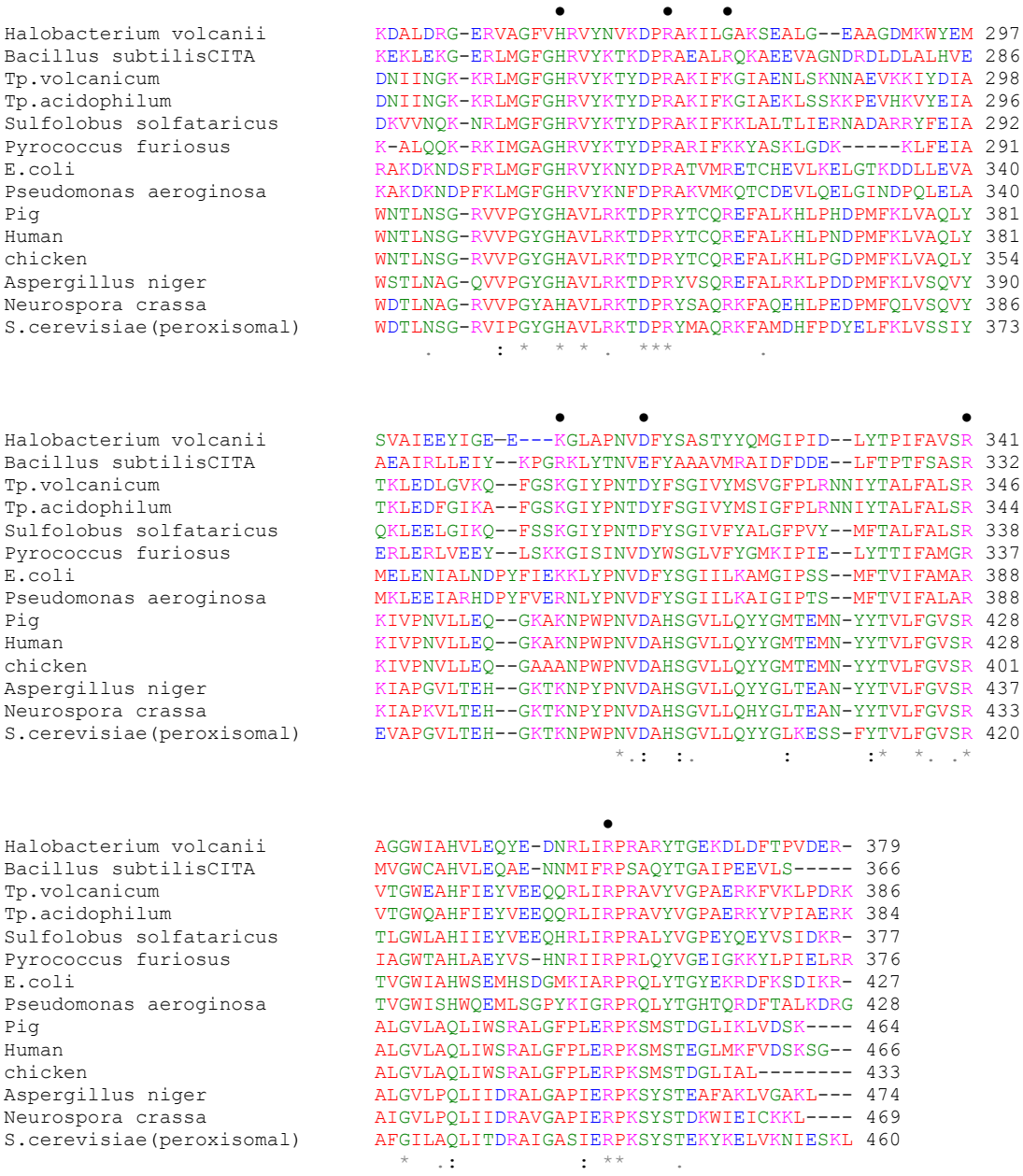


Figure 3.21 (continued)

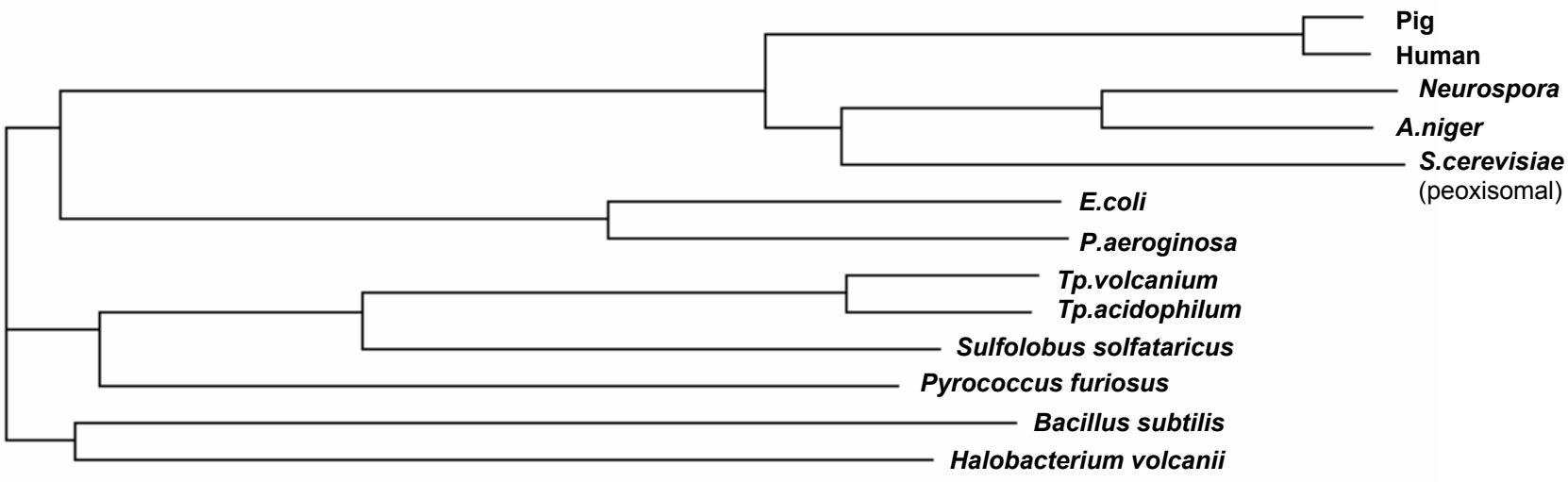


Figure 3.22 Phylogenetic tree which was constructed based on the homology between citrate synthase amino acid sequences from different source organisms: Eukaryotes; Human, Pig, *Neurospora crassa*, *Aspergillus niger*, *Saccharomyces cerevisiae*, Gram negative bacteria; *E.coli*, *Pseudomonas aeruginosa*, Archaea; *Tp.acidophilum, Tp.volcanium, Pyrococcus furiosus, Sulfolobus solfataricus*, *Halobacterium volcanii*, and a Gram positive bacteria *Bacillus subtilis*.

Table 3.5 Sequence identities between eukaryotic, eubacterial and archaeabacterial citrate synthases.

Citrate synthase source	Percent identity with citrate synthase from													
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
(1) Pig	100	18	22	19	24	21	95	61	91	24	59	20	55	19
(2) *Tp.volcanium*		100	30	87	40	57	18	19	18	31	18	30	20	31
(3) *E.coli*			100	30	34	33	21	20	21	31	21	69	19	30
(4) *Tp.acidophilum*				100	40	58	18	20	18	34	18	30	19	29
(5) *P.furiosus*					100	44	23	22	24	40	17	32	21	35
(6) *S.solfataricus*						100	21	20	21	36	19	34	20	33
(7) Human							100	61	89	22	59	21	56	19
(8) *A.niger*								100	62	22	80	20	62	20
(9) Chicken									100	22	60	20	56	19
(10) *H.volcanii*										100	20	32	19	39
(11) *N. crassa*											100	22	62	22
(12)*P.aeroginosa*												100	21	28
(13) *S.cerevisiae*													100	18
(14) *B.subtilis*														100

3.9 Over-expression and Purification of the Recombinant *Tp.volcanium* Citrate Synthase

The first step in the purification of recombinant *Tv*CS enzyme made use of high temperature stability of the *Thermoplasma* enzyme as compared to that of mesophilic host. Heat treatment of recombinant *E.coli* cell extract at 60^{o}C for 15 min resulted in precipitation of heat labile *E.coli* proteins, without any loss in the catalytic activity of *Tv*CS. Citrate synthase from crude extracts was purified in one-step by the use of bioaffinity chromatography on Reactive Red 120 Column. Citrate synthase was eluted from the column as a sharp single peak (Figure 3.23). The progress of the purification was followed by SDS-polyacrylamide gel electrophoresis (Figure 3.24). The apparent subunit molecular weight of about 43,500 Da was determined by SDS-PAGE (Figure 3.20). The steps of purification process are summarized in Table 3.5. After heat treatment almost all the activity was recovered. This step resulted in 1.1 fold purification, with a specific activity of 0,210U/mg. A purification of 1.2 fold was achieved after affinity chromatography, and 65.5% of the citrate synthase activity was obtained. The specific activity of the purified enzyme was 0,212 U/mg. The purified enzyme was homogenous as judged from SDS-polyacrylamide gel electrophoresis (Figure 3.24).

3.10 Kinetic Parameters

The purified enzyme followed classical Michaelis-Menten kinetics. The K_m values of 5.15 μM and 5.6 μM and V_{max} values of 1.74 μmol/ml/min and 1,6 μmol/ml/min were obtained from Lineweaver-Burk plots for acetyl-CoA and oxaloacetate, respectively (Figure 3.25 and Figure 3.26).

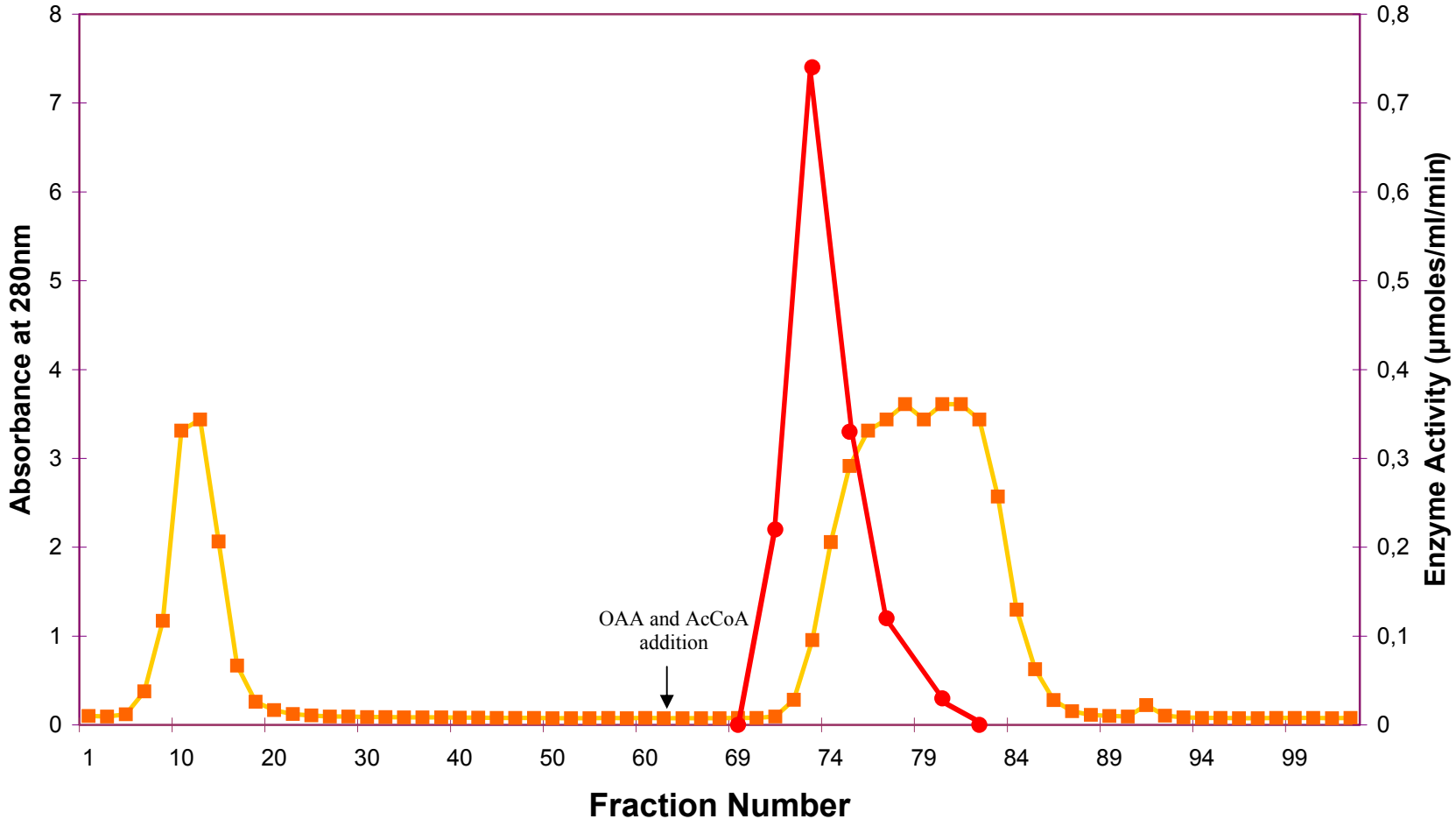


Figure 3.23 Purification of Recombinant *Tp.volcanium* Citrate Synthase by Affinity Chromatography on Matrix Gel Red A. OD280 ; Enzyme activity (µmoles/ml/min)

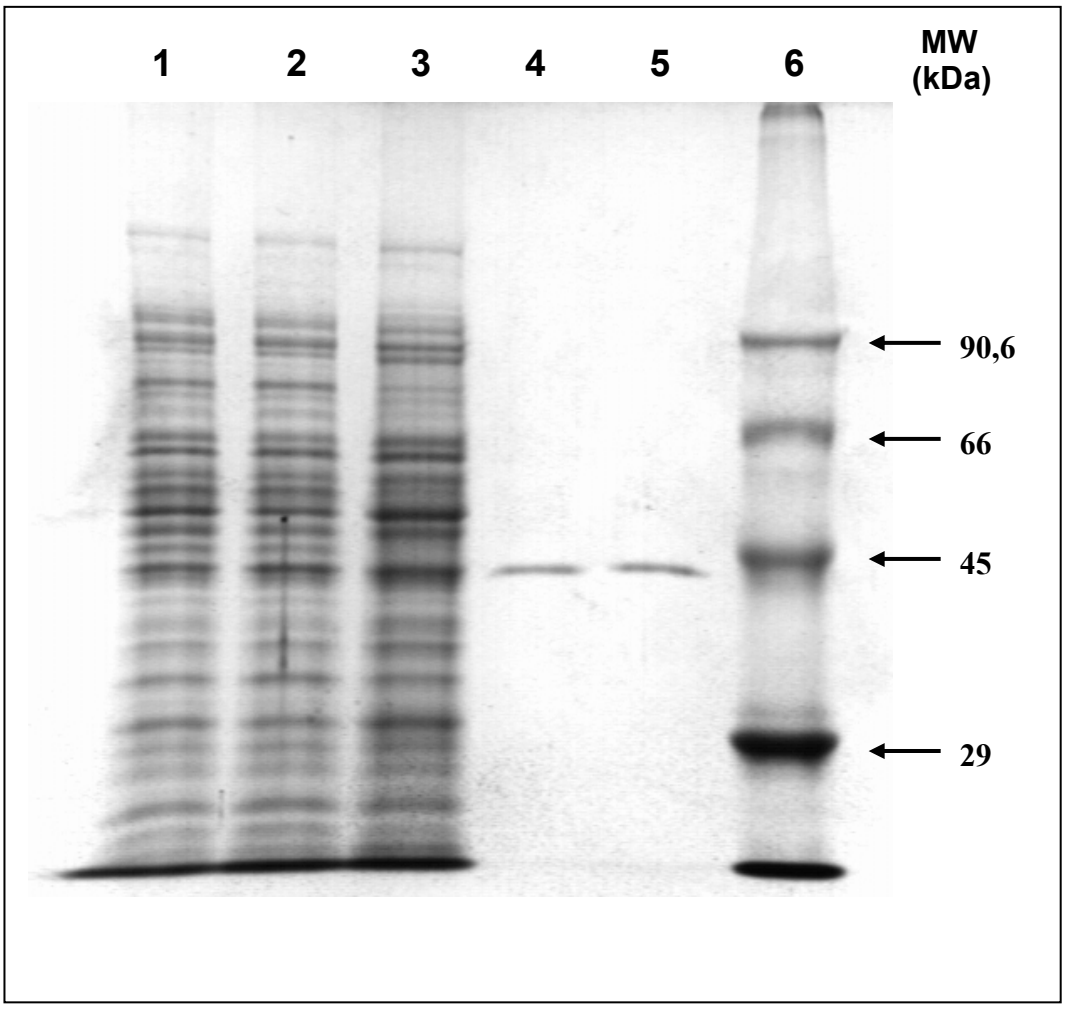


Figure 3.24 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis of recombinant *Tp.volcanium* citrate synthase. The samples are *E.coli* cell extract after heat treatment at 70^{o}C (lane 1), at 60^{o}C (lane 2), before heat treatment (lane3), pooled and concantrated fractions from Reactive Red 120 column (lane 4 and 5), and standard protein marker (lane 6). Standard proteins: Urease 90kDa (monomer); albumin, bovine, 66kDa (monomer); albumin, egg, 45 kDa; carbonic anhydrase, 29 kDa; α-lactalbumin, 14.2 kDa.

Table 3.5 Purification Steps of Recombinant *Tp.volcanium* Citrate Synthase

Step	Total volume (ml)	Protein concentration (mg/ml)	Total protein (mg)	Enzyme activity (U/ml)	Specific activity (U/mg)	Total activity (U)	Yield (%)	Purification (fold)
Cell-free Extract (Before Heat Treatment)	3.3	9.8	31,34	1.93	0.197	6.37	100	1
Heat treatement (at 60^{o}C,15 min)	3.0	9.2	27.60	1.93	0.210	6.30	98.9	1.1
Affinity Chromatography and concentration	2.3	8.3	19.09	1.76	0.212	4.05	65.5	1.2

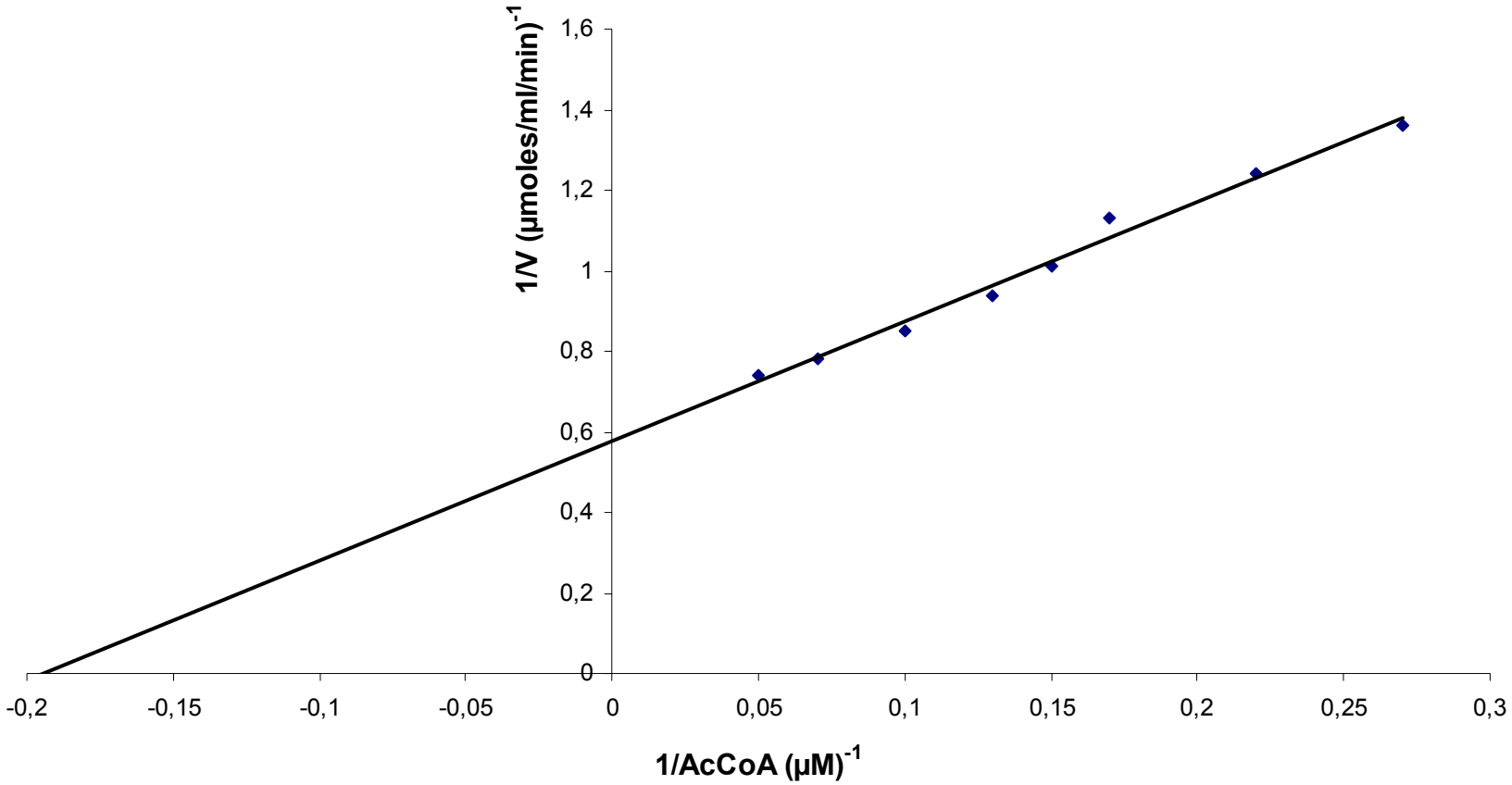


Figure 3.25 Double-reciprocal plot of citrate synthase activity *versus* the concentration of acetyl-CoA.

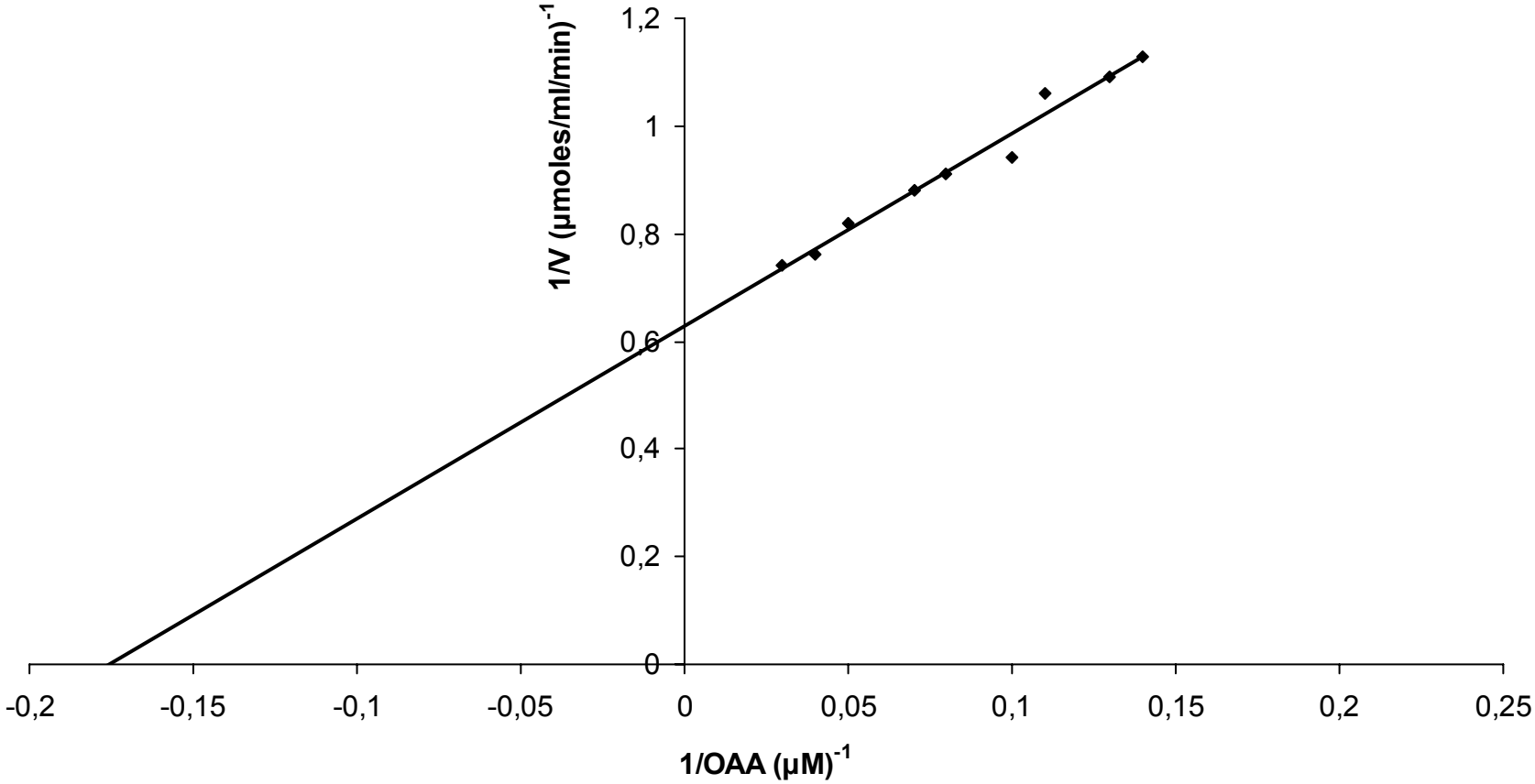


Figure 3.26 Double-reciprocal plot of citrate synthase activity *versus* the concentration of oxaloacetate.

3.11 Heat Treatment Effect

In order to study the effect of the heat treatment on the activity recombinant *Tv*CS, the purified enzyme was incubated at various temperatures, between 30°C and 95°C, for 60 min. The remaining enzyme activity was measured as described in the Materials and Methods and the activities are expressed as the percentage of the activity measured at 0 time (Figure 3.27). Enzyme activity has increased as the temperature for the heat treatment increased, up to 83°C. Enzyme activity increased about 3.5 fold, by incubation at 80°C for 60 min. About 71.5% and 85.5% of the CS activity was lost after 1h heating at 90°C and 95°C for 1hour, respectively.

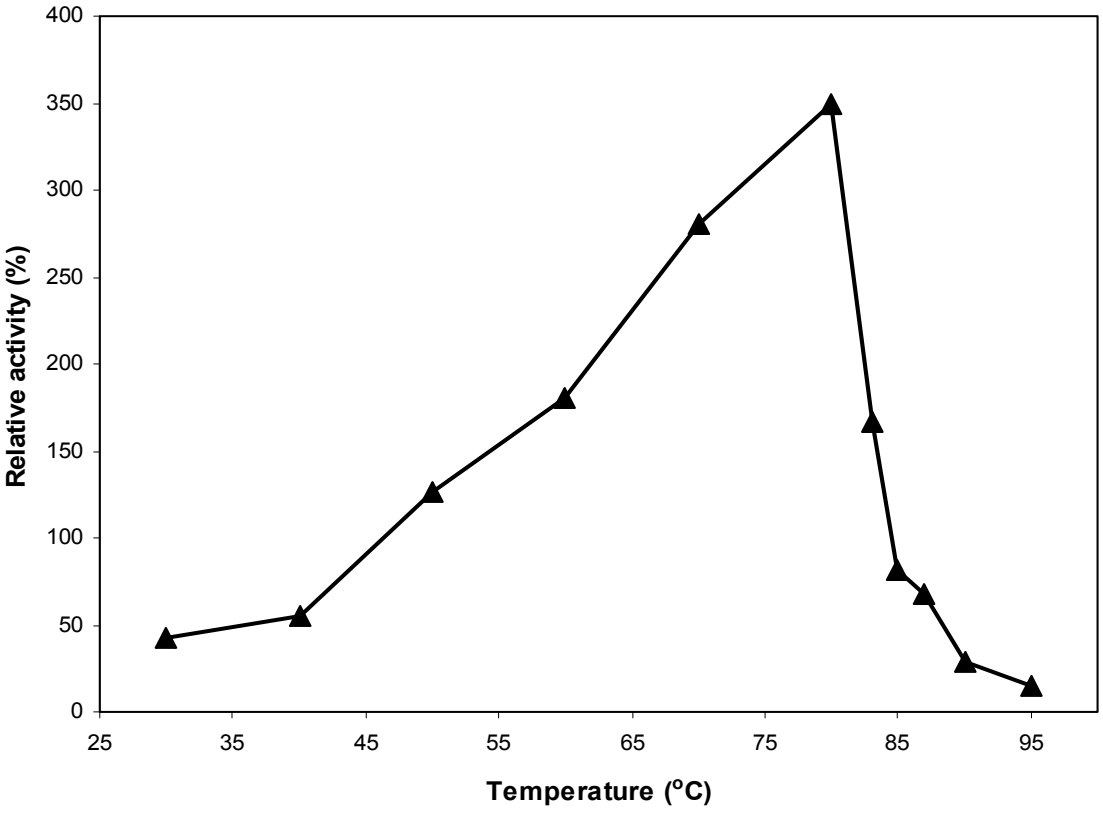


Figure 3.27 The residual activities at various temperatures after 1 h incubation were expressed as the percentage of the activity measured at 0 time.

CHAPTER 4

DISCUSSION

It is well known that thermophilic organisms live in environments with high temperature and their functional components, such as enzymes also resist high temperatures. Thermostable enzymes have potential applications in biotechnological processes. We have selected citrate synthase from a thermoacidophilic archaea *Thermoplasma sp.* as model enzyme to investigate mechanisms of thermal stability and activity. To this end, first we have used *Tp.acidophilum* citrate synthase for engineering thermostability *via* site-specific mutagenesis (Kocabıyık and Erduran 1998). As an alternative model, citrate synthase from the second species of *Thermoplasma* i.e. *Tp.volcanium* was selected to verify the thermal stability principles we have proposed for *Ta*CS. In this study, we have first cloned the gene coding for *Tp.volcanium* citrate synthase (*Tv*CS) using pDrive vector system, and then over-expressed this enzyme in *E.coli* which is a mezophilic host. For *Tv*CS expression *E.coli* system was selected due to its rapid growth and high biomass levels. Also heat sensitivity of the most *E.coli* proteins provided additional advantages in purification process as described in the Materials and Methods.

Activity and thermostability parameters of various archaea are shown in the Table 4.1. When their Michaelis constants are considered, catalytic features of recombinant *Ta*CS and *Tv*CS were almost same. However, thermal stability of *Tv*CS was considerably higher ($t_{1/2}$ at 90^{o}C, 45 min) as compared to that of *Ta*CS ($t_{1/2}$ at 90^{o}C, 2 min).

Table 4.1 Thermostability and catalytic properties of citrate synthases from *Tp.acidophilum* (*Ta*CS), *Tp.volcanium* (*Tv*CS), *S.solfataricus* (*Ss*CS) and *P.furiosus* (*Pf*CS).

Organism	K_m(μM)		V_{max} (μmoles/ml/min)		$t_{1/2}$ (min)	$T_{1/2}$ (^oC)	Reference
	K_mOAA	K_mAcCoA	V_{max}OAA	V_{max}AcCoA			
*Tv*CS	5.6	5.15	1.6	1.74	45	90	Present work
*Ta*CS	6.6	5.2	4.5	5.3	2	90	Kocabıyık *et al.* (2000)
*Ss*CS	5.1	2.3			8	95	Connaris *et al*. (1998)
*Pf*CS	5.1	1.7	150 ± 8 (μmoles/min/mg)		8	100	Muir *et al.* (1995)

There are two fundamental approaches to comprehend the basis of protein thermostability. The first approach is the comparative studies on homologous proteins from mezophiles and thermophiles. The second one is the engineering of proteins so that the effect of different combinations of amino acids or structural components could be identified (Querol *et al*., 1996). In this work, the sequence information, and alignments provided valuable data to commend on the stability features of *Tv*CS at the level of primary structure. Several of these sequence characteristics will be subject for our engineering studies that we are planning next.

For the thermostability analyses cloning of citrate synthase from *Tp. volcanium* (Optimum growth temperature=60^{o}C) fills a gap for the comparative approaches between the citrate synthases from mezophilic Pig (37^{o}C), moderately thermophilic *Thermoplasma acidophilum* (55^{o}C), thermophilic *Sulfolobus solfataricus* (85^{o}C), and hyperthermophilic *Pyrococcus furiosus* (100^{o}C). One structural property purposed to take part in thermal stability is compactness (Szilagyi *et al*., 2000). *Tv*CS structural gene codes for a protein of 386 amino acids. When sequence based alignment of *Tv*CS, PigCS and *Ta*CS (Figure 3.14) was examined it was observed that *Tv*CS is shorter in N-terminal, and Helix A of PigCS was not observed in

*Tv*CS. Same regions were also absent in *Ta*CS, *Ss*CS and *Pf*CS. This could indicate that when the growth temperature increases compactness of the enzyme also increases.

Thermostable enzymes tend to contain low number of thermolabile aminoacid residues (e.g. Cys, Asn and Gln; Three-letter amino acid code) (Russell and Taylor 1995). In table 4.2 percent amino acid contents of PigCS, *Ta*CS, *Tv*CS, *Ss*CS and *Pf*CS are given. *Tv*CS contains no Cys residue and about 3 percent Gln residues. *Ta*CS also contains no Cys and 3 percent of Gln residues. However, there is a two percent increase in Asn content in *Tv*CS as compared to *Ta*CS. Similarly, *Ss*CS has one percent more Asn content than *Ta*CS. The increase in this thermolabile residue in organisms with higher optimum growth temperature than *Tp.acidophilum* makes it obvious that thermostability is provided by more than one parameter as suggested by others (Querol *et al*., 1996, Jaenike and Böhm 1998, Kocabıyık and Erduran 1998, Szilagyi *et al*., 2000, Bell *et al*., 2002). Also this result could bring about a question of how these amino acids are prevented from chemical modifications (e.g. deamination) at high temperatures. One other difference with respect to amino acid content of *Tv*CS is the one percent reduction in the His content compared to *Ta*CS. As will be discussed later, although His content of *Tv*CS is about half of that of the *Ta*CS, all catalytically important His residues as indicated by Remington (1992), were conserved. Leu is the most abundant amino acid found in PigCS. In moderately thermophilic *Ta*CS the percentage of Leu residues is about half of that of PigCS. However, when the optimum growth temperature of the organisms increases the number of the Leu residues is also increases and in *Ss*CS and *Pf*CS their number reaches about 10 percent, which is very close to the percentage of Leu residues in PigCS (12%). In thermophilic *Tv*CS and *Ta*CS there is an increase in number of Ala residues in contrast to the decrease in Leu residues. It is well known that Ala residues have a helix-stabilizing effect (Connaris *et al*., 1998). In addition, by increasing the hydrophobicity of the enzyme, these Ala residues could provide a thermal

Table 4.2 Percentage amino acid composition of citrate synthase from pig (optimum temperature 37°C), *Tp.acidophilum* (optimum temperature 55°C), *Tp.volcanium* (optimum growth temperature 60°C), *Sulfolobus solfataricus* (optimum temperature 85°C), and *Pyrococcus furiosus* (optimum temperature 100°C)

Enzyme source	Amino acid (%)																			
	Cys	Gly	Pro	Ala	Val	Leu	Ile	Phe	Met	Trp	Tyr	His	Ser	Thr	Asn	Gln	Lys	Arg	Asp	Glu
Pig	1	7	5	7	6	12	4	3	3	2	4	3	7	5	4	4	6	4	5	5
Tp.acidophilum	0	7	4	11	6	6	8	4	3	1	5	2	5	5	4	3	8	5	5	8
Tp.volcanium	0	7	4	11	6	7	8	4	2	1	5	1	5	5	6	3	8	5	5	8
S.solfataricus	0	6	4	10	6	10	7	5	2	1	6	1	5	5	5	2	8	5	5	8
P.furiosus	0	8	4	7	6	10	9	3	2	2	7	2	5	4	3	1	9	4	3	10

resistance in polar solvent environment, thermodynamically (Kocabıyık and Erduran 1998). On the other hand the Ala content in hyperthermophilic *Pf*CS reduced drastically compared to those CSs from thermophilic organisms. In general hyperthermophilic *Pf*CS contains similar amount of polar and charged amino acid content with PigCS but the reduction in the thermolabile amino acids is more obvious. A very significant increase in the content of isoleucine residues was observed in all of the thermostable citrate synthases. This dramatic increase could be due to the improved packaging provided by these residues (Russell and Taylor 1995).

Citrate synthase is a dimeric enzyme, and therefore, subunit-subunit interaction is another important aspect of thermal-stability. Moreover, as both monomers are required for catalytic activity, the integrity of the monomers through interface interactions plays a crucial role in the thermal-activity (Kocabıyık and Erduran 1998, Querol *et al*., 1996). As indicated earlier, in *Tv*CS there is an increase in helix-stabilizing Ala residues in helix G. This could result in an increase in hydrophobic interactions, which could improve thermostability. In contrast, in *Tv*CS, there is a Ser substitution for Ala in the helix F and G in the positions 87 and 107, respectively when compared to *Ta*CS. However, thermostability analysis showed that recombinant *Tv*CS was more thermostable ($t_{1/2}$=45 min at 90^{o}C) compared to wild type recombinant *Ta*CS ($t_{1/2}$=2.0 min at 90^{o}C) (Kocabıyık and Erduran 1998). As reported by Kocabıyık and Erduran 1998, Ala97→Ser substitution on helix G on the subunit interface of *Ta*CS, resulted in an unexpected increase in the thermostability and catalytic activity, although this mutation lowers the hydrophobicity. This might indicate the importance of some other parameters in thermalstabilization (e.g. better hydrogen bonding, helix dipole stabilization and improved electrostatic interactions) beside the hydrophobic effect (Querol *et al*., 1996, Kocabıyık and Erduran 1998). Intrinsic biochemical factors such as amino acid content, the positions of some key aminoacids and small differences within the structure such as the number and sizes of cavities, hydrophobicity of surface area, all collaborate to

provide thermal stability and thermal activity in thermoresistant citrate synthases.

Although there is low sequence homology between citrate synthases from thermophiles and mezophiles the overall structure was found to be very similar (Bell *et al*., 2002). This structural similarity was also enhanced by the conservation of purposed catalytic amino acids (Remington 1992). Nine out of 11 catalytic residues were found to be conserved between *Tv*CS, *Ta*CS and PigCS. In fact remaining two catalytic residues were also conserved since two catalytic arginine residues in PigCS were replaced in *Ta*CS and *Tv*CS with lysine that has very similar biochemical property. Structural conservation is well implied in the structure-function relations of the enzymes. The conservation of catalytic amino acids along with structural conservation may indicate that there should be some basic similarities between the enzymes to catalyze same reaction.

Homology alignments between experimental and putative archaeal citrate synthases demonstrate that the phylogenetic positions of thermophilic archaea diverged at one point from methanogens and halogens. This could imply the presence of some distinct parameters, which provides these organisms with thermophilicity. When the similarity with respect to small and large domains of *Tp.acidophilum* CS, it is observed that the percentage of conserved residues in large domain is more than that of small domain. This seems to contradict the opinion that as the small domain is involved in catalytic action it is conserved more (Bell *et al*., 2002). However, for the archaeal species the requirement of stability due to environmental factors, especially high heat, could result in the conservation of residues, which are located in large domain and provide thermostability. Thus, the structural organization of the large domain may provide an important insight for future enzyme design.

As found in all organisms belonging to three Kingdoms: Eukarya*,*

Bacteria and Archaea, citrate synthase could also be a valuable enzyme for the determination of the phylogenetic positions of the source organisms. Thus, the sequence comparisons of citrate synthase sequences can be utilized so that the evolutionary relations between the source organisms may be identified. The first phylogenetic tree we have constructed by the comparison of CS sequences from different archaea (Figure 3.20) revealed that thermophilic archaea form a distinct group compared to methanogens and halogens. *Tp.volcanium* (Optimum growth temperature=60^{o}C) was found to be the most closely related to *Tp.acidophilum* (Optimum growth temperature=55^{o}C). The next organism found to be more close to *Tp.volcanium* and *Tp.acidophilum*, was *S.solfataricus* (Optimum growth temperature=85^{o}C). *Pyrococcus* (Optimum growth temperature=100^{o}C) and *S.tokodaii* (Optimum growth temperature=90^{o}C) could be the most ancient and the least related thermophiles to these three archaea: *Tp.volcanium*, *Tp.acidophilum* and *S.solfataricus*. From these observations it might be conferred that the degree of relationship between thermophilic archaea is well correlated with the optimum growth temperature of the source organism. This result differs from the one implicated by the phylogenetic tree of Woese *et al* (1990), where *Pyrococcus* forms a distinct linage group.

We have constructed second phylogenetic tree (Figure3.22) by comparing the citrate synthase aminoacid sequences from organisms representing all three Kingdoms. As expected Archaea, Eukarya Gram-negative bacteria and Gram-positive bacteria form four distinct phylogenetic groups. One interesting phenomena observed from such an alignment was the high similarity (%40) between the CS from *H.volcanii* and *B.subtilis* (Table 3.5). The phylogenetic analysis performed by Muir *et al*. (1995) indicated that the archaea were more close to Gram positive bacteria. Moreover, *Tp.acidophilum* and *P.furiosus* were found to be in the same phylogenetic group together with the Gram positive bacteria. This classification profile is not an unexpected one since citrate synthases from

both archaea and Gram positives are homodimeric and have the same regulation pattern.

In conclusion, the cloning and partial characterization of *Tv*CS provided new insights into biochemistry of extremozymes in terms of our understanding of the enzyme stability. Based on this information, a possible strategy could be the *in-vitro* genetic recombination to prove the mechanisms proposed in the present work. In this context, cloning of citrate synthase and determination of its nucleic acid sequence data provided basic information and material for our future work to engineer new protein products by domain swapping and direct mutagenesis. On the other hand, phylogenetic analyses based on sequence comparisons of CSs from various organisms and Archaea revealed critical features regarding their evolutionary and functional relatedness.

REFERENCES

Alter, G.M., Casazza, J.P. Zhi, W., Nemeth P., Srere, P.A., and Evans, C.T. (1990). Mutation of Essential Catalytic Residues in Pig Citrate Synthase. Biochemistry 29,7557-7563.

Bell, G.S., Russell, R.J., Connaris, H., Hough, D.W., Danson, M.J., Taylor, G.L. (2002) Stepwise adaptations of citrate synthase to survival at life`s extremes. Eur.J.Biochem. 269, 6250-6260.

Castro, I.M., Fietto, J.L.R., Vieira, R.X., Tropia M.J.M., Campos, L.M.M., Paniago, E.B., Brandao, R.L. (2000). Bioleaching of zinc and nickel from silicates using *Aspergillus niger* cultures. Hydrometallurgy 57, 39-49.

Connaris H., Chaudhuri J.B., Danson M.J., and Hough D.H (1999). Expression, Reactivation, and Purification of Enzymes from *Haloferax volcanii* in *Escherichia coli*. Biotechnology and Bioengineering 64(1), 38-45.

Connairs, H., West, S.M., Hough, D.W., Danson, M.J. (1998) Cloning and overexpression in *Escherichia coli* of the gene encoding citrate synthase form the hyperthermophilic Archaeon *Sulfolobus solfataricus*. Extremophiles 2, 61-66.

Crolla, A., Kennedy, K.J., (2001). Optimization of citric acid production from *Candida lipolytica* Y-1095 using *n*-paraffin. Journal of Biotechnology 89, 27-40.

Danson, M.J., Black, S.C., Woodland, D.L. and Wood, P.A (1985). Citric acid cycle enzymes of the archaebacterial: citrate synthase and succinate thiokinase. FEBS 179(1), 120-124.

Donald, L.J., Molgat, G.F., Duckworth, H.W. (1989). Cloning, Sequencing and Expression of the Gene for NADH-Sensitive Citrate Synthase of *Pseudomonas aeruginosa*. Journal of Bacteriology 171(10), 5542-5550.

Evans,C.T., Owens,D.D., Sumegi,B., Kispal,G., Srere,P.A. (1988) Isolation, nucleotide sequence, and expression of a cDNA encoding pig citrate synthase. Biochemistry 27:4680-4686.

Ferea, T., Contreras, E.T., Oung, T., Bowman E.J., Bowman B.J. (1994). Characterization of the *cit*-1 gene from *Neurospora crassa* encoding the mitochondrial form of citrate synthase. Mol. Gen. Genet. 242, 105-110.

Henneke, C.M., Danson, M.J., Hough, D.W., Osguthorpe, J. (1989) Sequence alignment of citrate synthase proteins using a multiple sequence alignment algorithm and multiple scoring matrices. Protein Engineering 9(8), 597-604.

Jaenicke, R., Böhm, G. (1998). The stability of proteins in extreme environments. Curr. Op. Str. Biol. 8, 738-748.

Jianlong, W., Xinmin, Z., Decai, D., Ding, Z. (2001). Bioadsorption of leas(II) from aqueous solution by fungal biomass of *Aspergillus niger*. Journal of Biotechnology 87, 273-277.

Jin, S., Sonenshein A.L. (1994). Identification of Two Distinct Bacillus subtilis Citrate Synthase Genes. Journal of Bacterology 176(15), 4669-4679.

Kawashima T., Amano, N., Koike, H., Makino, S., Higuchi, S., Ohya., Y.K., Watanabe, K., Yamazaki, M., Kanehori, K., Kawamoto, T., Nunushiba, T., Yamamoto, Y., Aramaki, H., Makino, Kozo., Suzuki, M. (2000) Archaeal adaptation to higher temperatures revealed by genomic sequence of *Thermoplasma volcanium*. PNAS 97(26), 14257-14262.

Kirimura, K. Yoda, Masashi, Ko, I., Oshida, Y., Miyake, K., Usami, S. (1999). Cloning and Sequencing of the Chromosomal DNA and cDNA Encoding the Mitochondrial Citrate Synthase of *Aspergillus niger* WU-2223L. Journal of Bioscience and Bioengineering 88 (3), 237-243.

Kocabıyık S., Erduran İ., Russel R.J.M., Danson, M.J., and Hough, D.W. (1996). The Effect of Cysteine-43 Mutation on Thermostability and Kinetic Properties of Citrate Synthase from *Thermoplasma acidophilum*. Biochemical and Biophysical Research Communications 224(1), 224-228.

Kocabıyık, S., Erduran, İ. (1998). Amino Acid Substitutions in the Subunit Interface Enhancing Thermostability of *Thermoplasma acidophilum* Citrate Synthase. Biochem, Biophys. Res. Commun. 249, 566-571.

Kocabıyık, S., Erduran, İ. (2000). The Effect of Valine Substitution for Glycine in the Dimer Interface of Citrate Synthase form *Thermoplasma acidophilum* on Stability and Activity. Biochem, Biophys. Res. Commun. 275, 460-465.

Kumar, S. and Nussinow, R. (2001). How do thermophilic proteins deal with heat? Cell. Mol. Life Sci. 58, 1216-1233.

Laemmli, U.K. (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature 227, 680-685.

Lesk, M., Chothia, C. (1984). Mechanisms of Domain Closure in Proteins. J. Mol. Biol. 174, 175-191.

Muir, J.M., Russell, R.J.M., Hough, D.W., Danson, M.J., (1995). Citrate synthase from the hyperthermophilic Archaeon, *Pyrococcus furiosus*. Protein Engineering 6, 583-592.

Mulholland, A.J., Richards W.G. (1998). A model of the condensation step in the citrate synthase reaction. Journal of Molecular Structure 427,175-184.

Ner, S.S., Bhayana,V., Bell, A.W., Giles, I.G., Duckworth, H.W., Bloxham, D.P., (1983). Complete sequence of the gltA gene encoding citrate synthase in *Escherichia coli*. Biochemistry 22:5243-5249.

Querol, E., Perez-Pons, J.A., Mozo-Villarias, A. (1996) Analysis of protein conformational characteristics related to thermostability. Protein Engineering 9(3), 265-271.

Remington, S.J. (1992). Structure and Mechanism of Citrate Synthase. Curr.Top Cell. Regul. 33, 209-229.

Remington, S. Wiegand, G. and Huber, R. (1982). Crystallographic Refinement and Atomic Models of Two Different Forms of Citrate Synthase at 2.7 and 1.7 Å Resolution. J.Mol.Biol. 158,111-152.

Remington, S. Wiegand, G., Deisenhofer, J. and Huber, R. (1982). Crystal Structure Analysis and Molecular Model of a Complex of Citrate Synthase with Oxaloacetate and *S*-Acetonyl-coenzyme A. J.Mol.Biol. 174, 205-219.

Reysenbah, A.L. (2001). Bergey`s Manual of Systemic Bacteriology, 2nd Edn. New York, Springer, 1,335-340.

Rosenkrantz, M., Alam, T., Kim, K.-S., Clark, B.J., Srere, P.A., Guarente,L.P. (1986). Mitochondrial and nonmitochondrial citrate synthases in *Saccharomyces cerevisiae* are encoded by distinct homologous genes. Mol. Cell. Biol. 6:4509-4515.

Russell, R.J., Hough, D.W., Danson, M.J. and Taylor G.L.,(1994) The crystal structure of the citrate synthase form the thermophilic Archaeon, *Thermoplasma volcanium*. Structure 2(12), 1157-11

Russell R.J., Taylor G.L. (1995) Engineering thermostability: lessons from thermophilic proteins. Curr. Op. In Biotech. 6,370-374

Russell, R.J. Ferguson, J.M., Hough, D.W., Danson M.J. and Taylor, G.L. (1997). The Crystal Structure of Citrate Synthase form the Hyperthermophilic Archaeon *Pyrococcus furiosus* at 1.9 Å Resolution. Biochemistry 36, 9983-9994.

Sailágyi, A. and Závodszky, P. (2000). Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. Structure 8(5), 493-504.

Saitoh, Y., Ogawa, A., Hori, T., Kunita, R., Mizuno, S. (1993). Identification and localization of two genes on the chicken Z chromosome: implication of evolutionary conservation of the Z chromosome among avian species. Chromosome Res. 1:239-251

Sambrook, J. and Russel, D.W. (2001). Molecular Cloning. Cold Spring Harbor Laboratory Press. 1(6),15

Segerer,A., Langworthy, T.A. & Steter, K.O. (1988). *Thermoplasma acidophilum* and *Thermoplasma volcanium* sp. nov. from Solfatara Fields. Syst. Appl. Microbiol. 10, 161-171.

Srere, P.A., Brasil, H., and Gonen, L. (1963). The citrate condensing enzyme of pigeon breast muscle and moth flight muscle. Acta Chem. Scand. 17, S129-S143.

Srere, P.A. (1966). Citrate-condensing Enzyme-Oxalacetate Binary Complex. J. Biol. Chem. 241, 2157-2160.

Sutherland, K.J., Henneke.C.M., Towner, P., Hough, D., Danson, M.J. (1990). Citrate synthase from the thermophilic archaebacterium *Thermoplasma acidophilum*. Eur. J. Biochem. 194, 839,844.

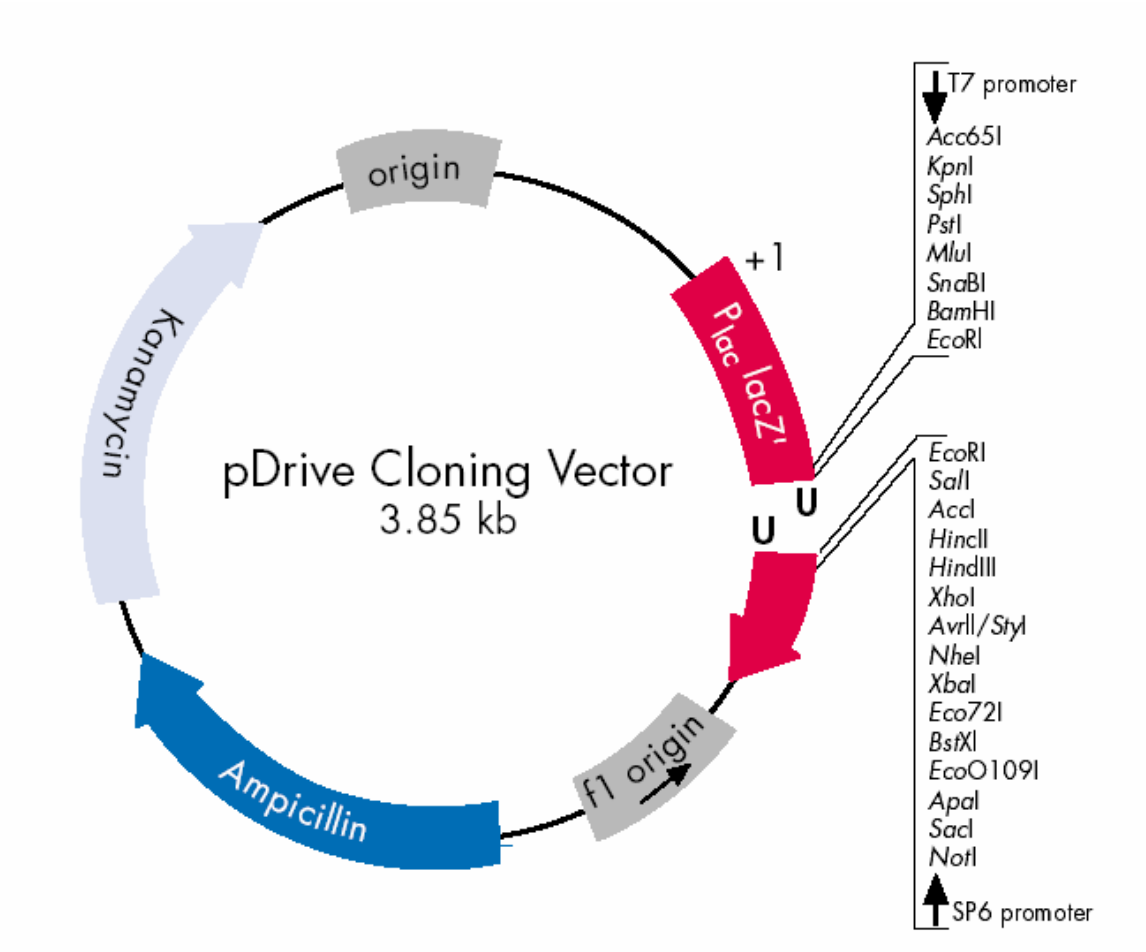
Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22:4673-4680.

Whitaker, J.R., Granum, P.E. (1980) An Absolute Method for Protein Determination Based on Difference in Absorbance at 235 and 280 nm. Analytical Biochemistry 109, 156-159.

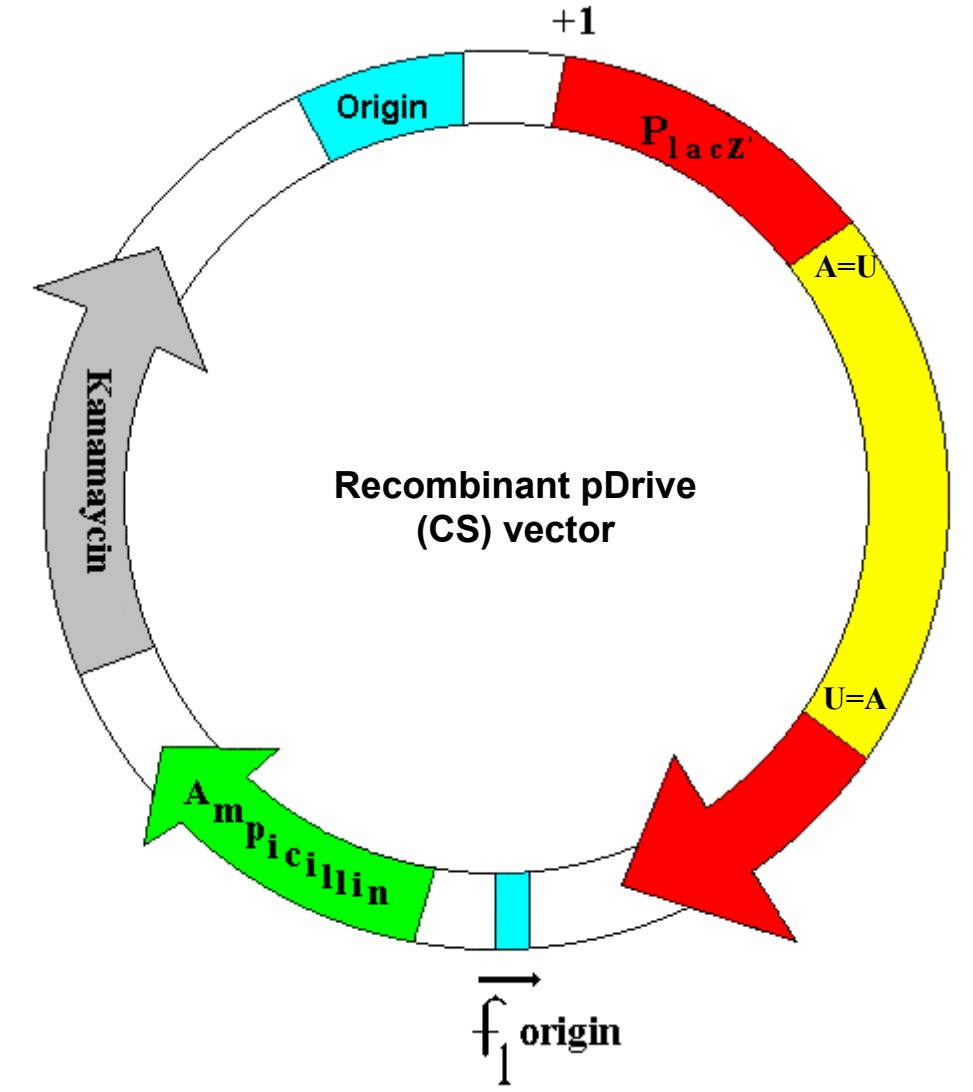
Woese, C.R. (1990). Towards a natural system of organisms: Proposal for the doains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci USA 87, 4576-4579.

APPENDIX A

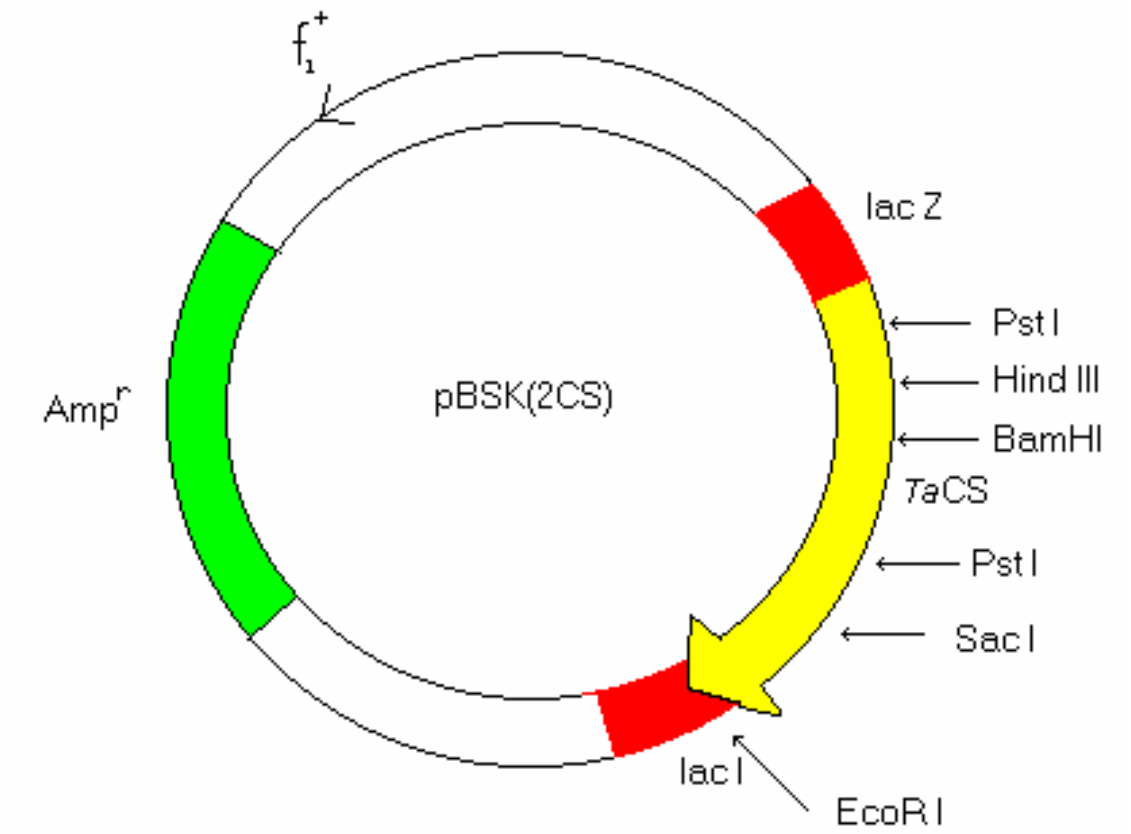
pDrive Cloning Vector



Recombinant pDrive-CS Vector
(Harboring *Tp.volcanium* citrate synthase gene)

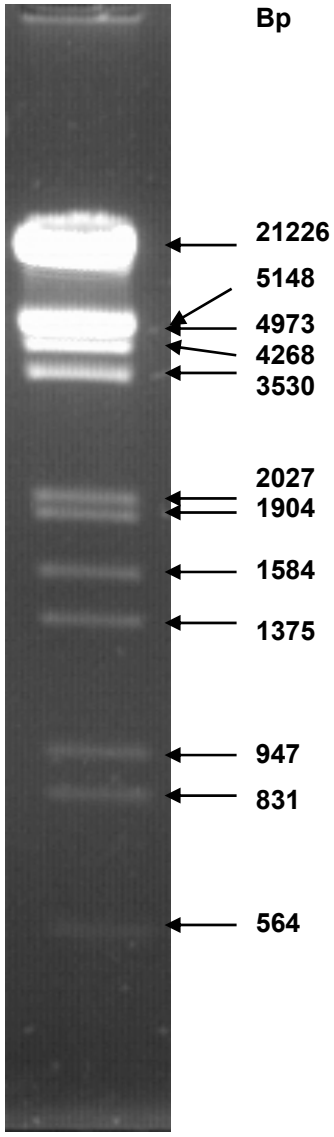


pBSK(2CS) Vector
(Harboring *Tp.acidophilum* citrate synthase gene)

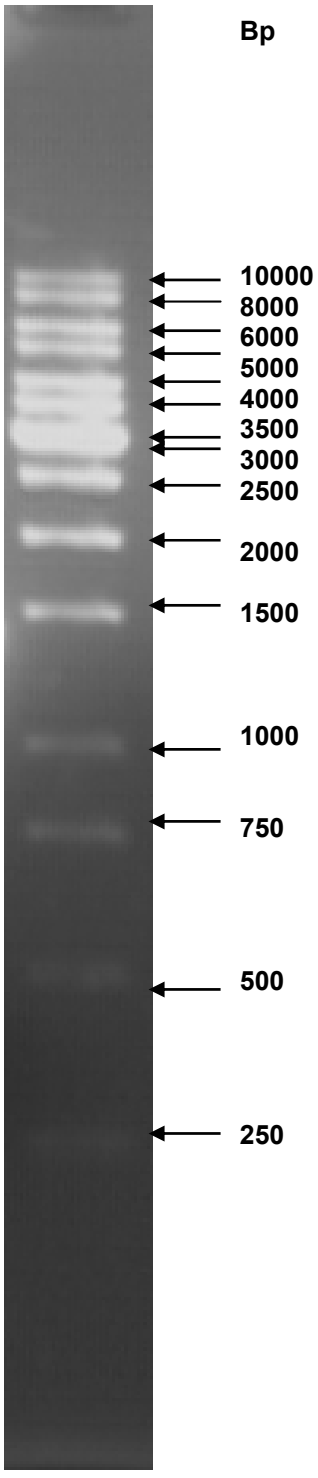


APPENDIX B

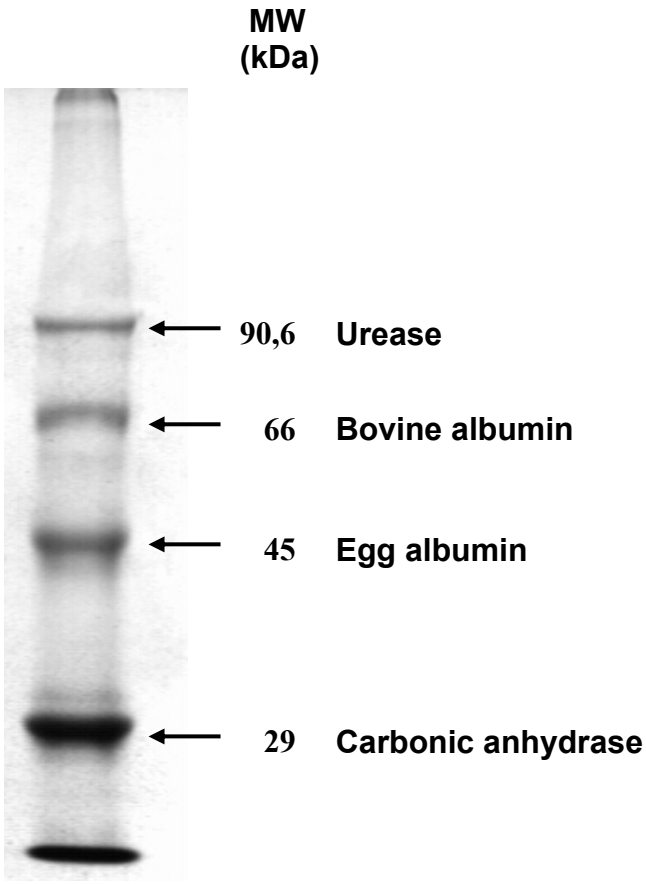
A) *Eco*RI and *Hin*dIII cut Lambda DNA (MBI Fermentas AB, Vilnius, Lithuania)



B) Gene Ruler, 1 kb DNA Ladder (MBI Fermentas AB, Vilnius, Lithuania)



C) Standard protein size markers (Sigma Chemical Co., St. Louis, Missouri, USA)



APPENDIX C

1. TE Buffer

10 mM Tris-HCl pH 7.4
1 mM EDTA pH 8.0

Autoclave the solution at 121^{o}C for 20 min. And store at 4^{o}C.

2. 20X SSC

NaCl	175.3 g/L
Na-Citrate	88.2 g/L

Adjust pH to 7.5 by 10 N NaOH and complete to 1 L by distilled water. Store at 4^{o}C.

3. TSS solution

0.5 M $MgCl_2$ (or $MgSO_4$)	10 ml
50% PEG-800	20.0 ml
DMSO	5.0 ml
LB-Broth	65 ml

Store the solution inside of a dark bottle, at 4^{o}C.

4. Reaction and Extraction Buffer

20 mM Tris-HCl, 1 mM EDTA, pH 8.0

Adjust pH to 8.0 with 10 M HCl, Autoclave the solution at 121^{o}C for 20 min. And store at 4^{o}C.

5. 50X TAE Buffer

Tris-HCl	242 g.
Glacial acetic-acid	57.10 ml
EDTA 0.5 M(pH:8.4)	100.00 ml
pH=8.4	

Complete the solution to 1L with distilled water, autoclave the solution at 121^{o}C for 20 min. And store at 4^{o}C.

6. Hybridization buffer for Southern Blot and Hybridization (for 100ml)

20X SSC	20 ml
Blocking reagent powder	5 g
10% (w/v) SDS	0.2 ml
N-lauroylsarcosine	100 mg
sdH_2O	50 ml

Add formamide before use. Keep the solution at -20^{o}C. Do not sterilize.

Formamide 100% (deionized)	50 ml

7. Maleic acid Buffer for Southern Blot and Hybridizaiton

Maleic acid	0.1 M
NaCl	0.15 M

Adjust pH to 7.5 by concentrated NaOH. Sterilize at 121^{o}C by autoclaving for 20 min. Store at 4^{o}C.

8. Washing Buffer for Southern Blot and Hybridization (for 100ml)

Maleic acid buffer	250 ml
Tween 20	750 µl

Do not autoclave. Store at 4^{o}C.

9. 10X Blocking Solution for Southern Blot and Hybridization (for 100ml)

Maleic acid buffer	100 ml
Blocking reagent powder (provided in detction kit)	10 g

Autoclave at 121 ^{o}C for 20 min.store at 4^{o}C.

10. Running Buffer for SDS-Polyacrylamide Gel Electrophoresis

Tris-HCl	3.02 g/L
Glycine	14.40 g/L
SDS	1.00 g/L

Complete the volume to 1 L by distilled water. Autoclave at 121^{o}C for 20min. And store at 4^{o}C.

11. 2X Sample Buffer in SDS-Polyacrylamide Gel Electrophoresis

Tris-HCl	1.51 g
Glycerol	20.00 ml

Dissolve in 35 ml distilled water and adjust pH to 6.75 with concentrated HCl; then add,

SDS	4.00 g
B-mercaptoethanol	10.00 ml
Bromophenol blue	2.00 mg

Dilute with distilled water to a final volume of 100 ml, and store the solution at -20^{o}C.

12. Staining Soltion

Coomassie Gel Stain

Coomassie Blue R-250	1.0 g/L
Methanol	450.0 ml/L
Glacial acetic acid	100.0 ml/L
dH_2O	450 ml/L

13. Destaining Solution

Coomassie Gel Destain

Methanol	100 ml
Glacial aceteic acid	100 ml
dH_2O	800 ml