Investigation of the structural basis of protein hyperthermostability in citrate synthase for the Archaea

Arnott, Michael A.

Award date:
1999

Awarding institution:
University of Bath

INVESTIGATION OF THE STRUCTURAL BASIS OF PROTEIN HYPERTHERMOSTABILITY IN CITRATE SYNTHASE FROM THE ARCHAEA

Submitted by Michael A Arnott
for the degree of PhD of the University of Bath
1999

M. A. Arnott.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

DEDICATED

TO

MY PARENTS, NIGHAT & ALISON

ABSTRACT

The ability of enzymes to withstand extreme temperatures and remain active is of great commercial interest. In seeking to understand the structural basis of enzyme thermostability many studies have compared and contrasted enzymes from mesophilic and thermophilic hosts. Russell et al. (1997) have made predictions of structural features responsible for the enhanced thermostability of citrate synthase (CS) from the archaeon *Pyrococcus furiosus* (*Pf*), optimum growth temperature 100°C, based on structural comparisons of citrate synthases from pig, *Thermoplasma acidophilum* (*Tp*), optimum growth temperature 55°C, and *P.furiosus.* . Some of the key findings from this comparison are that the structure of *Pf*CS relative to pig and *Tp*CS is more compact, has better complementarity between the subunits of the dimer, has fewer thermolabile residues, a greater number of intersubunit ionic interactions and an additional subunit interface interaction at its C-terminii.

Site-directed mutants have been made of *Pf*CS to test the importance of the additional C-terminal subunit interaction, and of the 5 membered subunit interface ionic network, to thermostability. Denaturation studies of these mutants have confirmed that these structural features do play a role in the thermostability of *Pf*CS.

The crystal structures determined for *Tp*CS and *Pf*CS show that each monomer of the enzyme consists of a large and a small domain. Chimaeric citrate synthases have been constructed, where the small domains have been swapped between *Pf*CS and *Tp*CS, to determine where the principal determinants of thermostability and catalytic activity lie. This study has shown that the key determinants of thermostability are located in the large domain of *Pf*CS, while the key determinants of catalytic activity are located within the small domain.

CONTENTS

CHAPTER 2

Materials and Methods

CHAPTER 3

Creation of citrate synthase domain swap mutants

CHAPTER 4

Production of *Pf*CS C-terminal mutants

CHAPTER 5

Production of *Pf*CS subunit interface ionic network mutants

FIGURES

TABLES

ACKNOWLEDGEMENTS

I would like to thank my supervisors Professor Mike Danson and Dr David Hough for allowing me the opportunity to do this PhD, and for the time that they have spent with me in a range of useful discussions. Thanks are also due for their guidance throughout the project, and help in the production of this manuscript.

I would also like to thank all of my friends at the Centre For Extremophile research, both past and present, for being such good fun, making it a pleasure to come to work. Special thanks to Dr Carl Thompson, Dr Rupert Russell, and Dr Cameron Dunn for their help with my many computing 'problems'.

I gratefully acknowledge the BBSRC who provided the sponsorship for this project.

Finally a big thank you to my wife Nighat for her patience, encouragement and understanding throughout this project. I would also like to thank my family for their support.

Abbreviations

Å	Angström
Ac-CoA	Acetyl coenzyme A
Asn	Asparagine
Asp	Aspartate
ATP	Adenosine triphosphate
bp	Base pairs
BSA	Bovine serum albumin
CoA	Coenzyme A
CS	Citrate synthase
Cys	Cysteine
ΔG_{stab}	Conformational stability
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DTNB	5, 5'-di-thiobis-2-nitrobenzoic acid
EDTA	Ethylenediamine tetraacetate
FDH	Formate dehydrogenase
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
Gln	Glutamine
GluDH	Glutamate dehydrogenase
His	Histidine
Ile	Isoleucine
IPMDH	3-isopropylmalate dehydrogenase
IPTG	Isopropyl-β-D-thiogalactoside
kb	Kilobase pairs
kcal	Kilocalorie

kDa	Kilodalton
LB	Luria Bertani medium
Leu	Leucine
Lys	Lysine
Met	Methionine
M_r	Relative molecular mass
NADH	Nicotinamide
dNTP	Deoxynucleoside triphosphate
OAA	Oxaloacetate
PCR	Polymerase chain reaction
Peg	Polyethylene glycol
Pf	*Pyrococcus furiosus*
Phe	Phenylalanine
PMSF	Phenylmethysulphonylfluoride
Pro	Proline
rRNA	Ribosomal RNA
RNAse	Ribonuclease
SDM	Site directed mutagenesis
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
TAE	40mMTris-acetate, 1mM EDTA
TEMED	Tetramethylethylenediamine
TLN	Thermolysin
TLP	Thermolysin like protein
Tp	*Thermoplasma acidophilum*
Val	Valine
Xgal	5-Bromo-4-chloro-3-indoyl-β-D-galactoside

CHAPTER 1

INTRODUCTION

1.1 General

The work presented in this thesis aims to contribute to our understanding of the intrinsic factors that contribute to the structural basis of protein thermostability. Citrate synthase from the Archaea has been chosen as the model protein in this study. This introduction will therefore give a description of citrate synthase, the Archaea, and intrinsic features that are thought to be important in protein stability.

1.2 Citrate Synthase

This section gives a general background on CS, explaining why we have chosen it as our model enzyme, describing its mechanism of action, and the relationship between its structure and the taxonomic status of the organism from which it was isolated.

1.2.1 Citrate synthase as the model protein

CS was chosen by our group to investigate the structural determinants of thermostability for the following reasons:

1. Citrate synthase plays an important role in cell metabolism, in that it is the first enzyme of the citric acid cycle (Krebs & Lowenstein, 1960). The citric acid cycle is shown in Fig 1.1. The importance of the role of CS in cellular

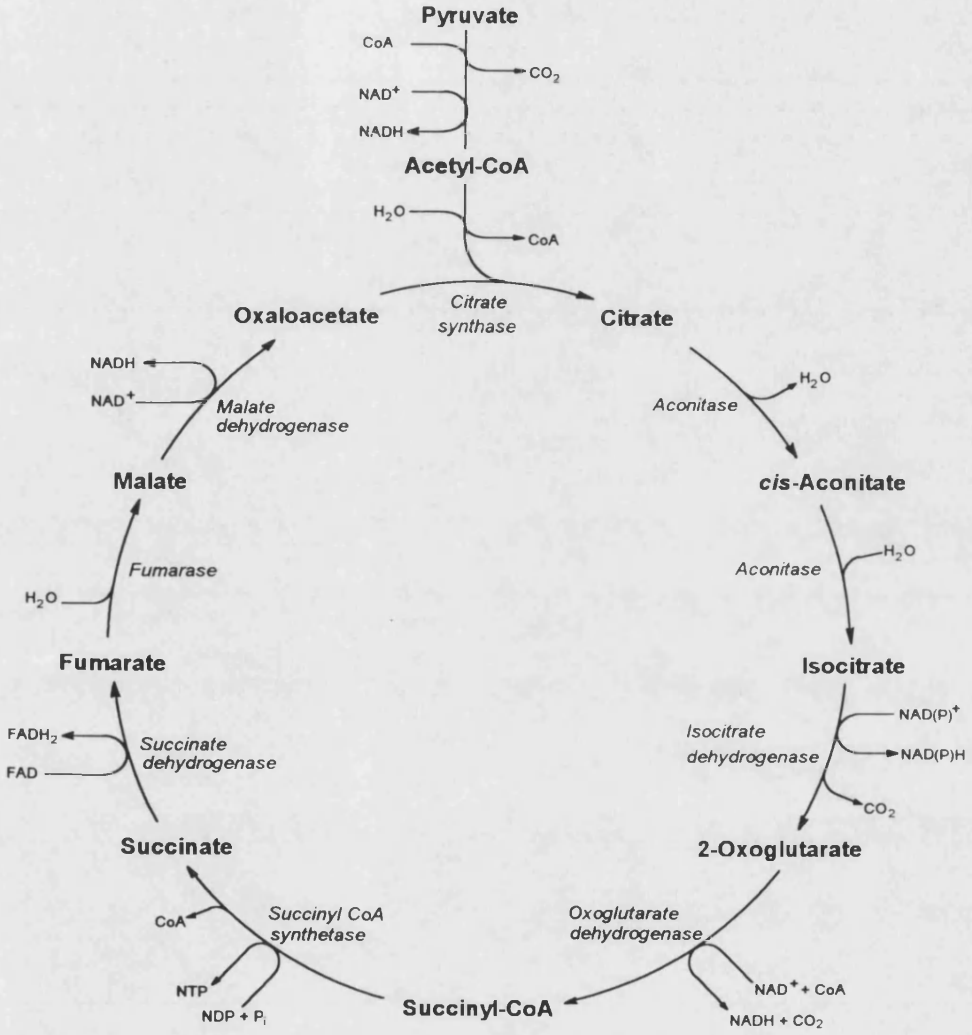


Figure 1.1 Position of citrate synthase within the citric acid cycle.

metabolism is reflected in the fact that the enzyme is conserved throughout the three domains of life, and has maintained the same function throughout evolution (Danson, 1988; Muir et al., 1994).

2) There is a wide range of sequence information on CS, with around 50 sequences listed in the database. Crystal structures have been determined for the open (ligand free) and closed (substrate bound) forms of the enzyme from pig and chicken heart (Remington, 1992; Remington et al., 1982; Weigand et al., 1984), the open form from the thermophilic archaeon *Thermoplasma acidophilum* (Russell et al., 1994), the closed form of the hyperthermophilic archaeon *Pyrococcus furiosus* (Russell et al., 1997), and the psychrophilic bacterium DS23r (Russell et al., 1998). Our group has studied CS for a number of years, cloning, sequencing, expressing, and crystallising CS's from organisms that inhabit environments which range from the Antarctic (DS23r) to geothermally heated marine sediments (*P.furiosus*).

3) The enzyme can be easily purified (James et al., 1994), and there is a good assay of its activity (Srere et al., 1963).

1.2.2 CS Mechanism

Citrate synthase catalyses the condensation of acetyl-CoA and oxaloacetate to form citrate and free CoA, and thereby effects the entry of carbon into the citric acid cycle.

The catalytic mechanism has been studied extensively in pig heart CS. The current working hypothesis for the mechanism of the CS catalysed reaction in pig, is that oxaloacetate binds first to the enzyme's active site, causing a large conformational change from the enzyme's open form to a closed form. His320

causes polarisation of the oxaloacetate carbonyl bond thereby increasing the positive charge on the 2-carbon of oxaloacetate, making it more susceptible to condensation with Ac-CoA. Binding of Ac-CoA to the CS-OAA complex is followed by a concerted acid -base reaction, where Asp375 is the general base and His 274 is the general acid. This causes an enolization of Ac-CoA (the methyl group being deprotonated) to yield an enolized intermediate, the proposed nucleophile that may be neutral or partially negatively charged. Additional conformational changes are believed to accompany this step. The nucleophile then condenses with the OAA to give citryl CoA. Citryl CoA is hydrolysed, possibly utilising Asp375, to yield citrate and free CoA. Finally a conformational change occurs to an open form to release the products. The proposed mechanism for CS is shown in Fig 1.2 (reproduced from M^{c}Cormack, 1995). This scheme has been derived primarily from detailed kinetic analysis of wiidtype, chemically modified, and site directed mutants of CS's (Bayer et al., 1981, Kurz et al., 1992, Evans et al., 1988, 1989, 1996; Alter et al., 1990, and references therein) and from analysis of CS crystal structures (Remington et al., 1982; Karpusas et al., 1990; Liao et al., 1991; Remington, 1992; Usher et al., 1994),

1.2.3 Relationship between enzyme structure and taxonomic status of organism

Studies of CS's from the eukaryal and bacterial domains have shown a correlation between the taxonomic status of the organism and the subunit composition, catalytic activity, and regulatory properties of the enzyme (Weitzman and Danson, 1976; Weitzman, 1981; Danson, 1988). The enzyme from Eukarya, Archaea, and Gram-positive bacteria is an isosterically regulated homodimer (inhibited by ATP) whereas the majority of Gram-negative bacteria possess an allosterically regulated hexamer (inhibited by NADH and showing no response to ATP). The exceptions to this rule are the dimeric CS's from the

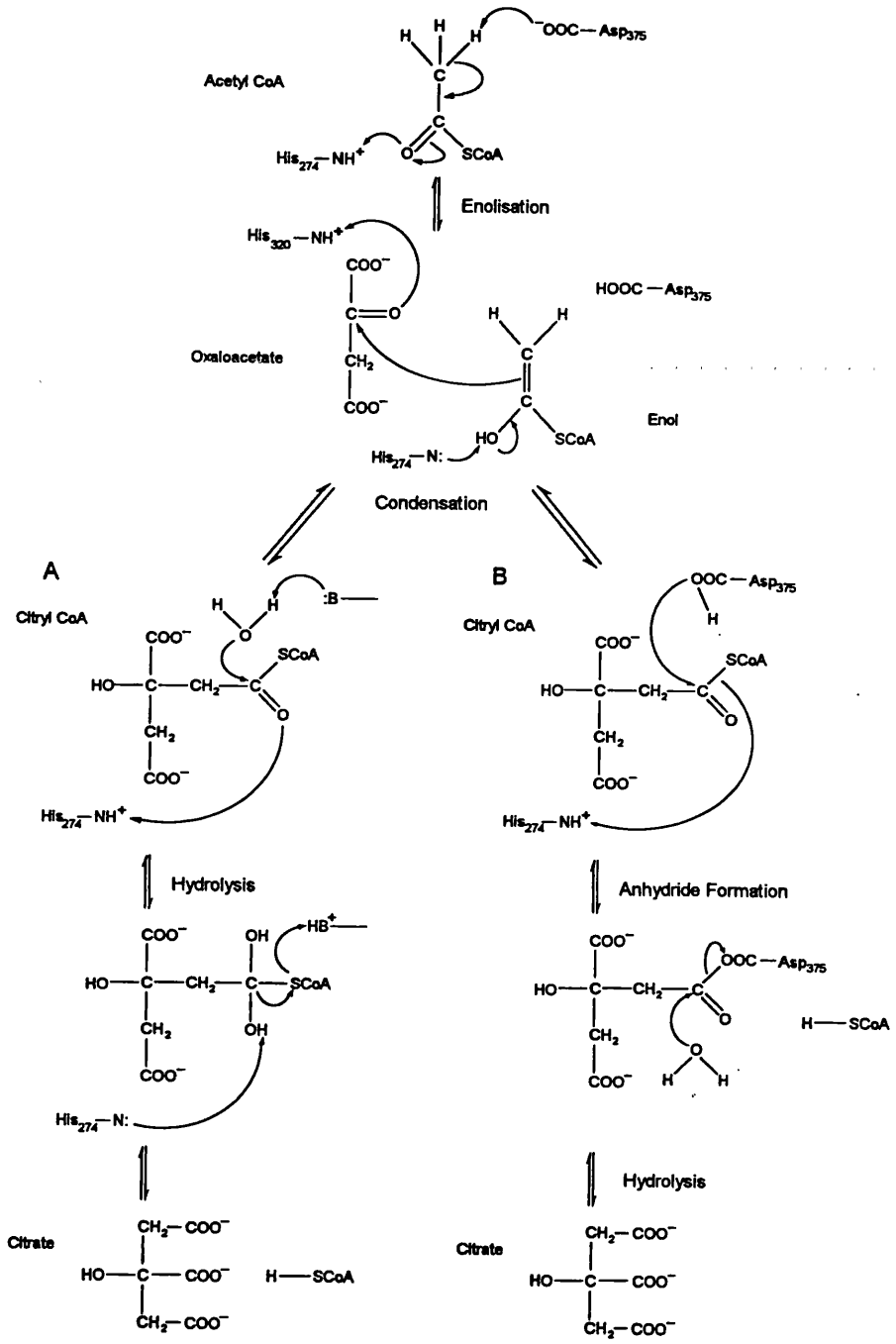


Figure 1.2 The Remington model for the catalytic mechanism of pig CS (Remington, 1992). There are three steps, enolization of Ac-CoA, condensation with oxaloacetate, hydrolysis to release citrate and coenzyme A. Two mechanisms have been proposed for the hydrolysis step, step A (Remington, 1992), and step B (Alter et al., 1990).

Gram-negative organisms *Thermus aquaticus, Rickettsia prowazekii*, and *Coxiella burnetii,* which are also isosterically regulated by ATP. Despite this oligomeric variation, the subunit M_r is between 42,000 and 50,000 in all cases.

Each monomer of CS consists of two domains, a large domain and a small domain. The large domain contains residues that make up the subunit interface, while the small domain is mainly concerned with the catalytic activity of the enzyme. In pig CS the small domain is known to rotate relative to the large domain upon binding of substrates (Remington et al., 1982). In CS's from pig, *Tp* and *Pf*, the average main chain β-factor for the small domain is higher than that of the large domain (Russell et al., 1997) indicating an increased flexibility of the small domain with respect to the large domain. In *Pf*CS each monomer is composed of 16 α-helices. Fig 1.3 shows the crystal structure of *Pf*CS, showing the dimeric nature of the enzyme, and bound citrate and CoA. Fig 1.4 shows the large and small domains, and the subunit interface of *Pf*CS. The main monomer-monomer association for all the dimeric CS's is an 8 α-helical sandwich, comprising 4 anti-parallel pairs of helices (F, G, L, and M). The enzyme is only active as a dimer as residues from each subunit contribute to each active site.

1.3 The Archaea.

On the basis of a comparison of 16SrRNA sequences, Woese et al (1990) proposed that life could be divided into three primary phylogenetic domains, the Archaea, the Eukarya, and the Bacteria. The Archaea are a phylogenetically distinct group of microrganisms that from comparison of the 16SrRNA genes, have been subdivided into two further groups, the *Crenarchaeota* and *Euryarchaeota.* The kingdom *Crenarchaeota* generally consists of hyperthermophiles or thermoacidophiles (some genera are *Pyrodictium* and

Figure 1.3 Crystal structure of *Pf*CS, one monomer is shown in red, the other in yellow. Bound citrate and CoA are shown in their binding sites in grey.

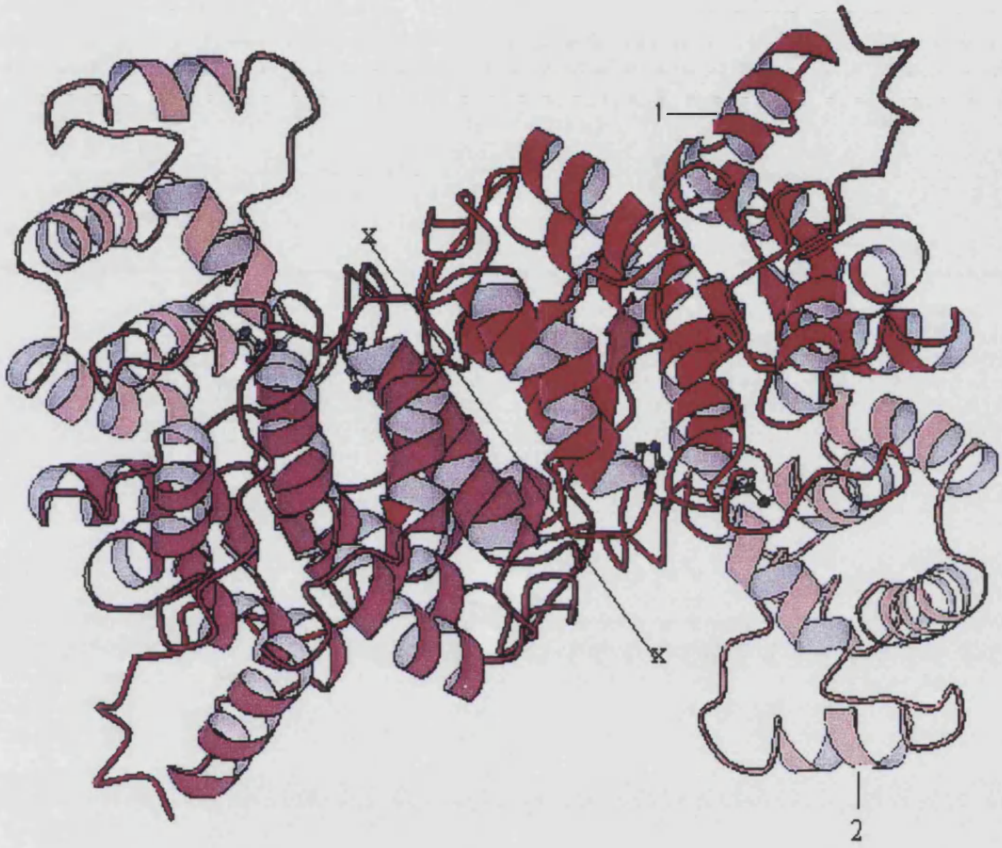


Figure 1.4 Structure of *Pf*CS viewed down the 2-fold axis of the dimer. Key. 1) Large domain coloured in red and mauve. 2) Small domains are coloured pink. X---X) Subunit interface.

Thermofilum), while the kingdom Euryarchaeota spans a broader ecological range and includes hyperthermophiles (eg *Pyrococcus* and *Thermoplasma*), methanogens (eg *Methanosarcina*), and halophiles (eg Haloferax and Halobacterium) (Brown & Doolittle, 1997).

The Archaea can also be grouped phenotypically according to the extreme environmental niche in which they live. The extremophilic phenotypes of the Archaea include thermophiles, hyperthermophiles, psychrophiles, halophiles, methanogens, acidophiles, alkaliphiles and barophiles, (Danson & Hough, 1998). However PCR amplification and analysis of DNA from soil samples has shown that the Archaea inhabit a much wider range of habitats than previously assumed (Bintrim et al., 1997; Service, 1997); this, coupled with the fact that there are many examples of bacterial extremophiles, means that the definition of an Archaeon should be a phylogenetic one, as opposed to a phenotypic one. However, in this introduction, the Archaea for convenience will be discussed in terms of their phenotypic groups and this will be limited to a brief description of the thermophiles.

The extremophilic nature of the Archaea has generated considerable interest in their physiological adaptations to such extreme environments. Biotechnology companies are aiming to exploit the enhanced stability properties of enzymes from these organisms (Cowan, 1996; Danson & Hough, 1998; Huber & Stetter, 1998; Johnson, 1998). The reason we have used enzymes isolated from the Archaea to study thermostability is that these proteins must contain adaptations/modifications, compared with enzymes isolated from mesophilic organisms, to withstand such extreme conditions.

1.3.1 Thermophiles

The archael thermophiles are a diverse group of organisms that are characterised by their high growth temperatures of between 50-110°C (Woese et al., 1990). This group can be subdivided into those that grow optimally at 55°C-80°C (thermophiles), and organisms that grow optimally at 80°C-110°C (hyperthermophiles). The most common thermophilic habitats are between 50-70°C, examples of these include solar heated soil, decaying vegetation, and self-heating coal tips. Habitats where the temperatures are greater than 70°C are relatively rare, and tend to be geothermally active regions.

Details of the Archaea from which the CS's used in this project were originally isolated are given below.

Pyrococcus furiosus, a hyperthermophile whose name translates as 'furious fireball' (furious referring to the rapid doubling time and strong motility of the organism), was originally isolated from geothermally-heated marine sediments in Italy, and grows between 70-103°C with an optimum growth temperature of 100°C (Stetter & Fiala, G, 1986). The organisms are spherically shaped, 0.8-2.5μm in width, and are anaerobic heterotrophs, growing on starch, maltose and peptone substrates.

Thermoplasma acidophilum, a thermophile whose name translates as heat and acid loving, was first isolated from burning coal refuse piles (Darland et al., 1970) but has since been isolated from naturally occurring hot springs and solfataric fields (Segerer et al., 1988). It is an aerobe, whose cells are motile and highly irregular in shape due to the lack of a cell wall. It grows at temperatures of between 45-63°C, with an optimum temperature for growth of 59°C, and an optimum pH of 1-2.

1.4 Protein Thermostability

1.4.1 What is thermostability ?

When exposed to sufficiently elevated temperatures all enzymes will eventually lose catalytic activity. Thermally induced enzyme inactivation is classified as reversible or irreversible depending on whether activity is restored following a return to ambient conditions (Lumry & Eyring, 1954):

$$N \overset{K}{\rightleftharpoons} U \xrightarrow{k} I$$

where *N* is a native catalytically active enzyme, *U* is a reversibly thermally unfolded enzyme, and I is an irreversibly thermoinactivated enzyme. The first step in this scheme is the reversible denaturation of an enzyme's native conformation, and K defines the equilibrium constant between the *N* and *U* forms of the enzyme. The subsequent irreversible thermoinactivation processes, described by the rate constant k, can involve both conformational and covalent changes that are specific for individual enzymes. These changes are frequently accompanied by aggregation and precipitation.

When an enzyme is described as being thermostable what is implicit in this statement is that it is able to withstand thermally induced enzyme inactivation to a degree. There is a considerable interest in the molecular mechanisms that contribute to stability, because by understanding these, biologists hope to be able to engineer stability into enzymes, so that they can be used in industrial and medical applications that employ conditions quite different to those in which the enzyme has evolved to function in nature. In some cases, lessons from nature have been applied in the engineering of more stable variants of existing industrial enzymes (Rojkova et al.,1999; Nosoh & Sekiguchi, 1990). Also an understanding

of the structural basis of enzyme stability may explain how some organisms can tolerate extreme conditions of temperature and pH.

The conformational stability of a protein can be defined as the difference in the free energy between the folded and unfolded conformations under physiological conditions. For most naturally occurring globular proteins this is remarkably low, being estimated at between 5-15 kcal/mol (Pace, 1990). It may be advantageous for organisms to have proteins for which the folded active conformation is only marginally more stable than the unfolded inactive conformations. The reasons for this are not clear but may be related to the protein's turnover or function (Fersht & Serrano, 1993; Arfin & Bradshaw, 1988). Studies comparing the conformational stability (ΔG_{stab}) values of homologous enzymes from mesophilic and thermophilic hosts show that the ΔG_{stab} values of enzymes from thermophilic hosts are higher than those of enzymes from mesophilic hosts by between 3 and 15 kcal/mol (Nojima et al., 1977, 1978). These ΔG_{stab} values were determined by reversible thermal unfolding experiments. Stability studies on mutant enzymes (Shih &Kirsch, 1995; Kawamura et al., 1996) have shown that changes in ΔG_{stab} values as small as 3-6.5 Kcal/mol can account for thermostability increases of up to 12°C. The ΔG_{stab} values have not been calculated for any enzymes from hyperthermophiles but predictions of a ΔG_{stab} value for α-glucosidase from *P.furiosus* (Dale & McBennett, 1992) suggest that this may be significantly higher than the ΔG_{stab} values for α-glucosidase from a mesophilic host. The difference in ΔG_{stab} values for homologous enzymes from mesophilic, thermophilic and hyperthermophilic hosts is a result of differences in specific amino acid sequences, and the structures to which these give rise.

1.4.2 Strategies for investigating thermostability

Two main strategies have been used to investigate enzyme thermostability. One is a comparative approach whereby either the amino acid sequences or crystal structures of homologous enzymes from mesophilic, thermophilic, or hyperthermophilic organisms are compared. Their differing thermostabilities can then be attributed these differences, although phylogenetic differences can confuse the picture. Several studies have sought to compare the sequences of mesophilic and thermophilic enzymes belonging to protein families for which a structure was available (Argos et al., 1979; Menendez-Arias & Argos, 1989; Britton et al., 1995). Suggested preferences in the selection of amino acids for thermophilic enzymes included exchanges such as lysine to arginine, glycine / serine to alanine, and serine to threonine. However sequence comparisons of mesophilic, and thermophilic enzymes are not always consistent with these suggestions (Pauptit et al., 1988). Also with sequence comparisons between mesophilic and thermophilic enzymes within homologous families, the number of sequence changes are generally so numerous that it is difficult to locate those that contribute to increased thermostability. Thus analysis of amino acid sequence alone is insufficient to allow prediction of features that give rise to enhanced thermostability.

The second approach is the use of site directed mutagenesis to create mutants to determine their effects on the protein's thermostability. This second approach is often combined with the first to test predictions made by structural / amino acid comparisons. One way in which site directed mutagenesis is used to investigate thermostability is in the mutation of a mesophilic enzyme to incorporate some of the structural features from the thermophilic homologue that have been implicated in its thermostability. A number of groups have used this approach successfully. Vetriani et al (1998) carried out a structural comparison

between hexameric glutamate dehydrogenases from the hyperthermophiles *P.furiosus*, and *Thermococcus litoralis*, which grow optimally at 100°C and 88°C respectively, and found that an intersubunit ion pair network was substantially reduced in the less stable enzyme from *T.litoralis.* Two residues were altered by site directed mutagenesis to restore these interactions. The single mutations both had adverse effects on the thermostability of the enzyme. However with both mutations in place to create an ionic interaction, a fourfold improvement in stability at 104°C was observed over the wildtype enzyme. Following comparison of the amino acid sequences from a thermolysin like protein (TLP) from *Bacillus stearothermophilus* (TLPst) with its more stable counterpart thermolysin (TLN), Van den Burg et al (1998) showed that the TLPst, could be made more stable than TLN by mutating TLPst to incorporate 5 amino acids that are found in TLN, and 3 rationally designed mutations. The advantage of the mutagenesis approach is that any amino acid change can be directly attributed to a change in stability.

1.4.3 Determinants of thermostability

Jaenicke & Bohm (1998) have reviewed the strategies of thermal stabilisation of 8 thermophilic enzymes and reported that in going from lower to higher growth temperatures numerous differences are frequently reported: the clustering of intrasubunit and/or intersubunit ion pairs; improved packing of the hydrophobic core; additional networks of hydrogen bonds, increased helix-dipole stabilisation; an increased polar surface area; a decreased number and total volume of cavities, and burying the hydrophobic surface area. Querol et al. (1996) have also listed a number of different physical and chemical features that researchers have reported to explain enhanced thermostabilization in proteins, and these include improved hydrogen bonding, better hydrophobic internal packing, enhanced secondary structure propensity, helix dipole stabilization,

removal of residues sensitive to oxidation or deamination, burying of hydrophobic accessible area, improved electrostatic interactions, strengthening of the inter-subunit association, decreased backbone strain, salt bridge optimisation, enhanced stability of secondary structures, decreased glycines and increased prolines in loops, shorter loops, additional disulphide bridges, decrease in number and volume of internal cavities, and an increase in aromatic interactions.

Proteins may be engineered or evolve in vivo to achieve greater stability using one or more of these strategies, although no single or preferred method has yet been found. A brief description of some of the intrinsic features determined to play a role in protein stability is now given.

1.4.3.1 Intrinsic determinants of thermostability

1.4.3.1 a Ionic interactions

The importance of ionic interactions in thermostability has been demonstrated by increasing the number of ionic interactions within a protein by site directed mutagenesis (Vetriani et al., 1998), see section 1.4.2, and by the disruption of ionic networks leading to a decrease in thermostability (Pappenberger et al., 1997). Structural comparisons have also highlighted the importance of ionic interactions in thermostability (Wallon et al., 1997; Korndorfer et al., 1995; Yip et al., 1995, 1998; Russell et al., 1997).

However a number of studies have questioned the importance of ion pair networks to thermostability. For example, the replacement by mutagenesis of three charged residues that form part of a small ion pair network in the Arc repressor by three hydrophobic residues resulted in an increase in stability (Waldburger et al., 1995). Furthermore, the comparative structural analysis of the neutral protease from *Bacillus cereus* and the thermostable enzyme

thermolysin failed to reveal a dramatic increase in the number of ion pairs, suggesting that this type of interaction is of little importance in generating stability in this enzyme (Pauptit et al., 1988).

1.4.3.1 b Hydrophobic effect

Hydrophobic interactions are often thought to provide the energy required for proteins to fold in aqueous solutions. From a study of 72 side-chain aliphatic mutations in four enzymes, Pace (1992) calculated that each buried $-CH_2$ group contributes an average 1.3 kcal mol^{-1} to the conformational stability of the enzyme. Hydrophobic interactions have been predicted to play a role in protein stability from a comparison of the amino acid sequences (Haney et al., 1999) and crystal structures of thermophilic enzymes with their mesophilic counterparts (Davies et al., 1993; Koralev et al., 1995). The role of hydrophobic interactions in stability has also been shown by work on 3-isopropylmalate dehydrogenase; when two residues at the dimer interface of the *E.coli* enzyme (Glu246 and Met249) were replaced by the more hydrophobic, homologous residues from *T.thermophilus* IPMDH (Leu246 and Val249) the resulting mutant was stabilised against dissociation and denaturation by urea (Imada et al., 1991; Kirino et al., 1994). Furthermore Rojkova et al. (1999) in work on formate dehydrogenase (FDH) from a *Pseudomonas* sp, showed that the stability of the enzyme could be increased 1.5 fold by substituting α-helical residues with more hydrophobic residues.

Closely packed aromatic ring-ring interactions in proteins represent a type of hydrophobic interaction, and their stabilising potential is approximately 1kcal/mol (Burley & Petsko, 1985). A large number of aromatic interactions is thought to be one of the stabilizing factors of thermitase from *Thermoactinomyces vulgaris* (Teplyakov et al., 1990).

1.4.3.1c Hydrogen bonds

In a hydrogen bond, a hydrogen atom is shared by two other electronegative atoms. The atom to which the hydrogen is more tightly linked is called the hydrogen donor whereas the other atom is the hydrogen acceptor. The donor atom in a hydrogen bond is oxygen or a nitrogen that has a covalently attached hydrogen atom. The acceptor atom is either oxygen or nitrogen. The contribution of hydrogen bonds to a protein's native state has been investigated by disruption of the bonds by mutating one or both of the residues involved in a pair. Shirley et al. (1992) carried out thermal and unfolding studies on 12 mutants of ribonuclease T1 (with disrupted hydrogen bonds), and found that the hydrogen bonds made a significant contribution to the protein's conformational stability. An average decrease in stability of 1.3 kcal mol^{-1} was found per hydrogen bond destroyed. Vogt et al. (1997) noticed in their structural comparison study of proteins from 16 families that in 80% of the proteins there was an average increase of 11.7 hydrogen bonds per chain per 10°C rise in stability.

1.4.3.1d Disulphide bonds

A disulphide bond is formed when the sulphydryl groups of two cysteine residues are oxidised to produce a cystine residue. Disulphide bonds might be expected to stabilise the conformation of a folded protein by forming a physical constraint to unfolding. However Wetzel (1987) suggests that disulphide bonds actually stabilise by decreasing the degrees of freedom of the unfolded state, thus disulphide crosslinks decrease the amount of entropy that can be gained in the unfolding reaction. Pace et al. (1988) suggested that a disulphide bridge can contribute up to 4 kcal/mol to stability. Attempts at stabilising proteins by introducing disulphide bridges have met with mixed success. In many cases introduction of a disulphide bridge has resulted in destabilisation of the protein

(Villafranca et al.,1988; Wetzel et al., 1988). In other cases introduced disulphide bridges have been shown to be stabilising. Mansfeld et al. (1997) showed the importance of disulphide bridges in thermostability by engineering one into a thermolysin-like protease from *B.stearothermophilus*. The thermostability of the mutant was increased significantly (the half life of the double mutant enzyme being increased more than 120 fold at 92°C compared with the wildtype enzyme).

1.4.3.1e Helix Stabilising factors

The α helix is one of two prominent types of secondary structure found in globular proteins, and for this reason many studies have concentrated on factors that affect helix stability. Facchiano et al. (1998) compared the Xray structures of 13 thermophilic proteins with their mesophilic counterparts (using a computational method to estimate stabilizing contributions associated with helix formation) to determine the difference in the stability of their helices. They found that overall the helices of thermophilic proteins were more stable than their mesophilic homologues in 69% of cases.

Munoz and Serrano (1995) have grouped α-helix stabilising factors into sequence motifs (the N-capping box, the hydrophobic staple, the Schellman motif, and the alpha L motif, single residue effects (N-capping residue, β-branched effects, charge-dipole interaction) and interactions between specific pairs of side chains (Lys-Glu, Phe-Cys/Met, hydrophobic effect). The modification of these stabilising factors in helices can be used to improve stability (Nicholson et al., 1988; Zhukovsky et al., 1994).

Facchiano et al. (1998) however have disputed the importance of some α-helical stabilising factors in their study, finding little evidence for the importance of the N-capping box, the Schellman, alphaL and hydrophobic staple motifs, and

concluded that these factors were not related to the stability of thermophilic proteins, although they did not exclude the fact that they could play a stabilising role in specific cases.

1.4.3.1f Reduction in the entropy of unfolding

Mathews et al. (1987) proposed that proteins could be stabilised by decreasing the entropy of their unfolded state. In the unfolded state Gly is the residue with the highest conformational entropy and in the folding process much more energy is required to restrict the conformation of Gly than any other residue. Pro residues, with their pyrrolidine ring, can only form a few configurations and they restrict the configurations allowed for the preceding residue. Thus by substituting any other residue for Gly (or Pro for any other residue) one should reduce the entropy of the unfolded state. However, for such a mutation to stabilise the folded protein, its location should be carefully chosen so as not to reduce any noncovalent interactions and to ensure a similar conformation to the residue that was replaced. Mathews et al. (1987) replaced glycine residues with alanine residues to reduce the entropy of unfolded T4 lysozyme, and found these mutations stabilised the protein.

The effect of Pro residues on protein stabilisation has been studied by site directed mutagenesis in T4 lysozyme. Mathews et al. (1987) made a Lys-Pro substitution in T4 lysozyme which was stabilising. They also made an Ile-Pro substition in T4 lysozyme which reduced the stability of the protein. It was thought that this was because the Ile-Pro mutation created a cavity in the molecule, with a resulting loss of van der Waals interactions (Dixon et al., 1992). The degree of enthalpic destabilisation in the Ile-Pro mutant was greater than the degree of entropy gained by introducing the Pro residue, therefore the mutation was destabilising.

1.4.3.1g Resistance to covalent destruction.

Irreversible enzyme denaturation by temperature induced covalent modifications (i.e. cysteine destruction, asparagine and glutamine deamidation and the hydrolysis of peptide bonds [Volkin & Middaugh, 1992]) becomes significant only at elevated temperatures. In their studies of α-amylases Tomazic & Klibanov, (1988), showed that the enzymes from *Bacillus amyloliquefaciens* and *Bacillus stearothermophilus* were inactivated by unfolding, while deamidation was responsible for inactivating the more rigid, and more stable enzyme from *Bacillus licheniformis.* Hensel et al. (1993) investigated glyceraldehyde-3-phosphate dehydrogenase (GAPDH) from a thermophile and a hyperthermophile, and concluded that the hyperthermophilic GAPDH had a greater resistance to covalent modification than the thermophilic GAPDH.

Cysteine residues, which have the potential to form disulphide bridges and which have been shown to stabilise thermophilc enzymes at temperatures up to 60-80°C (Matsumara et al., 1989), are often absent in enzymes isolated from hyperthermophiles (Volkin & Klibanov, 1987; Burley & Petsko, 1985; Russell et al., 1997). Gln and Asn residues are also reduced in number and or buried from solvent in hyperthermophilic enzymes compared to their mesophilic counterparts (Russell et al., 1997).

1.4.3.1h Stabilisation of loops/reduction in flexibility

Loops are believed to be important in protein stability as they might provide initiation points for unfolding (Leszczynski & Rose, 1986). Consequently, one might expect to find shorter loops in thermophilic enzymes. This has been observed in the structural comparison of IPMDH from two mesophiles and a thermophile (Wallon et al.,1997). In the comparison of the crystal structures of CS from a mesophile and a thermophile (Russell et al., 1994), four loops present

in pig CS are absent in *Tp*CS, and other loops located next to highly flexible α-helices in pig CS are shortened in *Tp*CS. *Pf*CS has the same four shortened loops, and a further two loops that are significantly shorter compared to pig CS. It is believed that these changes increase the compactness of a protein reducing its flexibility.

1.4.3.2 Extrinsic determinants of thermostability

All of the previously described thermostabilising features are of an intrinsic nature. However, in some cases where a hyperthermophilic enzyme has been expressed in a mesophilic host, the recombinant protein has a decreased stability to temperature and pressure (Purcarea et al., 1997). Such observations indicate that some extrinsic factors play a role in protein stability in vivo.

It is thought that some hyperthermophilic Archaea accumulate compatible organic solutes at high temperatures. For example, *P.furiosus* (Martins & Santos, 1995), and *Methanococcus igneus* (Ciulla et al., 1994) both accumulate di-myo-inositol-1,1'(3,3')-phosphate under such conditions, and other Archaea accumulate cyclic-2,3-bisphosphoglycerate (Hensel & Konig, 1988), or trehalose (Martins et al., 1997). It is thought that these solutes may have a role as intracellular thermoprotectants, and there is evidence that they have a direct effect on enzyme stability (Hensel, 1993; Ramos et al., 1997). Additionally, the stability of enzymes *in vitro* is often increased in the presence of ligands, suggesting another mechanism of *in vivo* stabilisation. This has been observed in our group where the presence of citrate enhances the thermostability of citrate synthases.

Another extrinsic determinant of thermostability is pressure. A number of groups have investigated the effects of pressure on the thermostability of proteins

from organisms from a range of habitats. Athes et al. (1997) found that moderate pressures of between 50-250Mpa exerted a protective effect against thermal denaturation for three β galactosidases tested. Michels & Clark (1997) found that the thermostability and activity of a protease from *Methanococcus jannaschii*, an extremely thermophilic deep sea methanogen, was greatly affected by pressure. When they raised the ambient pressure to 500MPa the reaction rate at 125°C increased 3.4 fold and the thermostability 2.7 fold. Summit et al. (1998) postulated that stabilization by pressure may result from pressure effects on thermal unfolding or pressure retardation of unimolecular inactivation of the unfolded state.

1.5 Citrate Synthase: a case study;

To identify potential thermostabilising features of CS from the Archaea Russell et al. (1997) compared the structures of CS from a mesophile (pig, 37°C), a thermophile (*T.acidophilum*, 60°C) and a hyperthermophile (*P.furiosus*, 100°C). Fig 1.5, an α carbon trace (orthogonal view) of CS from pig, *Tp*, and *Pf*, shows the strong structural homology between CS from *Pf*, *Tp* and pig despite their amino acid sequence differences (*Pf*CS exhibits 30% and 42% identity to pig and *Tp*CS respectively).Given the overall similarity in the folds, active sites and subunit interfaces of the CS's it has been possible to observe trends in the CS structure moving from a mesophilic CS to a thermophilic CS, and to hyperthermophilic CS. The most significant of these are an increased compactness of the enzyme, a more intimate association of the subunits, an increase in the number of intersubunit ion pairs, and a reduction in the number of thermolabile residues. This comparison has taken into account the fact that the structure of *Tp*CS is in the open form (ligand free) while the other structures were elucidated from the closed form (ligand bound).

Increased compactness: *Tp*CS has four loops that are significantly shorter than the corresponding loops in pig CS; *Pf*CS has the same four shortened loops, and a further two loops that are significantly shorter compared to pig CS. Figure 1.6 shows a schematic representation of the loop region of pig, *Tp* and *Pf*CS between helices B and C. One can see from this figure that there is a reduction in loop size from the pig to *Tp* to *Pf*CS. As one goes up the temperature scale there are also fewer cavities. Cavities exist in virtually all large proteins and may be required to make the protein functional by reducing the strains caused by dense packing of the hydrophobic side chains. Cavity filling studies have shown that cavities are a source of instability within a protein (Akasako et al., 1997; Ishikawa et al., 1993). *Tp*CS has only 3 cavities compared to 7 in pig while *Pf*CS had none. Thus the absence of large cavities in *Pf*CS may contribute to its thermostability by eliminating areas in the structure where non-optimal packing occurs. *Pf*CS has also a reduced accessible surface area and volume compared to pig CS and the percentage of atoms buried is greater. These differences are not just because of the 35 amino acid N-terminal extension in pig CS since identical calculations for an N-terminally deleted (35 residue) pigCS shows that *Pf*CS is still more compact.

Improved monomer-monomer association: In *Pf*CS the monomer-monomer association is made up of two parts: the 8 α-helical sandwich, which is the main component, and the interaction of the C-terminus of one monomer with the partner subunit. The complementarity of packing can be viewed by way of a gap volume index, defined as the interface gap volume divided by the interface surface accessible area (Jones and Thornton.,1995). The gap volume index for pig is 1.86 Å , while that for *Tp*CS is 1.10 Å and for *Pf*CS is 1.17 Å, a smaller figure being indicative of closer packing. However, in the 3D structure of the

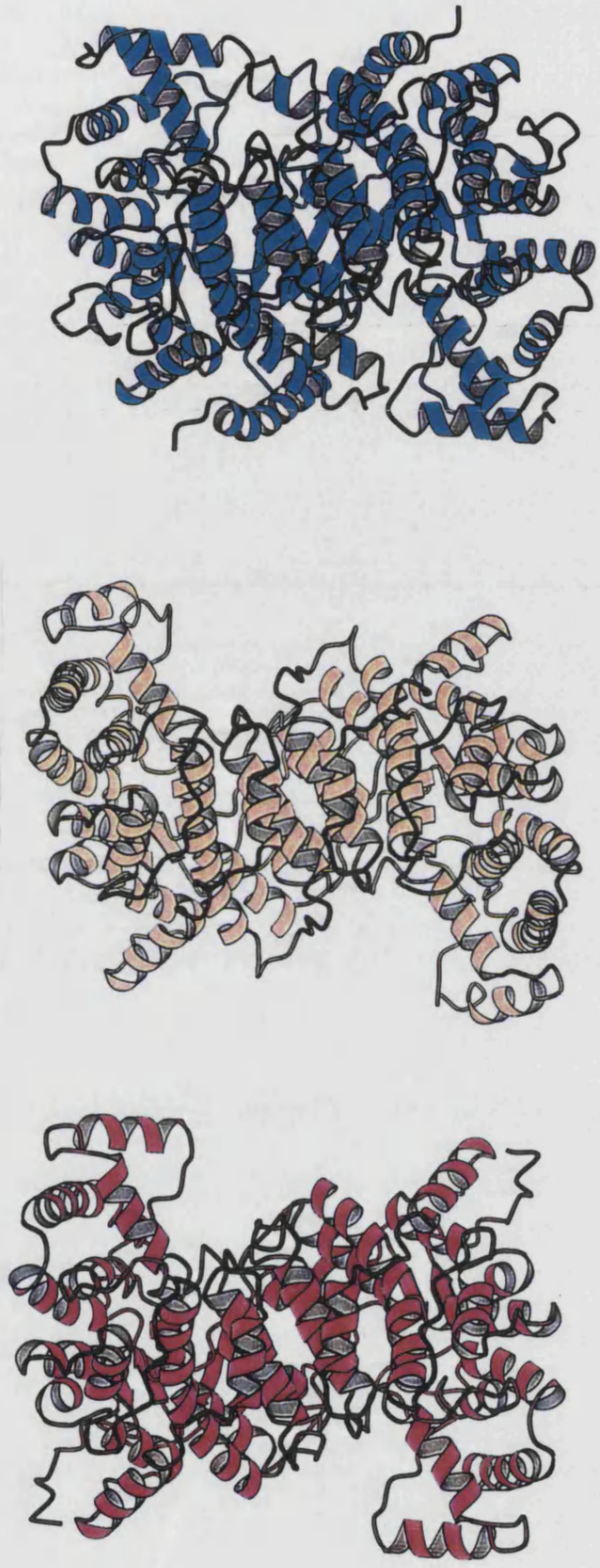


Figure 1.5 Ribbon diagrams of an orthogonal view of pigCS, *Tp*CS, and *Pf*CS. Pig CS is coloured in blue, *Tp*CS in gold and *Pf*CS in red.

open form of *Tp*CS the final 14 C-terminal residues are missing, and if a gap volume index is recalculated for a 14 C-terminal deleted *Pf*CS, then a figure of 1.02 Å is obtained. This improved complementarity may allow the two monomers to remain in an intact dimer at high temperatures through optimised packing interactions.

In addition to the improved complementarity the interactions at the subunit interface in *Pf*CS are strengthened through an increased number of intersubunit ionic interactions within the 8 α-helical sandwich region. Using a cut-off distance of 4 Å there are no ion pairs in pigCS in this intersubunit region, while in *Tp*CS there is one ion pair. At the subunit interface of *P.furiosus*, however, there are 5 residues that make up a 5 membered ionic ring network at either end of helices G and M. Further details of this network are given in the introduction to chapter five.

In *Pf*CS the C-terminus of one monomer embraces the N-terminus of the other monomer, and this interaction is thought to be anchored by an ionic interaction between residues Glu48-Arg375'. There are no similar interactions observed in pig or *Tp*CS, although until a closed structure is obtained for *Tp*CS this type of interaction cannot be ruled out (further details of the C-terminal interaction are given in the introduction to chapter 4). A calculation of the number of hydrogen bonds for the 3 enzymes showed that there was no significant difference in hydrogen bonding between the three enzymes. However there is a significant increase in the number of intersubunit hydrogen bonds for *Pf*CS (51) compared with pigCS (40).

Increased number of ionic bonds: When the total number of ion pairs was calculated for pigCS, *Tp*CS, and *Pf*CS, using cut off distances between oppositely charged residues of 4, 5 and 6 Å, there was observed to be a

significantly increased number of total ion pairs found in *Pf*CS compared with *Tp*CS, and pigCS (see Table 1.1). There is also a significantly greater number of intersubunit ionic bonds in *Pf*CS than in pig CS or *Tp*CS.

Ion Pairs in Citrate Synthase			
	pigCS	*Tp*CS	*Pf*CS
6 Å Cut-off			
Total ion pairs	82	90	105
Intra ion pairs in A	35	39	42
Intra ion pairs in B	35	41	39
Inter A to B-	12	10	24
5 Å Cut-off			
Total ion pairs	64	65	69
Intra ion pairs in A	26	30	25
Intra ion pairs in B	26	29	24
Inter A to B-	12	6	20
4 Å Cut-off			
Total ion pairs	36	43	43
Intra ion pairs in A	12	21	14
Intra ion pairs in B	12	18	12
Inter A to B-	12	4	17

Table 1.1 The number of ion pairs in CS from pig, *Tp*, and *Pf*, using a 6, 5, and 4 Å cut-off distance between charged groups. A and B refer to the two monomers of the native dimer

A number of ion pairs in *Pf*CS are located in loop regions, so that in non-helical regions of the hyperthermostable enzyme ion pairs help to stabilize these flexible regions, in addition to loop stabilization via shortening. Fig 1.7 shows the ionic interactions present in loop regions of CS from pig, *Tp*, and *Pf*.

Fewer thermolabile residues: of the 20 amino acid residues, four (Asn, Gln, Met, and Cys) may be classified as thermolabile due to their tendency to undergo deamidation or oxidation at high temperatures. Many of the thermolabile residues found in pigCS are replaced in *Pf*CS by a residue that forms part of an

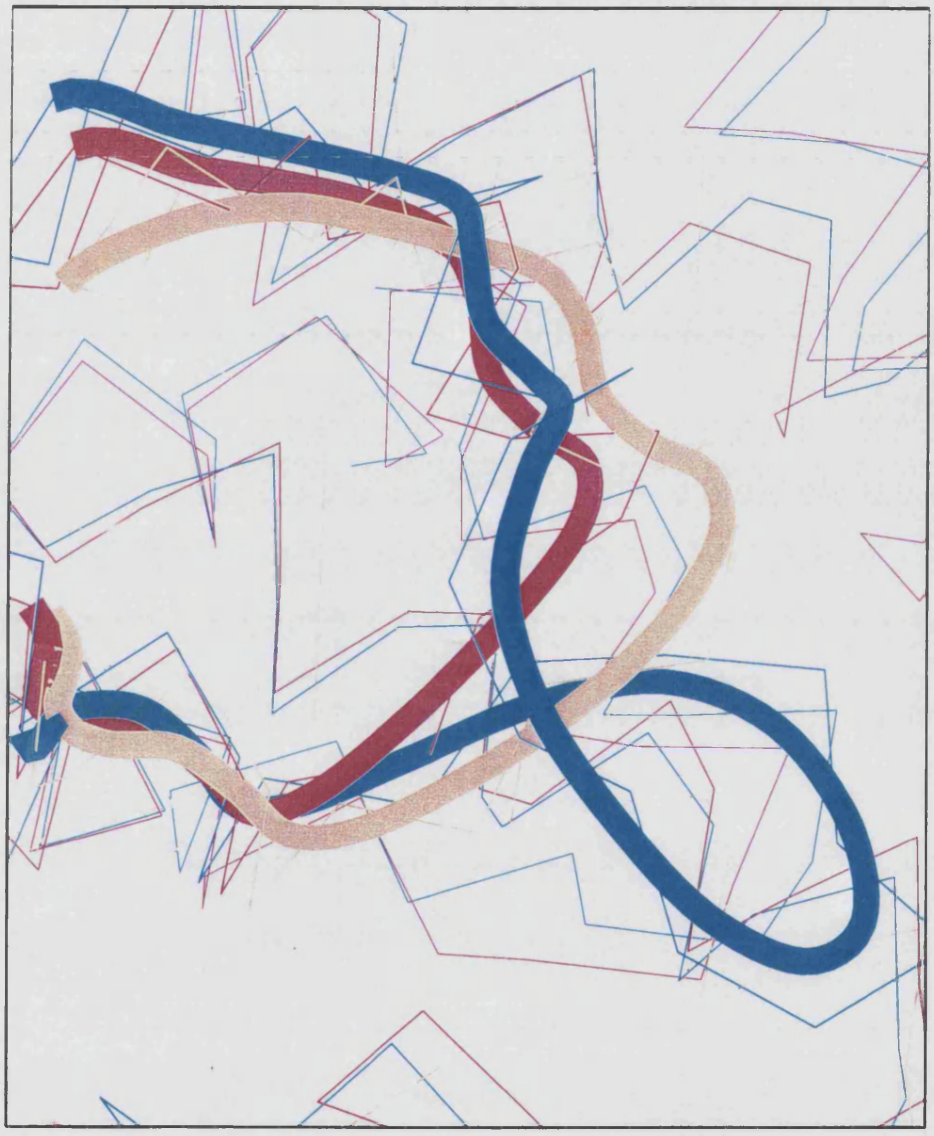
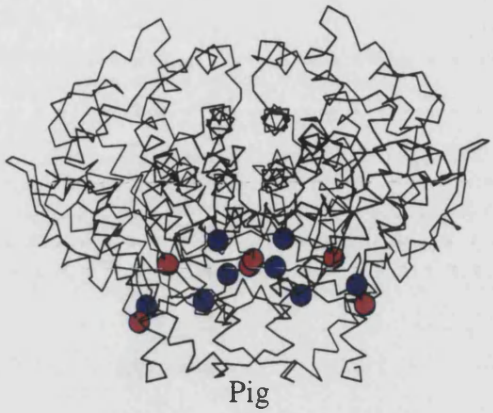
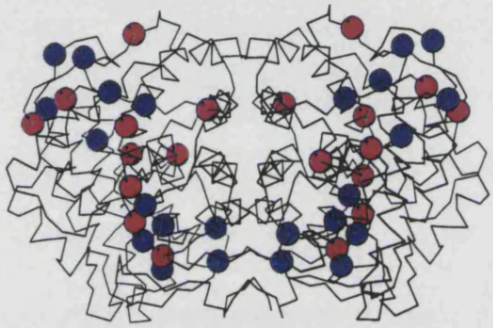


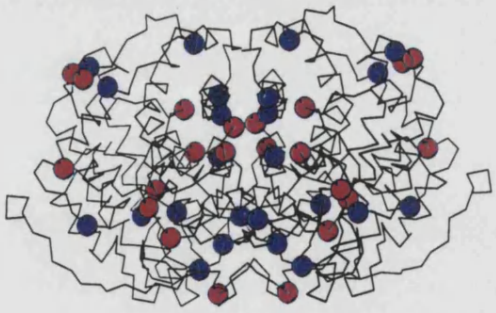
Figure 1.6 Schematic representation of the loop region of pig CS, *Tp*CS and *Pf*CS between helices B and C. The pig CS loop is coloured in blue, the *Tp*CS loop is coloured in gold, and the *Pf*CS loop is in red.



Pig



Thermoplasma

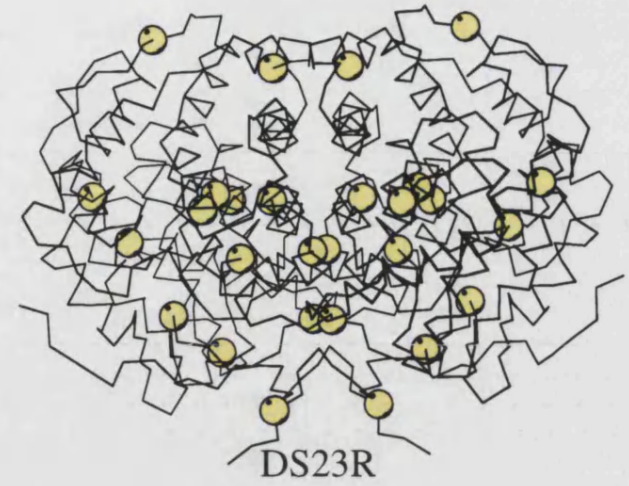


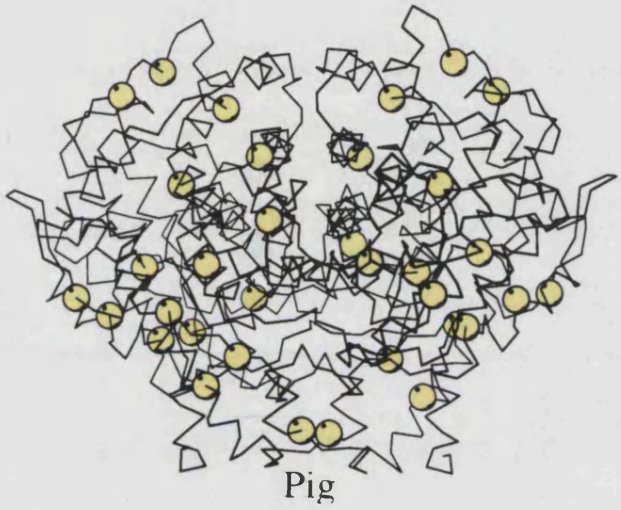
Pyrococcus

Figure 1.7 Ionic interactions present in the loop regions of CS from pig, *Tp*, and *Pf*. The blue and red dots indicate positively and negatively charged residues respectively.

ion pair, also some thermolabile residues are eliminated via loop shortening. *Pf*CS thus has fewer thermolabile residues than *Tp*CS, which in turn has fewer than pigCS. *Pf*CS has fewer Met residues than either pig or *Tp*CS and the Met residues that it has are all effectively buried from solvent. Similarly *Pf*CS has only 5 Asn residues exposed to solvent as opposed to 13 and 14 for pig and *Tp*CS, respectively. *Pf*CS has only 5 Gln as opposed to 10 and 17 Gln residues in *Tp*CS and pig CS, respectively.

Isoleucine Clusters: a significant increase in the number of isoleucine residues is observed in *Pf*CS compared with pigCS and DS23r, see Fig 1.8. These substitutions result in the formation of isoleucine clusters sandwiched between the intersubunit helices (F, G, L, and M) and between helices I and S in each monomer. This clustering of isoleucine residues allows for a very well packed hydrophobic core of the protein. A similar clustering is observed in glutamate dehydrogenase from *P.furiosus* and has been implicated in the stability of this enzyme (Yip et al.,1995).





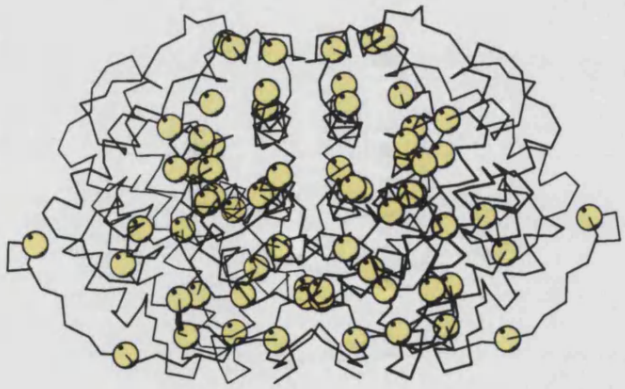


Figure 1.8 Isoleucine residues present in CS from, DS23r, pig and *Pf*CS.

1.6 Other Detailed Comparisons

A number of other groups have carried out structural comparison studies (Wallon et al., 1997; Vogt et al., 1997; Yip et al., 1995,1998). Wallon et al. (1997) compared the structures of 3-iospropylmalate dehydrogenase (IPMDH) from two mesophiles, *Escherichia coli* and *Salmonella typhimurium*, with their thermophilic counterpart from *Thermus thermophilus*, to elucidate the basis of the higher thermostability of the *T.thermophilus* enzyme. They predicted from this comparison that the main stabilising features in the thermophilic enzyme were an increased number of salt bridges, additional hydrogen bonds, a proportionately larger and more hydrophobic subunit interface, shortened N and C termini, and a larger number of proline residues

Vogt et al., (1997) compared 16 families of proteins with different thermostabilities, including subtilisin, glyceraldehyde 3-phosphate dehydrogenase, and lactate dehydrogenase. The proteins were compared on the basis of their respective fractional polar atom surface areas, and the number and type of hydrogen bonds and salt links between explicit protein atoms. In over 80% of the families studied, correlations were found between the thermostability of the familial members and an increase in the numbers of hydrogen bonds, as well as an increase in the fractional polar surface which results in added hydrogen bonded density to water. The number of ion pairs were also found to increase in two-thirds of the families tested.

In their study of the structure of glutamate dehydrogenase (GluDH) from the hyperthermophile *P.furiosus*, Yip et al. (1995) observed that this structure contained a number of ion pair networks, not found in the equivalent mesophilic enzyme (Rice et al., 1996). These were proposed to be involved in maintaining

stability at high temperatures. From this comparison of the structure of *Pf* GluDH with those of mesophilic GluDH's from *E.coli* and *Clostridium symbiosum,* Yip et al. (1998) then went on to determine the amino acid sequence similarity between the *Pf* GluDH and other members of the GluDH family, from species which span a wide range of optimum growth temperatures. The results of this analysis indicated that the ion pair networks became more fragmented as the enzyme stability decreased. These findings were consistent with the role of ion pair networks in the adaptation of enzymes to extreme temperatures.

A problem with the comparative approach is that it is suggestive but does not actually prove whether features predicted to be responsible for an enzyme's thermostability, actually are. Therefore this approach is most useful when it is combined with site directed mutagenesis.

1.7 Are there general rules for protein thermostability ?

On the basis of comparative studies of the primary structures of proteins from mesophilic and thermophilic organisms general "rules" for stability have been proposed (Argos et al., 1979; Menendez-Arias & Argos, 1989). Mutational studies however have shown that many of these purely statistical rules do not apply and the effects of site directed mutations need to be evaluated in a context dependent way. From the work of Van Den Berg et al. (1998) it appears that improvements in thermostability are additive, and that a very small number of changes to a molecule can dramatically alter its thermostability, but these are very site specific. Usher et al. (1998) echo this view stating that an improvement in thermostability is as a result of small changes such as an increase in the number of proline residues and that no single factor predominates. Jaenicke & Bohm (1998) view proteins as 'individuals' that accumulate increments of

stabilization; in thermophiles these come from charge clusters, networks of hydrogen bonds, optimization of packing and hydrophobic interactions.

1.8 Aims and Objectives

The aim of this project was to investigate the structural factors that determine protein thermostability in CS from the Archaea. The genes encoding the citrate synthase from the archaea *P.furiosus* and *T.acidophilum* have previously been cloned, sequenced and expressed in our research group, and the 3D crystal structures have been determined.

From structural comparisons of the CS's from a psychrophile to a hyperthermophile (Russell et al., 1997, 1998) it has been possible to observe trends in the CS structures as one moves up the temperature scale, and to make predictions as to the determinants of thermostability.

The work that is presented here falls into two main areas:

1) The swapping of the small domains between *Tp*CS and *Pf*CS to produce chimaeric synthases. Analysis of these mutants will provide information on the relative contributions to thermostability and catalytic activity of each domain.

2) Mutation of structural regions within the large domain of *Pf*CS that were predicted to play a role in the thermostability of *Pf*CS (from the crystal structure of *Pf*CS).

Analysis of these mutants will test the predictions of thermostabilising features made from comparing the crystal structures of *Pf*CS, *Tp*CS and pigCS.

CHAPTER 2

Materials and Methods

This chapter details the general techniques used in the experiments described in this thesis. Methods that are specific to a particular chapter are described within that chapter.

2.1. Reagents, enzymes and other materials

Molecular Biology: Taq DNA polymerase, T4 polynucleotide kinase, T4 DNA Ligase, and the Rapid Ligase Kit were supplied by Boehringer Mannheim, (Mannheim, Germany). Lysozyme, RNAse A, and phenol: chloroform: isoamyl alcohol (25:24:1) were supplied by Sigma Aldrich Chemical Company, (Poole, UK). Restriction enzymes were supplied by New England Biolabs, (Boston, USA). Oligonucleotides were supplied by Perkin Elmer Biosystems Ltd., (Warrington, UK). Geneclean® III DNA purification kits were obtained from Bio 101 Inc, (La Jolla, USA). The Altered Sites II®mutagenesis kit was supplied by Promega UK Ltd., (Southampton, UK). Quiagen DNA purification kit™ was supplied by Quiagen Ltd., (Crawley, UK). The Snap Miniprep Kit™ was supplied by Invitrogen, (Carlsbad, USA). Ready To Go™ PCR beads were supplied by Amersham Pharmacia Biotech Ltd, (St Albans, UK)

Cell culture: Bacteriological peptone was supplied by Oxoid Ltd, (Basingstoke UK). Bacto-agar, yeast extract and tryptone were supplied by

Difco Laboratories Ltd., (Detroit, USA). Tetracycline, nalidixic acid, and ampicillin (sodium salt) were supplied by the Sigma Aldrich Chemical Company, (Poole, UK). X-Gal, and IPTG were supplied by the Calbiochem-Novabiochem Corporation, (La Jolla, USA)

Electrophoresis: Low melting point agarose, standard agarose, ethidium bromide, and TEMED were obtained from Sigma Aldrich Chemical Company, (Poole, UK). Lambda HindIII DNA markers and 1 kb ladder were supplied by Life Technologies Ltd, (Paisley UK). Pst-I digested lambda DNA, was obtained from Northumbria Biologicals Ltd., (Cramlington, UK). Ammonium persulphate, broad, and low range protein standards were supplied by Bio-rad Laboratories Ltd., (Hemel Hempsted, UK). Protogel™ acrylamide stock solution was supplied by National Diagnostics, (Georgia USA). BioDesign GelWrap ™ was supplied by BioDesign Inc., (New York, USA)

CS purification and assay: Dye matrex red gel A was supplied by Millipore UK Ltd., (Watford, UK). PD-10 Sephadex G-25M columns were supplied by Amersham Pharmacia Biotech Ltd, (St Albans, UK). Amicon centriplus concentrators were supplied by Amicon, (Beverley, USA). Vivaspin 15ml concentrators were supplied by Vivascience, (Lincoln, UK). BSA, Coomassie Brilliant Blue G-250, Coomassie Brilliant Blue R, Coenzyme A (lithium salt from yeast), and oxaloacetic acid, were supplied by the Sigma Aldrich Chemical Company, (Poole, UK). PMSF was supplied by Fluka Chemicals (Gillingham, UK). DTNB was supplied by ICN ImmunoBiologicals, (Lisle, USA). 3MM paper was supplied by Whatman Scientific Ltd.,(Maidstone, UK). Polyethylene glycol grade 400 was supplied by Fisher Scientific Ltd., (Loughborough, UK).

All of the other chemicals used were supplied by Fisons, (Loughborough,

UK), Sigma Aldrich Chemical Company (Poole, UK), Fluka Chemicals, (Gillingham, UK), and Merck, (Lutterworth, UK).

Bacterial strains, culture conditions and plasmids

The *E.coli* strains JM109 (end A1, recA1, gyr A96, thi, hsdR17 (rk-, mk+), relA1, supE44, λ-, Δ (lac-proAB), [F', traD36, proAB, laclq ZΔM15]); and DH5α (Sup E44, Δ lac U169, (ϕ 80 lac Z M15), hsdR17, recA1, end A1, gyr A96, thi-1, relA1); MOB154 (gltA6, galK30, pyrD36, relA1, rpsL129, thi-1 supE44, hsdR4, recA λ- [Wood et al., 1983]), were grown in LB broth, 1.0% (w/v) Bacto tryptone, 0.5% (w/v) yeast extract, and 1.0% (w/v) NaCl) at 37°C, shaking at 200 rpm, or on LB plates, 1.5% (w/v) Bacto agar at 37°C. For the growth of *E.coli* strains carrying plasmids conferring ampicillin resistance, media were supplemented with ampicillin to give a final concentration of 125μg/ml. Bacterial strains expressing tetracycline resistance genes were grown in media and on plates supplemented with 15μg/ml of tetracycline. All bacterial strains were stored as glycerol stocks (24% v/v) at -70°C.

The plasmids pUC19 and pKK223-3 were obtained from Amersham Pharmacia Biotech Ltd. The pALTER-1 mutagenesis vector came with the Altered Sites II mutagenesis kit™. The pREC7 NDE vector was kindly supplied by Linda Kurtz of Washington University School of Medicine, (St Louis, USA).

2.2 Molecular Biology Methods

2.2.1 Small scale preparation of plasmid DNA -'Miniprep'

E.coli strains containing the plasmid of interest were cultured overnight in 1.5ml of LB (containing the appropriate selective antibiotic). Plasmid DNA was

then prepared by a method based on the alkaline lysis method (as described by Sambrook et al.,1989). The overnight culture was centrifuged at 13000 x g for 5 min, the supernatant discarded, and the cells resuspended in 150µl of pre-chilled lysis buffer (50 mM glucose, 10 mM EDTA, 25 mM Tris-HC, pH 8.0). A 200µl volume of 0.2 M NaOH, 1% (w/v) SDS was added and the cells left on ice for 5 min. Then a 200 µl volume of 3M potassium acetate, pH 4.5 was added, and the precipitate removed by centrifugation at 13000 x g for 10 min. The supernatant was extracted twice with an equal volume of 25:24:1 (v:v:v) phenol:chloroform:isoamyl alcohol and the DNA precipitated with ethanol. The DNA was then resuspended in 25µl of double distilled water.

A Snap Miniprep Kit™ was also used for the 'mini' preparation of plasmid DNA. The manufacturers' instructions supplied with the kit were followed.

2.2.2 Large scale preparation of plasmid DNA -'Maxiprep'

This method is based on that described by Ausubel et al. (1987). *E.coli* cells containing the plasmid of interest were cultured in 100ml of LB (containing an appropriate antibiotic) overnight on an orbital shaker at 37°C. The culture was spun down at 6000 x g for 10 min. The pellet was resuspended in 4ml of 50mM Tris-HCl, pH 8.0, 10mM EDTA, 100 µg/ml RNAse A. A 4ml volume of 200mM NaOH, 1% SDS (v/v) was added and incubated at room temperature for 5 min (mixture being inverted several times to ensure good mixing). A 4ml volume of chilled 3.0 M potassium acetate, pH 5.5, was added and left on ice for 15 min. The solution was spun in a Sorvall rotor at 13000 x g for 30 min at 4°C, the supernatant was then transferred to a clean Corex tube and spun for a further 15 min at 13000 x g. The final supernatant was then loaded onto a

Quiagen™ column -100 (pre-equilibrated with 4 ml of 750 mM NaCl, 50mM MOPS, pH 7.0, 15 % ethanol, 0.15% Triton X-100). The column was washed with 2x 10ml of 1.0M NaCl, 50 mM MOPS, pH 7.0, 15% ethanol. DNA was eluted from the column with 5ml of 1.25M NaCl, 50 mM Tris-HCl, pH 8.5. DNA was precipitated by the addition of 0.7 volumes of isopropanol at room temperature. The microfuge tubes were spun at 13000 x g for 30 min at 4°C, the supernatant was discarded, and the DNA pellets were washed twice with 1.5ml of 70% (v/v) ethanol. The microfuge tubes were air dried under vacuum in a rotary evaporator for 5 min before the DNA was redissolved in 80µl of 1 x TE buffer, pH 8.0.

2.2.3 Precipitation of DNA

Ethanol precipitation was used to concentrate a DNA sample, or to bring about a buffer exchange. DNA was precipitated by adding 3M sodium acetate pH 5.2, to give a final concentration of 0.3M, and three volumes of prechilled absolute ethanol. The sample was left in a beaker of dry ice for 5 min before being spun at 13000 x g for 10 min. The DNA pellet was washed twice with 70% (v/v) ethanol and air dried under vacuum in a rotary evaporator for 5 min before being redissolved in an appropriate volume of buffer.

2.2.4 Determining the concentration and purity of DNA

The concentration of a solution of DNA was determined by measuring the absorbance at 260 nm (Sambrook et al., 1989). A 50µg/ml solution of double stranded DNA has an absorbance of 1 unit at 260 nm. Purity of a DNA sample was assessed by calculating the A_{260}/A_{280} ratio of the sample. A ratio of 1.75 or more indicated that the DNA sample was pure.

2.2.5 Restriction Digestion of DNA

Restriction digestion was carried out in accordance with the enzyme manufacturer's instructions. The incubation time and the number of units of enzyme used were varied depending on the amount of DNA being cut. When double digests were carried out, restriction buffers that were compatible with both enzymes were chosen. If the restriction buffers of two enzymes in a double digest were incompatible, then the DNA was cut first with one enzyme, the DNA precipitated to remove the buffer, and resuspended in double distilled water, before being digested with the second enzyme. When restricting DNA prepared by the 'miniprep' method, DNAse-free RNAse A was added to a concentration of 0.1mg/ml, for the last 30 min of the restriction digestion.

2.2.6 5' Phosphorylation of Oligonucleotides

100pmoles of oligonucleotide were phosphorylated by the addition of 5 units of T4 polynucleotide kinase, X10 kinase buffer, 1mM ATP, and sterile distilled water to a final volume of 25µl . Each reaction was incubated at 37°C for 30 min, and was stopped by heating at 70°C for 10 min. Phosphorylated oligonucleotide was stored at -20°C

2.2.7 DNA Ligation

Ligations were carried out in a final volume of 30µl. The insert to vector ratio varied between 3-5:1. The ligations were incubated either at room temperature for 1 hour followed by 37°C for 1 hour, or overnight at 15°C. Each reaction contained 4 units of ligase and 3µl of 10 x ligase buffer. Controls: 1) no ligase and no insert (checks for uncut vector DNA), 2) no insert (checks for vector religation).

2.2.8 Transformation of Competent cells.

Bacterial cells were made competent and transformed with double stranded DNA using the calcium chloride method. A 250µl volume of an overnight culture of cells was used to inoculate 20ml of LB media in a 50ml Falcon tube. This was incubated at 37°C, shaking at 200 rpm in an incubator. After 2-3 hours, when the A_{600} was between 0.3-0.5, the cells were spun down at 3000 x g for 10 min. The supernatant was discarded and the cells were resuspended in 1ml of ice cold 50mM $CaCl_2$, and were then left on ice for a minimum of 15 min prior to starting the transformation.

200µl of freshly-prepared competent cells were used in each transformation reaction. DNA, either from a ligation reaction or uncut, was added to the competent cells and left on ice for a minimum of 30 min, before being heat shocked in a water bath at 42°C for 45 sec. The cells were placed on ice for 2 min before being aliquotted into 1ml of LB in a 10ml Falcon tube. The cells were incubated at 37°C at 200 rpm in an incubator for 90 min before 200µl of cells were plated onto LB plates containing the appropriate supplements. As a control, no vector DNA was added to the competent cells to check for contamination. The plates were incubated overnight at 37°C and transformed colonies selected. (Modified from Sambrook et al., 1989).

2.2.9 Agarose gel electrophoresis

Horizontal agarose gel electrophoresis was used for purification and analysis of small DNA fragments. A 1% (w/v) agarose gel gave efficient separation of linear, duplex DNA of 0.5-8.0 kb. Agarose was dissolved in 1 x TAE buffer (40mM Tris-acetate, pH 8, 1 mM EDTA) by heating in a microwave oven. When the agarose had cooled to approximately 45°C, ethidium bromide

was added to a concentration of 0.5μg/ml. The agarose was then poured into a gel mould, a well forming comb clamped into place and the gel left at room temperature to set. Once set, the comb was removed and the gel placed in a gel tank and covered with 1 x TAE buffer. DNA samples in 1 x loading buffer (0.04 %(w/v) bromophenol blue, 0.04 %(w/v) xylene cyanol, 5 % (v/v) glycerol), were loaded into the wells, and electrophoresed at a constant voltage of 100 V until the xylene cyanol dye front was 1 cm from the edge of the gel. The size separated DNA was then visualised under UV illumination. Markers used included 1 kb ladder, and HindIII and Pst-I digested λ DNA. When a DNA band was to be cut out from an agarose gel, a 1% low melting point agarose gel was used.

2.2.10 DNA purification from agarose gels

Following horizontal gel electrophoresis, the relevant band was visualised using a UV transilluminator and excised from the gel using a scalpel. The DNA was then purified from the gel using a Gene Clean® III Kit. This works in the following manner. The volume of the gel slice containing the DNA is calculated and 3 X this volume of NaI is added to the gel slice. DNA and agarose are highly soluble in NaI. Silica beads (glass milk) are then added to the solution. DNA has a high affinity for the silica beads at alkaline pH's but RNA and proteins do not bind and are therefore eliminated during subsequent washing steps with an ethanol based buffer. DNA is unbound from the silica beads by the addition of water or a low salt tris buffer (pH 8.0).

2.2.11 Polymerase Chain Reaction (PCR)

Reactions were carried out in 0.5ml Eppendorf tubes, which were placed in a Perkin Elmer Applied Biosystems Cetus thermal PCR machine. The PCR

reaction mix contained 12pmoles of a forward and a reverse primer, $MgCl_2$ (at concentrations which varied from 1-6mM), 1 unit of Taq (unless otherwise stated) polymerase, X10 polymerase buffer (supplied with the polymerase), the target DNA to be amplified (0.1μg), and 50μM of each dNTP, unless otherwise stated. Each reaction was made up to a final volume of 100μl with sterile distilled water, and the PCR mix was covered with 50μl of mineral oil to stop evaporation. The PCR machine was programmed to heat for 2 min at 94°C, and then 30 cycles of the following conditions:

94°C for 45 sec (denaturing)

55°C for 2 min (annealing)

72°C for 90 sec (extension)

72°C for 10 min (final step after 30 cycles). This was followed by a cooling step of 4°C until the samples were retrieved for analysis. Where the PCR conditions were different, details are given in the relevant chapters. Samples of the PCR products were taken, bromophenol blue dye added, and the products run on a 1% (w/v) agarose minigel.

Ready To Go™ PCR beads were also used instead of Taq polymerase. 0.1μg of template DNA and 12 pmoles of each primer were added to each 0.5 ml Eppendorf tube. Sterile distilled water was added to give a final volume of 25μl. Each reaction was overlaid with 50μl of mineral oil.

2.2.12 DNA Sequencing

DNA was sequenced using the Perkin Elmer ABI prism 377 Automated DNA Sequencer. Samples were prepared for sequencing by aliquotting 0.5μg

of 'maxiprep' DNA (or Snap miniprep purified DNA), and 3 pmoles of oligonucleotide primer into an 0.5ml microfuge tube. The volume was made up to 12μl using sterile distilled water. Each reaction was covered with 40μl of mineral oil. Samples were elongated using T7 DNA polymerase and fluorescent labels.

2.2.13 Software

The Wisconsin GCG Sequence Analysis software package was used to find unique restriction sites within *Pf*CS and *Tp*CS. The O programme (Jones et al.,1991) was used for looking at the crystal structures of *Pf*CS and *Tp*CS. The DNA sequence analysis package produced in association with the Perkin Elmer Biosystems ABI prism 377 Automated DNA Sequencer was used for interpreting the results from the automatic sequencer. The Enzpack Version 3 programme, written by P. A Williams and B.N. Zaba, School of Biological Sciences, University College of North Wales, Bangor, UK., was used for kinetic analysis of data from enzyme assays. In the case of substrate inhibition the Micro Math® Scientist® for Windows™ programme was used for analysing the kinetic data.

2.3 Biochemical Techniques

2.3.1 Determination of protein concentration

Protein concentration was determined by the method of Bradford (1976). A 0.9ml volume of filtered Bradford reagent, 0.01%(w/v) Coomassie G-250, 4.8%(v/v) ethanol, 8.5%(v/v) H_3PO_4, was incubated with a 0.1ml volume of protein sample at room temperature for 10 min. The absorbance at 595 nm was measured and the protein concentration calculated from a calibration curve prepared with BSA standard solutions (0 - 25 μg/ml)

2.3.2 Protein Expression

A 50ml Falcon tube containing 25 ml of LB (containing ampicillin at a concentration of 125μg/ml) was inoculated with a single colony from a plate. This culture was incubated overnight at 37°C on a shaking incubator at 200rpm. The following day this culture was used to inoculate two 1 litre flasks, each containing 500 ml of LB (ampillin 125μg/ml). This was incubated at 37°C on a shaking incubator overnight and the following day a compound was added to induce expression of protein; the incubation then continued for a further 24 hours. The inducing compound added depended on the vector that host cells contained. For example, if the vector was pRec7-Nde1 based the inducer added was nalidixic acid to a concentration of 50μg/ml. After 24 hours of induction the cells were harvested by spinning in GSA rotor at 8000 x g.

2.3.3 Preparation of cell extracts

An overnight culture of *E.coli* cells, ranging in volume from 10ml-1litre, was centrifuged at 6000 x g for 10 min in a Sorvall GSA rotor at 4°C. The supernatant was discarded and the cells were resuspended in pre-chilled 2 x TE buffer, pH 8.0, 1 mM PMSF, at a cell density of 0.2g cells /ml . The cells were sonicated on ice in 20 second bursts with 30 second intervals, five times. The cell debris was removed by centrifugation at 13000 x g for 20 min at 4°C, in a Sorvall SS34 rotor. The supernatant was stored at 4°C.

2.3.4 Protein purification

The prepared cell extracts were subjected to a heat step (see respective experiments), followed by centrifugation at 13000 x g for 20 min. Each extract was then loaded either onto a 1ml, or 30ml, Matrex Gel Red A column, equilibrated with 50mM Tris-HCl, pH 8, 2mM EDTA. In large scale CS

purifications unbound protein was removed by washing the column with loading buffer until the A_{260} nm baseline returned to zero. In small scale purifications, unbound protein was removed by washing the column with loading buffer, using 3 X the volume of sample loaded onto the column. Citrate synthase was eluted with loading buffer containing 5mM oxaloacetate and 1mM CoA (James et al., 1994). Eluted fractions were assayed for CS activity, and samples were run on a 10% (w/v) SDS-PAGE gel to check for purity. The pure CS fractions were pooled. To remove Ac-CoA and oxaloacetate from the purified enzyme, the pooled fraction was buffer exchanged into 50mM phosphate, pH7.2, 2mM EDTA, using a PD-10 Sephadex G-25M column. The sample was then concentrated using either an Amicon centriplus concentrator, or a Vivascience 15ml Vivaspin concentrator.

2.3.5 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were analysed by the SDS-PAGE method (Laemmli, 1970). A 20ml resolving gel solution (10% (w/v) Protogel acrylamide, 0.375M Tris-HCl pH 8.9, 0.1% (w/v) SDS) was polymerised by the addition of 5 mg of ammonium persulphate (from a 10% (w/v) solution), and 20µl of TEMED, and poured into a gel mould. The gel solution was overlaid with water-saturated n-butanol and left to polymerise at room temperature. The n-butanol layer was removed and the top of the resolving gel washed with distilled water. A 10ml stacking gel solution (5% (w/v) Protogel™ acrylamide, 0.13 M Tris-HCl, pH 6.5, 0.1% (w/v) SDS) was polymerised with 3 mg ammonium persulphate (from a 10% (w/v) solution) and 10 µl of TEMED. The stacking gel was poured to the top of the mould, and a comb placed into the gel, and allowed to set for 1 hour at room temperature. The comb was then removed carefully, and the gel placed

in an electrophoresis tank, which was filled with electrophoresis buffer (0.025 M Tris-HCl, pH 8.8, 0.096 M glycine, 0.1% (w/v) SDS). Bubbles were removed from the lower edge of the gel and unpolymerised acrylamide was washed from the wells of the stacking gel with electrophoresis buffer prior to loading of the protein samples.

Protein samples were prepared for electrophoresis by the addition of an equal volume of sample buffer(0.18M Tris -HCl, pH 6.5, 2% (w/v) SDS, 10% (w/v) glycerol, 2.6%(w/v) β-mercaptoethanol, 0.005% (w/v) bromophenol blue), and boiling for 3 min. Each protein sample was loaded onto the gel in a final volume of between 10 and 15μl. The gel was run at a constant current of 10 mA through the stacking gel and 20 mA through the resolving gel until the dye front was 0.5 cm from the lower edge of the gel. The gel was then removed from the electrophoresis tank and stained with 0.5% (w/v) Coomassie Brilliant Blue R in 9:2:9 (v:v:v) methanol:acetic acid:water for 30 min and then destained overnight with 2:1:7 (v:v:v) methanol:acetic acid:water

2.3.6 Spectrophotometric assay for citrate synthase

Citrate synthase was assayed at 412nm in a Perkin Elmer Lambda Bio or Lambda 11 spectophotometer, using the method of Srere et al.(1963). The thiol group of the reaction product (free CoA) reacts with 5,5'-di-thiobis-2-nitrobenzoic acid (DTNB) to release thionitrobenzoate which absorbs at 412nm and has an absorption coefficient of 13,600 M^{-1} cm^{-1}.

Coenzyme A was acetylated by dissolving 10 mg of coenzyme A in 1ml of Milli-Q water (on ice), and adding $KHCO_3$ to a concentration of 0.17M, and acetic anhydride to a concentration of 11mM. The solution was incubated on ice for 10 min to allow complete acetylation of the coenzyme A.

Enzyme assays were performed at 55°C in a 1ml volume of 20mM EPPS, pH 8.0, 2mM EDTA containing 150μM Ac-CoA, 200μM oxaloacetate and 0.2mM DTNB. One unit of enzyme activity was defined as 1μmol CoA produced per min. Non-enzymic generation of CoA was measured by the omission of oxaloacetate from the assay. Note the EPPS buffer was made up to a pH of 8.45 at 25°C, and the spectrophotometer water bath was set to 60°C (as the temperature of the reaction mix was found to decrease to 55°C after 1 min, the period of the assay).

2.4 Calculation of change in Gibbs free energy of activation of thermal inactivation.

The change in the ΔG of activation of thermal inactivation ($\Delta\Delta G_u^{\ddagger}$) between the wildtype *Pf*CS, and a mutant *Pf*CS, was calculated by using the following equation (Eyring, 1935):

$$\Delta\Delta G_u^{\ddagger} = \Delta G_{u.}{}^{\ddagger}{}_{wt} - \Delta G_u{}^{\ddagger}{}_{mut} = -RT \ln(k_{u.wt}/k_{u.mut})$$

where $\Delta G_{u.}{}^{\ddagger}{}_{wt}$ is the free energy of activation of inactivation of the wildtype enzyme, $\Delta G_u{}^{\ddagger}{}_{mut}$ is the free energy of activation of inactivation of the mutant enzyme, R is the gas constant, T is the temperature in Kelvin, and k_u is the rate constant for inactivation.

CHAPTER 3

Creation of Citrate Synthase Domain Swap Mutants

3.1 Introduction

Complementary to studying the role of specific residues in thermostability, the creation of chimaeric enzymes, in which large parts of the polypeptide chain are exchanged between homologous proteins, offers the possibility of investigating the role of larger substructures in conferring functionality and thermostability to a protein. Using this approach to investigate the structural basis of thermostability of CS from the archaea, it was decided to create CS chimaeric mutants by swapping the region coding for the small domain of *Pf*CS, with that coding for the small domain of *Tp*CS, and vice versa. This chapter describes the site directed mutagenesis carried out on *Pf*CS and *Tp*CS, to allow the region coding for the small domain to be cut out of each gene, and the subsequent ligation of the region coding for the *Tp*CS small domain into the region coding for the *Pf*CS large domain, creating the PYR::tp chimaeric CS. The ligation of the region coding for the *Pf*CS small domain into the region coding for the *Tp*CS large domain, creating the TP::pyr chimaeric CS, is then described. Both chimaeric CS mutants were further mutagenised by PCR to create restriction sites to allow their cloning into the expression vector pREC7-Nde1. The characterisation of both chimaeric CS's, and *Pf*CS and *Tp*CS is then described.

The reason for swapping the small domains between *Pf*CS and *Tp*CS was to determine whether the large domain (which contains the subunit interface and C terminus region), or the small domain (which makes up the active site, and is more flexible than the large domain) plays the dominant role in protein thermostability.

Several groups have used the strategy of making chimaeric enzymes between a mesophilic enzyme and its thermophilic counterpart to investigate the structural basis of thermostability. Numata et al. (1995) made chimaeric 3-isopropylmate dehydrogenases, by gene fusion between the enzymes from the mesophile *Bacillus subtilis* and the extremophile, *Thermus thermophilus*. The stability of each enzyme was approximately proportional to the content of the amino acid sequence from the *T.thermophilus* enzyme. The results suggested that amino acid residues contributing to thermostability distribute themselves, in general, evenly at least in the N-terminal half of the amino acid sequence of *T.thermophilus* isopropylamate dehydrogenase. Eidsness et al. (1997) made chimaeric enzymes by swapping β strand regions of rubredoxins (Rds) between *P.furiosus* and *Clostridium difficile*. The results from this study suggested that a global alignment, which optimised both main chain and side chain interactions between β-sheet strands and core residues, was more important than a few interactions within the β-sheet in conferring *P.furiosus* Rd-like thermostability. Singh & Hiyashi. (1995) made chimaeric β-glucosidases between *Cellvibrio gilvus* and *Agrobacterium tumefaciens.* The temperature optima for the chimaeras were found to be midway between that of both parents, and the C-terminal region of β-glucosidases was found to play an important role in determining enzyme characteristics.

A number of research groups have produced chimaeric enzymes by swapping whole domains. Molgat et al. (1992) swapped the domains between *E.coli* CS and *A.anitratum* CS, and determined from these that the large domains are involved in subunit interactions while the small domains are involved in catalytic activity, and that it was the large domains which determined the chemical stability of the CS. Lebbink et al. (1995) swapped the domains of glutamate dehydrogenase from *P.furiosus* with *Clostridium difficile.* They found the effect of domain exchange on thermostability dramatic. Instead of the hybrid enzymes displaying a thermostability in between those of the parental enzymes, as found in the case of chemical stability, the hybrids were even less thermostable than *C.difficile* glutamate dehydrogenase

In *Pf*CS and *Tp*CS a single segment of DNA sequence codes for the small domain, flanked by segments coding for the large domain. Therefore, to swap both small domains, is a relatively simple manipulation.

3.2 Methods

3.2.1 Site Directed Mutagenesis

The Altered sites 11® in vitro mutagenesis kit, supplied by Promega UK Ltd, was used for the site directed mutagenesis of *Tp*CS and *Pf*CS. The kit uses a vector called pAlter-1, which contains tetracycline and ampicillin resistance genes, however the ampicillin resistance gene has been inactivated by a frame shift mutation. To increase the frequency of mutants identified following the mutagenesis reactions, a mutagenic oligonucleotide supplied with the kit, which restores the ampicillin resistance gene, is used. Theoretically this oligonucleotide should bind to the single stranded DNA, at the same time as the oligonucleotide mutating the gene of interest. Thus, when the mutagenised vector DNA is plated out onto a plate containing ampicillin, the cells that grow should contain the desired mutation. Figure 3.1 shows the steps in the altered sites mutagenesis procedure. The first step in the procedure was to clone the *Tp*CS and *Pf*CS genes to be mutagenised into the pAlter-1 vector. After this the kit manufacturer's instructions, supplied with the kit, were followed with some modifications. Firstly, single stranded template DNA was prepared by the M13 infection method because of problems with mutating denatured double-stranded DNA. Secondly due to problems encountered trying to transform into the *E.coli* strain *ES1301*, another repair$^-$ strain of *E.coli, DH*α , was used. These cells were made competent by the calcium chloride method.

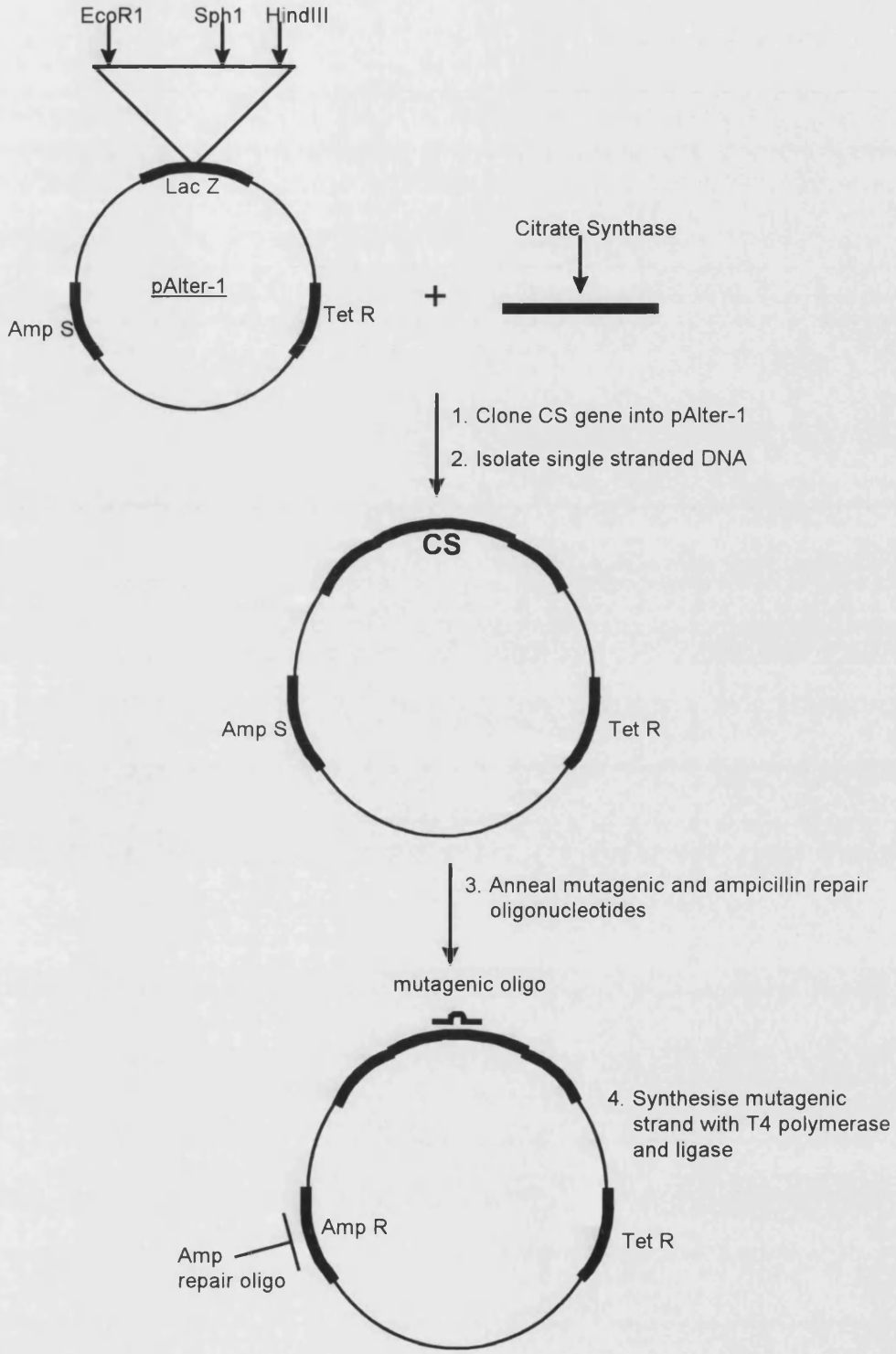


Figure 3.1 The altered sites®II in vitro mutagenesis procedure.

3.2.2 Transfer of proteins for N-terminal sequencing

Following SDS-PAGE (section 2.3.4) the acrylamide gel was equilibrated in transfer buffer pH 8.3, 50mM Tris, 50mM glycine, 20% (v/v) methanol, for 30 min. Twelve pieces of Whatman 3MM paper, and one piece of PVDF membrane (Immobilon), were cut to the same size as the gel, and were equilibrated in transfer buffer for 15 min. Six pieces of the wetted Whatman 3MM paper were placed on the anode plate of the electroblotter, the PVDF membrane was placed on top of this stack, and the SDS-PAGE gel on top of the PVDF membrane, and finally six pieces of Whatman 3MM paper were placed on top of the gel. The gel was electroblotted at 50mA for 3 hours, the membrane was then stained in 0.1% (w/v) Coomassie brilliant R, 50% (v/v) methanol, 7% (v/v) acetic acid for 2 min and destained in 50% (v/v) methanol, 7% (v/v) acetic acid for 1 hour. Finally the membrane was dried in an oven at 40°C for 1 hour and the protein band of interest cut out of the membrane.

Protein samples were sent for N-terminal sequencing to Zeneca Pharmaceuticals , Mereside, Alderley Park, Macclesfield, Cheshire, SK104TG.

3.3 Results

3.3.1 Design of chimaeric domain swap mutants

The GCG programme was used to search for unique enzyme restriction sites flanking the regions coding for the small domain of *Pf*CS and *Tp*CS. To cut out the DNA region coding for the small domain from that coding for the large domain in the *Tp*CS and *Pf*CS genes, the restriction sites Apa1 and BspE1 were selected. These sites were chosen as the Apa1 site occurred naturally once in the *Pf*CS gene, and allowed a cut to be made in the region coding for the loop region between helix M and helix N, and the BspE1 site occurred naturally once in the *Tp*CS gene, and allowed a cut to be made in the region coding for the end of helix R. To allow the region coding for the small domains to be cut out and swapped it was necessary to create an Apa1 site in *Tp*CS (at the same position as it is found in *Pf*CS), and a BspE1 site in *PfCS* (at the same position as it is found in *Tp*CS), see Fig 3.2.

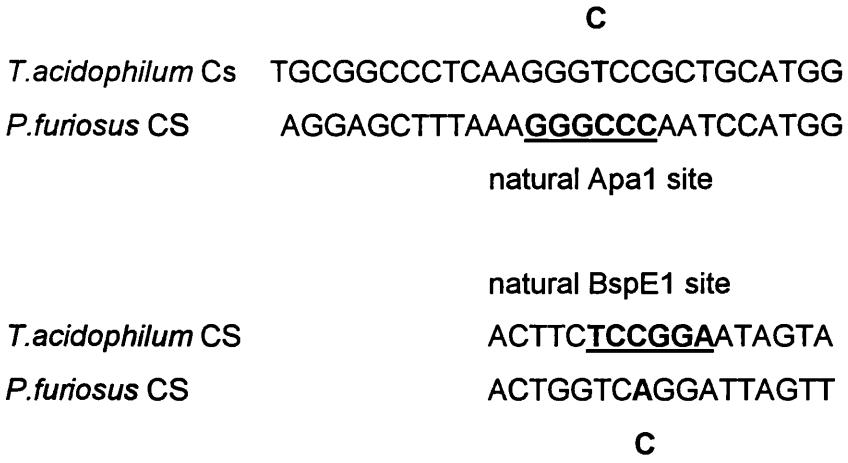


Figure 3.2 The DNA sequence adjacent to the Apa1 and BspE1 sites in *Pf*CS and *Tp*CS, and the base changes necessary to create an Apa1site in *Tp*CS and a BspE1 site in *Pf*CS.

An amino acid alignment between *Pf*CS, and *TpCS* is shown in Fig 3.3. The region shown in italic represents the small domain, the arrows indicate the positions where the small domains were swapped, using the restriction enzymes Apa1 and BspE1.

3.4 Site Directed Mutagenesis

The Altered Sites® II in vitro mutagenesis kit was used to create the Apa1 and BspE1 sites in *Tp*CS and *Pf*CS, respectively. The first step in the altered sites procedure was to clone the *Tp*CS and *Pf*CS genes into the pAlter-1 vector.

3.4.1 Cloning of the *P.furiosus* CS gene into pALTER-1

The *Pf*CS gene was present in a pKK223-3 vector (pBAP3103) into which it had been cloned as an EcoR1-HindIII fragment (Muir et al., 1995). It was decided to use EcoR1 and HindIII restriction enzymes to cut the gene out of the pKK223-3 vector, and clone the EcoR1-HindIII fragment into the pAlter-1 vector.

Approximately 10 μg of pKK223-3 *Pf*CS vector (pBap3103) was digested with EcoR1 and HindIII and the products were separated on a 1% (w/v) agarose gel. A band of approximately 1200bp was excised from the gel and purified using the Gene Clean method. Approximately 5μg of p-Alter-1 DNA was also digested with EcoR1 and HindIII, and the products were separated on a 1% (w/v) agarose gel. A band of 5.7kb in size was excised from the gel and purified using the Gene Clean method. The *Pf*CS gene was then ligated into the pAlter-1 vector, and transformed into JM109 cells, which were plated out onto LB plates containing 0.5mM IPTG, 12.5 μg tetracycline/ml and 40 μg X-Gal/ml.

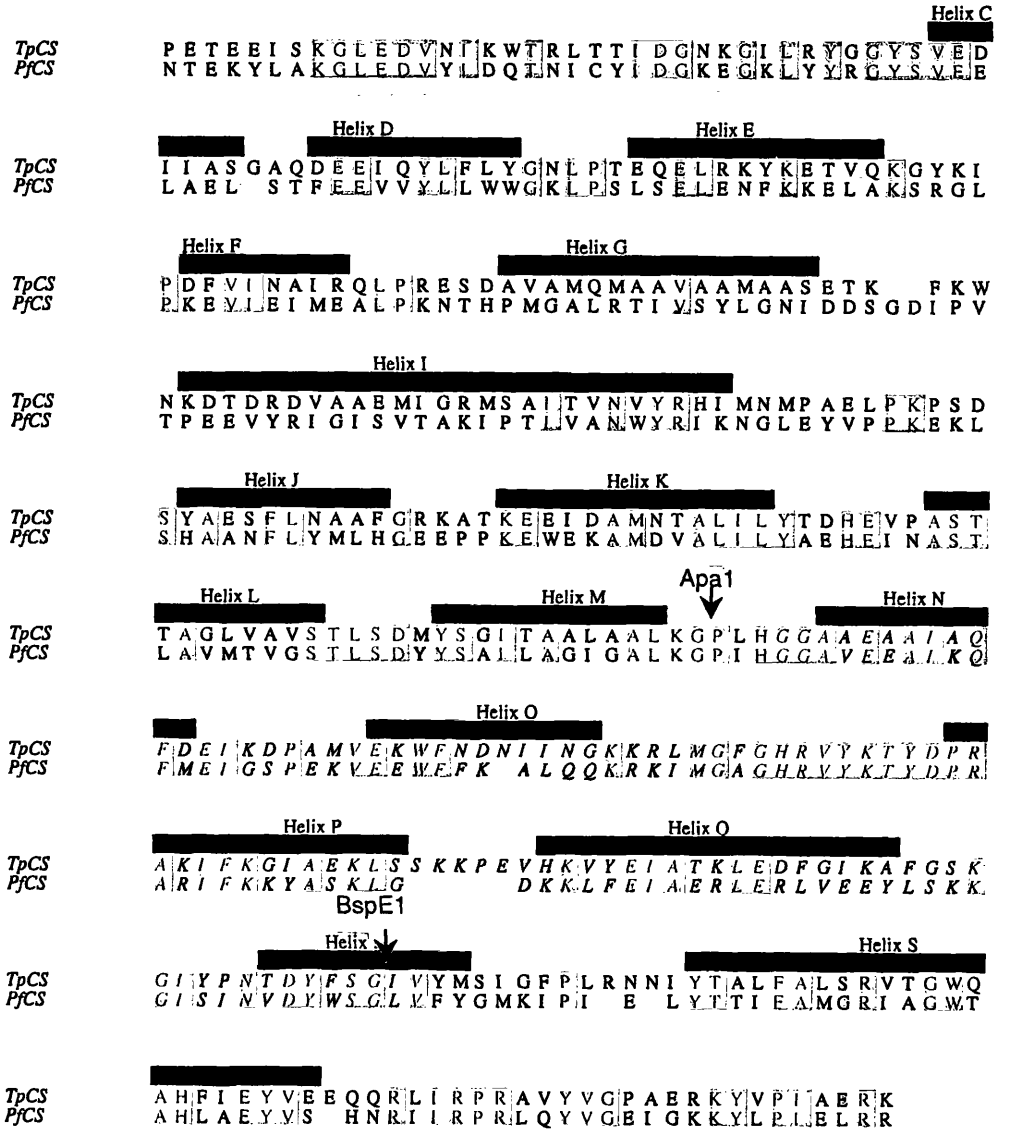


Figure 3.3 Amino acid alignment between *Pf*CS, and *Tp*CS. The sequence shown in italic represents the small domain, and the arrows indicate the positions where the small domains were swapped, using the restriction enzymes Apa1, and BspE1

White colonies were picked, restreaked and grown up overnight for the purpose of preparing 'miniprep' DNA (see section 2.2.1). The DNA from each clone was digested with EcoR1 and HindIII, and the digests run on a 1% (w/v) agarose gel. One clone could be seen to contain an EcoR1-HindIII band of approximately 1200bp; this pALTER-1 vector containing the *Pf*CS insert was designated pBAP3104.

3.4.2 Cloning of the *T.acidophilum* CS gene into pALTER-1

The *Tp*CS gene was present in a pUC19 vector (pCSEH), into which it had been cloned as an SpH1-EcoR1 fragment (Sutherland et al., 1990). The EcoR1 site was at the 3' end of the *Tp*CS gene, as opposed to the 5' end of the gene in the case of the *Pf*CS construct (pBAP3103), and therefore the *Tp*CS gene had to be cloned into the pAlter-1 vector in the opposite orientation to the *Pf*CS gene.

Approximately 10 μg of pCSEH DNA was digested with SpH1 and EcoR1 and separated on a 1% (w/v) agarose gel. A band of approximately 1300bp was excised from the gel and purified using the Gene Clean method. Approximately 5μg of pAlter-1 DNA was also digested with Sph1 and EcoR1, and separated on a 1% (w/v) agarose gel. A band of 5.7 kb in size was excised from the gel and purified using the Gene Clean method. The *Tp*CS gene was then ligated into the pAlter-1 vector, and transformed into JM109 cells, which were plated out onto LB plates containing 0.5mM IPTG, 12.5 μg/ml tetracycline and 40 μg X-Gal/ml. White colonies were picked, restreaked and grown up overnight for the purpose of preparing 'miniprep' DNA (see section 2.2.1). The DNA from each clone was digested with Sph1 and EcoR1 enzymes, and the digests were run on a 1% (w/v) agarose gel. Clone 7 contained a EcoR1-SpH1

fragment corresponding to 1300bp, and this pALTER-1 vector clone containing the *Tp*CS insert was designated pBAP3004.

3.4.3 Mutagenesis Reactions

Two mutagenic oligonucleotides were designed to create the required restriction sites (see Fig 3.4): MAA-1, a 29 base pair oligonucleotide designed to create an Apa1 site in *Tp*CS, and MAA-2, a 30 base pair oligonucleotide designed to create a BspE1 site in *Pf*CS. The restriction sites are highlighted in bold.

MAA-1 5' CGCCGGGAGTT**CCCGGG**CGACGTACCGCC 3'

MAA-2 5' GTTGACTACTGG**TCCGGA**TTAGTTTTCTAT 3'

Figure 3.4. Primers designed to introduce an Apa1 site into *Tp*CS, and a BspE1 site into *Pf*CS.

Both oligonucleotides were supplied by R&D Systems Ltd in an unphosphorylated form, and thus they had to be phosphorylated before being used in the mutagenesis reactions. The phosphorylation and mutagenesis reactions were carried out as detailed in the kit manufacturer's instructions. Following the mutagenesis reactions, transformation into *E.coli* strain DH5α, and transformation into into *E.coli* strain JM109, colonies were picked and used to prepare 'minipreps' of DNA. The *Tp*CS mutated to introduce an Apa1 site was designated pBAP3005, while the *Pf*CS mutated to introduce a BspE1 site was designated pBAP3105. When clones 3005 and 3105 were singly digested with Apa1 and BspE1, respectively, a single band was obtained, showing that the Apa1 and BspE1 restriction sites had been created (gel not shown).

3.4.4 Exchange of the regions coding for the small domains within pALTER-1

The previous section described the creation of restriction sites to allow the small domains to be swapped. This section describes how the regions coding for the small domains were cut out and swapped between *Tp*CS and *Pf*CS.

Large scale restriction digests of clones pBAP3005 and pBAP3105 were set up with the enzymes Apa1 and BspE1. These enzymes use incompatible restriction enzyme buffers and digestion temperatures, so the clones were first digested with Apa1, the DNA was precipitated, redissolved in water, digested with BspE1, and separated on a 1% (w/v) agarose gel (Fig 3.5). The Apa1-BspE1 300bp and 6.8kb fragments were cut out and gel purified using the Gene Clean method.

Overnight ligations were set up with the *Pf*CS 300bp Apa1-BspE1 fragment and the *Tp*CS 6.8kb Apa1-BspE1 fragment, and the *Tp*CS 300bp Apa1-BspE1 fragment with the *Pf*CS 6.8kb Apa1-BspE1 fragment. The ligations were transformed into JM109 cells, which were plated out onto LB plates containing ampicillin at 100μg/ml. White colonies were picked from these plates, restreaked and grown up overnight for the purpose of preparing 'miniprep' DNA (see section 2.2.1). A diagrammatic representation of the Apa1-BspE1 restriction digestion of pBAP3005 and 3105, and swapping of the small domains is shown in Fig 3.6

'Minipreps' from cells transformed with the *Tp*CS large domain / *Pf*CS small domain ligation, and with the *Pf*CS large domain / *Tp*CS small domain

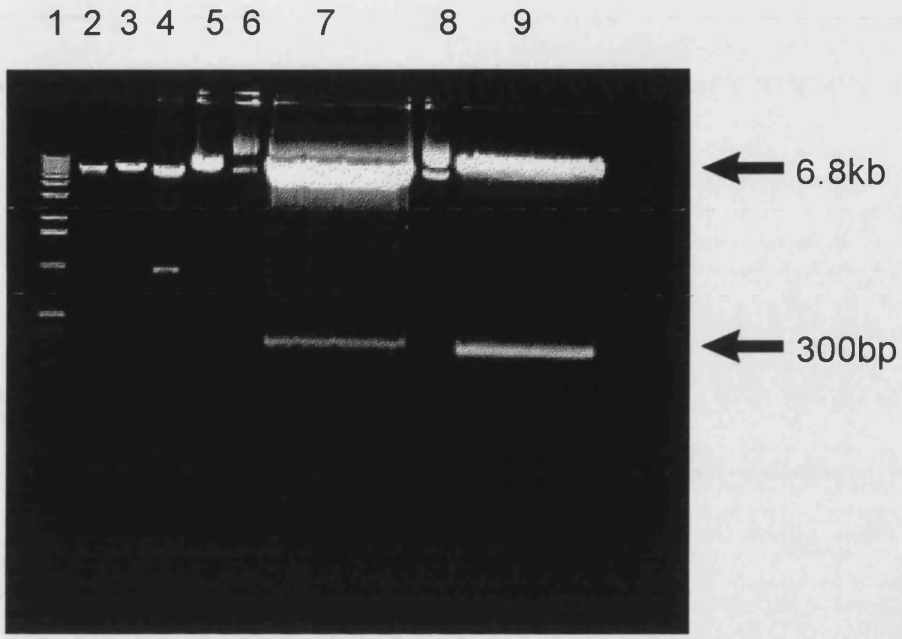
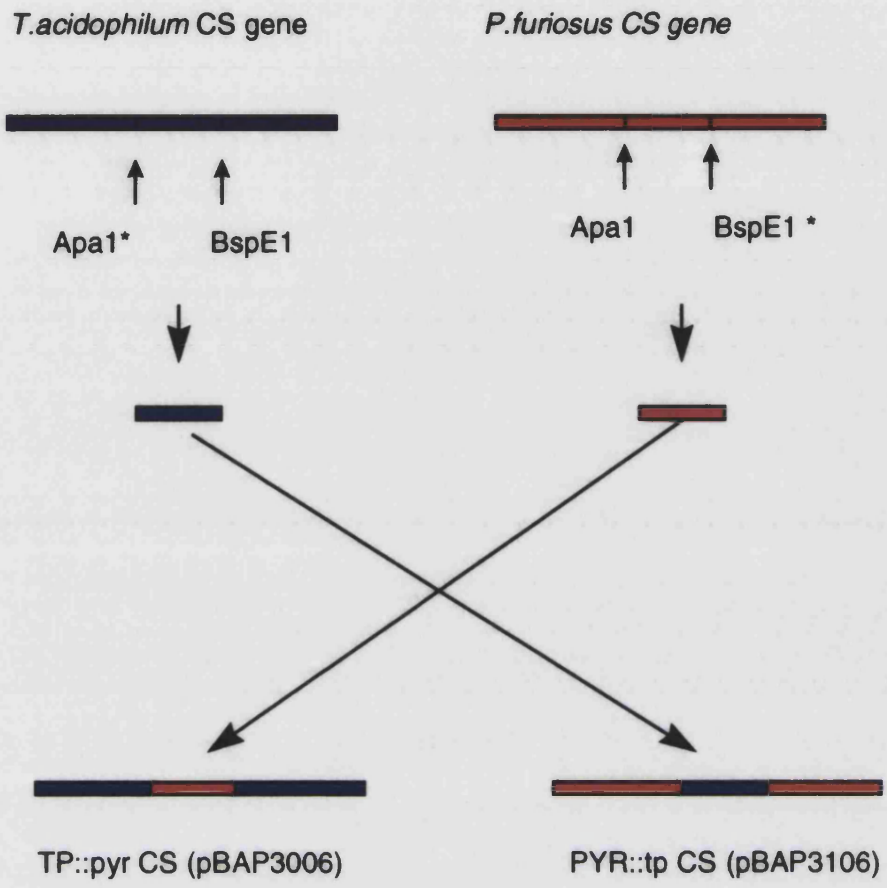


Figure 3.5: Restriction digestion of pBAP3005 and pBAP3105 clones with Apa1, BspE1, and HindIII restriction enzymes. 1) 1 kilobase ladder DNA. 2) pBAP3005 digested with Apa1. 3) pBAP3105 digested with Apa1. 4) pBAP3005 digested with HindIII. 5) pBAP3105 digested with HindIII. 6) Uncut pBAP3005. 7) Large scale digest of pBAP3005 with Apa1 and BspE1. 8) Uncut pBAP3105. 9) Large scale digest of pBAP3105 with Apa1 and BspE1.

ligation, were digested with HindIII restriction enzyme to identify the domain swapped mutants. HindIII enzyme was used as wildtype *Tp*CS has one HindIII site which is located in the small domain, and pAlter-1 also contains a HindIII site; thus the movement of the *Tp*CS can be followed by digestion with HindIII, (see Fig 3.7). A photograph of the gel of a HindIII digest of pBAP3005, pBAP3105, and the two domain swap recombinants is shown in Fig 3.8.



* Site created by site directed mutagenesis

Figure 3.6 Diagrammatic representation of the restriction digestion of *Tp*CS and *Pf*CS with Apa1 and BspE1, to cut out the regions coding for the small domains, (whole genes are in one colour) and swap them to create the chimaeric CS's pBAP3006, and pBAP3106.

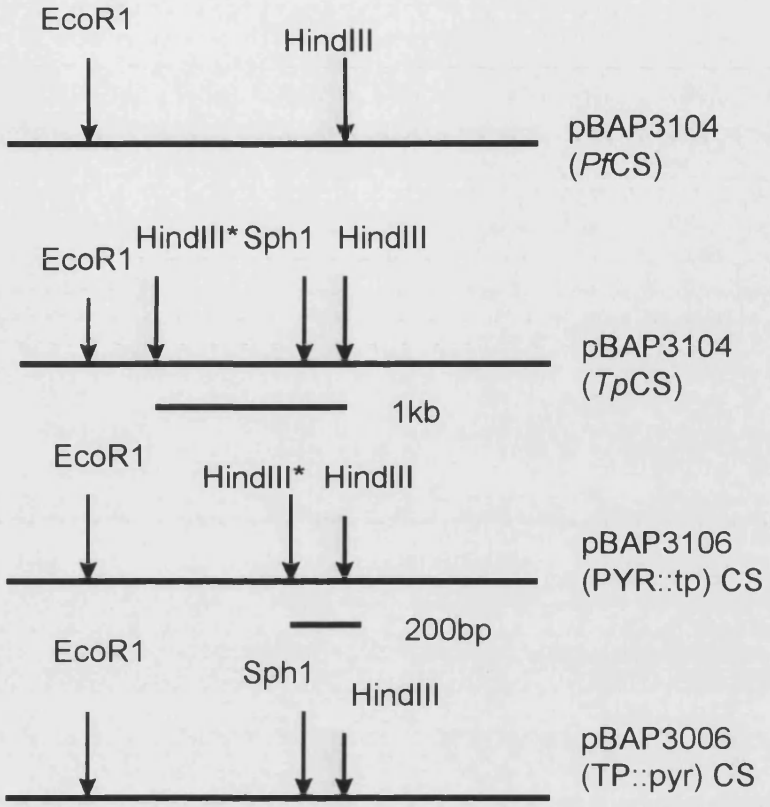


Figure 3.7 Diagram showing the position of the HindIII sites in the CS's. HindIII* represents the HindIII site within the *Tp*CS small domain.

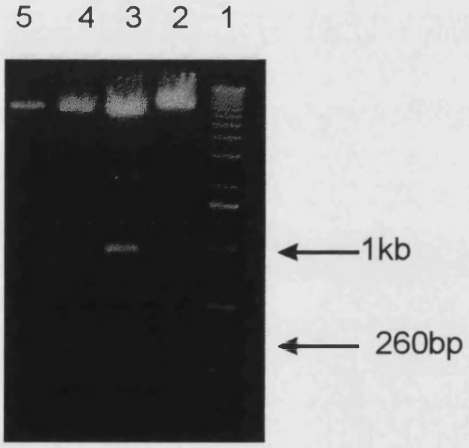


Figure 3.8 HindIII digestion of 4CS's within the pAlter-1 vector. 1) 1kb ladder. 2) pBAP3105. 3) pBAP3005. 4) PYR::tp (pBAP3106). 5) TP::pyr (pBAP3006).

From Fig 3.8 it can be seen that *Tp*CS within pAlter-1 (pBAP3005) digested with HindIII gave bands of approximately 6kb and 1kb, HindIII digestion of *Pf*CS in pAlter-1 (pBAP3105) gave one band of 7kb, while a *Tp*CS large domain/*Pf*CS small domain construct gave one HindIII fragment of 7kb, since this construct will have lost a HindIII site. On the other hand, the HindIII digest of the *Pf*CS large domain/*Tp*CS small domain construct gave a small 260 bp fragment since it had gained an extra HindIII site and the *Pf*CS gene was cloned into pAlter in the opposite orientation direction to *Tp*CS. This was proof that two clones had been created where the small domains had been swapped. The *Tp*CS large/*Pf*CS small domain construct (TP::pyr) within pAlter-1 vector was designated pBAP3006. The *Pf*CS large domain/*Tp*CS small domain construct (PYR::th) within the Alter-1 vector was designated pBAP3106.

3.4.5 Sequencing of pBAP 3006 and pBAP 3106 CS mutants

To determine whether any additional mutations had been created at the domain boundary regions in the domain swap mutants pBAP3006 and pBAP3106, a series of sequencing reactions were carried out. A number of oligonucleotide primers were designed to sequence across the boundary regions between the large and small domains (full details of the sequencing primers used are given in Appendix A).

Fig 3.9 shows the binding positions of the sequencing primers used to sequence clones pBAP3006, and pBAP3106. Primer MAA-3 binds 65bp upstream of the introduced Apa1 site in pBAP3006, primer MAA-6 binds 65bp upstream of the BspE1 site in pBAP3006, primer MAA-5 binds 65bp upstream of the Apa1 site of pBAP3106, and primer MAA-4 binds 65bp upstream of the

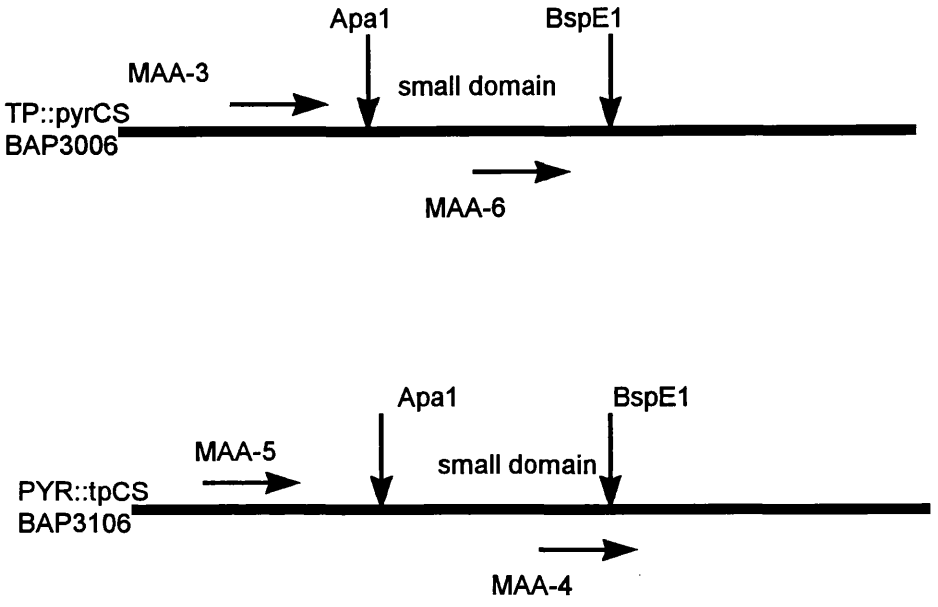


Figure 3.9. Positions at which sequencing primers bind on clones pBAP3006 and pBAP3106.

introduced BspE1 site of pBAP3106. Sequencing of clone pBAP3006 with MAA-3 and MAA-4 primers confirmed that the *Pf*CS small domain had been successfully cloned into the *Tp*CS large domain. The small domain was completely sequenced as were the inter-domain boundary regions; no other mutations were found.

Sequencing of clone pBAP3106 with primers MAA-5 and MAA-6 confirmed that the region coding for the small domain of *Tp*CS had been cloned into the *Pf*CS large domain. From the sequence obtained around the inter-domain boundary region no other mutations were found.

3.5 Expression of Domain Swapped Mutants

Having produced the domain-swapped CS mutants, the next objective was to express the chimaeric CS's, in order to obtain sufficient quantities of protein for characterisation of the mutants. To achieve this it was initially decided to clone the TP::pyr CS into pUC19, and PYR::tp CS into pKK223-3, as these were the vectors used previously for the expression of *Tp*CS and *Pf*CS. However later a decision was made to clone both the TP::pyr CS and PYR::tp CS into a pREC7-Nde1 vector, because this vector gives much higher levels of expression of protein than pUC19 or pKK223-3.

3.5.1 Cloning of PYR::tp chimaeric citrate synthase into pRec7-Nde1

To clone PYR::tp CS into pRec7-Nde1 it was necessary to replace the EcoR1 site with a Nde1 site at the 5' end of the gene in order to use the start codon from the pRec7-Nde1 vector. At the same time It was decided to replace the 3' HindIII site with a Kpn1 site, to facilitate the cloning of the PYR::tp CS gene by eliminating the need to perform partial HindIII digests (because PYR::tp CS {pBAP3106} contained two HindIII sites).

The oligonucleotides shown in Fig 3.10 were designed to add an Nde1 site and a Kpn1 site to the 5' and 3' ends, respectively, of PYR::th CS by PCR.

5' GGGAATTC**CATATG**AATACGGAAAAATACCTTGCT MAA-7 (Nde1)

5' CGG**GGTACC**AATGCTCATTTTATCTACCTCC MAA-8 (Kpn1)

Figure 3.10 Oligonucleotides designed to add an Nde1 and a Kpn1 site to Pyr::tp CS. The DNA sequences in bold represent the Nde1 and Kpn1 restriction sites.

The conditions used in the PCR were:

0.5μg of pBAP3106 DNA (PYR::th cloned into p-Alter-1)

18 pmoles of each primer, except where stated.

The concentration of magnesium chloride in each reaction was varied, from 1-5mM. The other PCR conditions are given in section 2.2.11

Following the PCR amplification, 10μl of each reaction was run on a 1% agarose gel to check whether the PCR had been successful (Fig 3.11). It can be seen that the PCR fragments produced were approximately 1300bp, which is the size of PYR::tp CS gene.

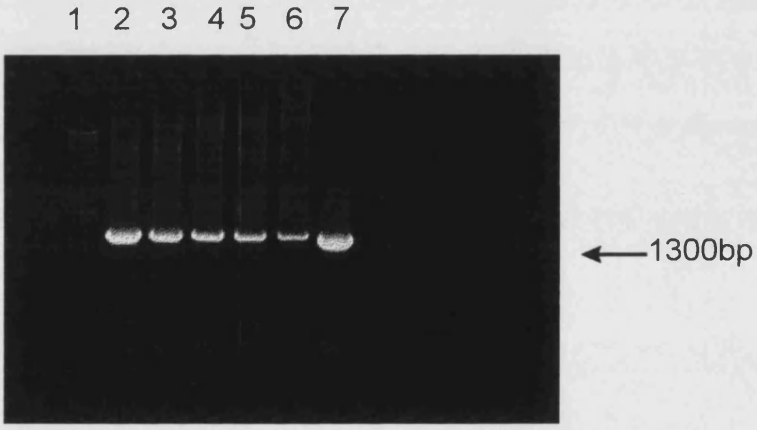


Figure 3.11 PCR products from the amplification of pBAP3106 to add Nde1 and Kpn1 sites to PYR::tp CS. 1) kb ladder DNA. 2) Lanes 2-6 varying concentrations of magnesium chloride from 1-5mM. 7) 2mM magnesium sulphate, 60 pmoles of each primer.

The products from the PCR reactions containing 1 to 5 mM magnesium sulphate, and 60 pmoles of each primer were pooled and purified by the ethanol precipitation method (see section 2.2.3). Restriction digestion of the purified

PCR product and the pRec7-Nde1 vector took place overnight at 37°C. The digests were run on a 1% (w/v) agarose minigel, and the pRrec7-Nde1 and PYR::tp Nde1-Kpn1 bands were cut out of the gel and purified using the Gene Clean method. A ligation was set up using ligase supplied by Boehringer Mannheim and left for 1 hour at room temperature, and 1 hour at 37°C, before being transformed into the CS-minus strain MOB154 (see sections, 2.2.7, and 2.2.8). The transformants were grown overnight at 37°C on LB plates containing ampicillin at a concentration of 125μg/ml; eight colonies were selected, and 'miniprep' DNA was prepared from these for the purpose of restriction digestion. The 'miniprep' samples were digested with both Kpn1 and Nde1 and the digests run on a 1% (w/v) agarose gel). The results of this gel were inconclusive so 'maxiprep' DNA was prepared from some of the clones. When DNA from the 'maxiprep' was digested with Kpn1 a band of 3.8 kb could be seen. Since the pRec7-Nde1 vector is 2.5 kb this showed that the pRec7-Nde1 vector contained a 1.3 kb insert. When this DNA was digested with Nde1 and Kpn1 a band of 1300bp could clearly be seen. This corresponded to the CS gene (see Fig 3.12). From this gel it was concluded that the PYR::tp CS had been cloned into the pRec7-Nde1 vector. This clone was designated pBAP3110.

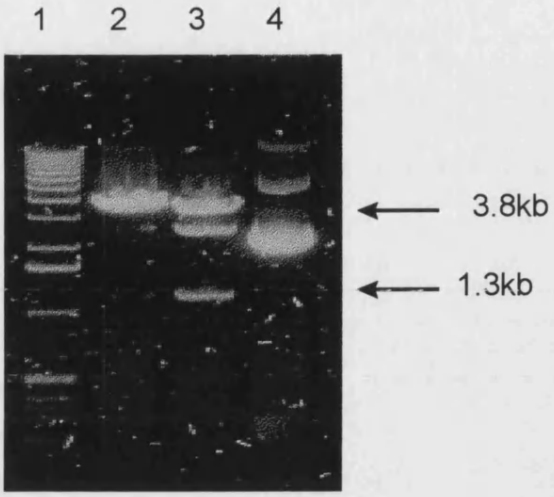


Figure 3.12 Nde1, Kpn1 digest of pBAP3110 construct. 1) 1kb ladder, 2) Kpn1 only, 3) Kpn1 and Nde1 digest of construct, 4) uncut DNA.

3.5.2 Cloning of the wildtype *Pf*CS into the pRec7-Nde1 vector.

To clone *Pf*CS into pRec7-Nde1 as an Nde1-Kpn1 fragment, the same PCR strategy (primers and conditions) was used as for the cloning of PYR::tp CS into pRec7-Nde1 (section 3.5.1).. The overall cloning procedure was also very similar in terms of the purification of PCR product, digestion of product with Nde1 and Kpn1, and ligation into pRec7-Nde1. pRec7-Nde1 vectors containing *Pf*CS were identified by restriction digestion with Kpn1 and Nde1. The pRec7-Nde1 construct containing *Pf*CS was designated pBAP3031.

3.5.3 DNA Sequencing of pBap 3110, and pBap 3031

Having cloned the PCR amplified PYR::tp CS and *Pf*CS into the pRec7-Nde vector, it was important to establish whether any sequence errors had been introduced during the PCR process and subsequent cloning.

A number of primers were designed to obtain the full DNA sequence of

these clones. The DNA sequences of these primers are given in Appendix A. MAA-9 was designed to bind 20 bp upstream of the Nde1 site (on the sense strand) of pRec7-Nde1, to allow read through from the 5' end of inserts cloned into the Nde1 site of pRec7-Nde1. MAA-10 was designed to bind 20 bp downstream from the Kpn1 site (on the antisense strand of the pRec7-Nde1 vector), so as to allow read through from the 3' end of inserts cloned into the Kpn1 site. MAA-11 was designed to bind to DNA adjacent to base number 440 of the sense strand of *Pf*CS to allow the complete sequencing of *Pf*CS related inserts.

0.5μg of pBAP3110 and pBAP3031 'maxiprep' DNA was used in each sequencing reaction along with the above primers. Sequencing was carried out using the ABI automated sequencer (see section 2.2.12). All of the pBAP3110, and pBAP3031 constructs were sequenced and the sequence obtained was that expected for *Pf*CS and the PYR::tp CS chimaera.

3.5.4 Cloning of TP::pyr CS into pUC19

A large scale EcoR1/SpH1 double digest was set up of the vector pUC19 and the TP::pyr CS construct (pBAP3006). The 1300bp CS EcoR1-SpH1 fragment and the double digested pUC19 vector DNA were gel purified using the Gene Clean method, and the purified 1300bp TP::pyr CS EcoR1-SpH1 fragment was ligated into the purified digested vector. Transformants were analysed by restriction digestion with EcoR1 and Sph1; positive clones gave 1300bp and 2.7kb fragments with these enzymes. The TP::pyr CS construct cloned into pUC19 as an Sph1-EcoR1 fragment was designated pBAP3007, Fig 3.13 shows a photograph of a gel of the EcoR1-Sph1 digest of pBAP3007.

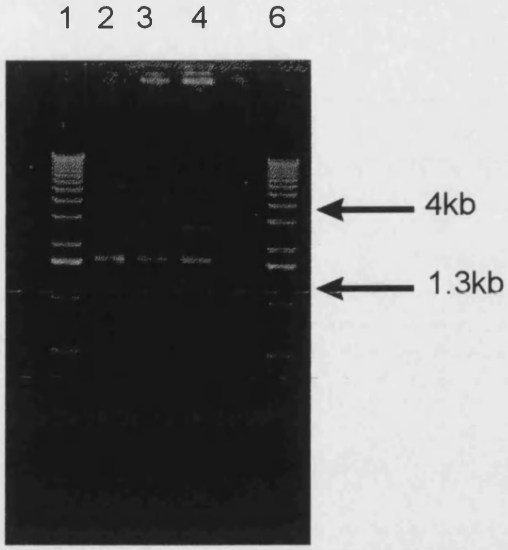


Figure 3.13 EcoR1/Sph1 digest of pBAP3007 construct. 1 and 6) 1 Kilobase ladder DNA. 2) Uncut pBAP3007. 3 and 4) EcoR1-Sph1 digested pBAP3007.

The 1.3kb TP::pyr CS gene and the 4kb single cut pBAP3007 construct bands can be clearly seen, while the 2.7kb band corresponding to pUC19 is weakly visible in Fig 3.15.

3.5.5 Cloning of the TP::pyr CS into the pRec7-Nde vector

The aim of this experiment was to put the TP::pyr CS into a vector that would give a good level of protein expression. The oligonucleotides MAA-12 and MAA-13 were designed to add an Nde1 site onto the 5'end of TP::pyr CS, and a Kpn1 site to the 3' end of TP::pyr CS, to allow TP::pyr CS to be cloned into the pRec7-Nde vector (Fig 3.14)

MAA-12 5' TGTAGAGGTGTA**CATATG**CCAGAAACTG 3'sense primer

MAA-13 5' TGCAGCCCGGG**GGTACC**GGATCCGAAAACAC 3' antisense primer

Figure 3.14 Oligonucleotides designed to add an Nde1 and a Kpn1 site to TP::pyr CS. The Nde1 and Kpn1 sites are shown in bold.

The template DNA used for each reaction was 0.5μg of pBAP3007 (TP::pyr CS in pUC19), 18pmoles of each primer were used, the concentration of $MgCl_2$ was 5mM. For the other PCR conditions see section 2.2.11.

Products of the PCR reaction were pooled together and the DNA was ethanol precipitated (see section 2.2.3), before being digested overnight at 37°C with Nde1 and Kpn1. The pRec7-Nde1 vector was also digested with Kpn1 and Nde1. Both the digested vector and the PCR products were run on a 1%(w/v) agarose minigel, and the Nde1-Kpn1 fragments were purified using the Gene Clean method. The insert and vector DNA were ligated together at room temperature for 1 hour, and then at 37°C for 1 hour, before being transformed into the CS-minus *E.coli* strain MOB154 and plated out onto LB plates containing ampicillin at a concentration of 125μg/ml. The plates were incubated at 37°C overnight and the next day colonies were picked, restreaked and grown up overnight for the preparation of DNA. DNA was prepared by the 'miniprep' method, and was digested with Nde1; the digests were run on a 1% (w/v) agarose gel, (Fig 3.15). Seven clones were screened to select one with the appropriate insert. From Fig 3.15 it could be seen that lanes 3, 5, and 7 corresponded to clones where the TP::pyr CS gene had been successfully cloned into the pRec7-Nde1 vector, whilst lanes 2, 6, and 8 corresponded to

vector only. The TP::pyr CS gene cloned into pRec7-Nde1 construct was designated pBAP3008.

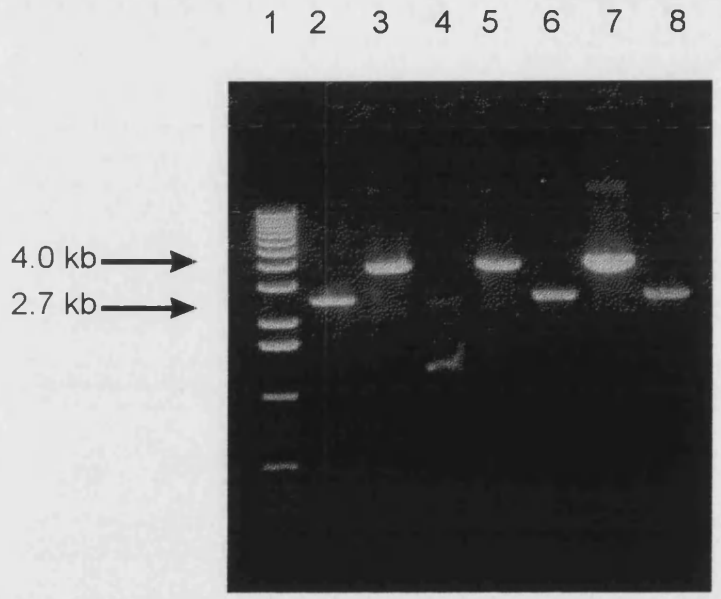


Figure 3.15 Nde1 digest of TP::pyr CS in pRec7-Nde1 recombinants. 1) 1 kilobase ladder. 2-8) Clones 1-7

3.5.6 DNA sequencing of TP::pyr in pUC19 (pBAP3007) and in pRec7-Nde1 (pBAP3008)

To determine whether any sequence errors had been introduced into pBAP3007 by the mutagenesis reaction and subsequent cloning into pUC19, or into pBAP3008 by the PCR reaction, a number of primers were used to enable all of the DNA sequence of pBAP3007 and pBAP3008 to be determined. These primers were MAA-3, MAA-6, MAA-9, MAA-10, and MAA-14. Full details of the primers designed for sequencing are shown in Appendix A, and the binding position of the sequencing primers within the TP::pyr CS mutants are shown in Fig 3.9. MAA-3 binds 65bp upstream of the *Tp*CS Apa1 site, MAA-6 binds 65bp upstream of the *Pf*CS bspE1 site, MAA-14 binds 30bp downstream of the EcoR1 site to the antisense strand of the pUC19 vector (allowing read through

of inserts cloned into the EcoR1 site), MAA-9 was designed to bind 20bp upstream of the sense strand of pRec7-Nde1, to allow read through from the 5' end of inserts cloned into the Nde1 site of pRec7-Nde1. MAA-10 was designed to bind 20bp downstream from the Kpn1 site on the antisense strand of the pRec7-Nde1 vector, so as to allow read through from the 3' end of inserts cloned into the Kpn1 site,

Three pmoles of each primer and 0.5μg of pBAP3007 and pBAP3008 'maxiprep' DNA were sent away for sequencing. The sequence obtained showed that the sequence of both clones was as expected with no other mutations observed throughout the entire sequence.

3.5.7 N-terminal sequencing of PYR::th CS (pBAP3110) and TP::pyr CS (pBAP3007) mutant proteins

While the DNA sequence had been verified for both mutant constructs, it was not known whether the proteins would be expressed using the correct start codon, whether the N-terminal amino acid sequences of the expressed proteins would have the methionine residue cleaved or not, or if any other N-terminal cleavage of the proteins would occur. To answer these questions, large scale cultures (1 litre of LB containing ampicillin at a concentration of 125μg/ml, in 2 litre conical flasks) of pBAP3007 and pBAP3110 were grown in a 37°C incubator for 48 hours. The cells were spun down and the CS's purified on a Matrex Red Gel A column (see methods sections 2.3.2, 2.3.3, and 2.3.4). Samples of both proteins were loaded and run on 10% SDS polyacrylamide gels (see methods section 2.3.5), which were then electroblotted for 2 hours at 50 mAmps to transfer the separated proteins on to nitrocellulose membranes (see section 3.2.2). These membranes were stained and destained as in section 3.2.2., the protein bands were cut out and sent for sequencing.

The amino acid sequences obtained are shown below beside the published N terminal sequences of *Pf* CS, and *Tp*CS.

Published *Tp* CS	Met Pro Glu Thr Glu Glu Ile Ser Lys
TP::pyr (BAP3007)	Pro Glu Thr Glu Glu Ile Ser Lys
Published *Pf* CS	Met Asp Thr Glu Lys Tyr Leu Ala Lys Gly
PYR::tp (BAP3110)	Met Asp Thr Glu Lys Tyr Leu Ala Lys Gly

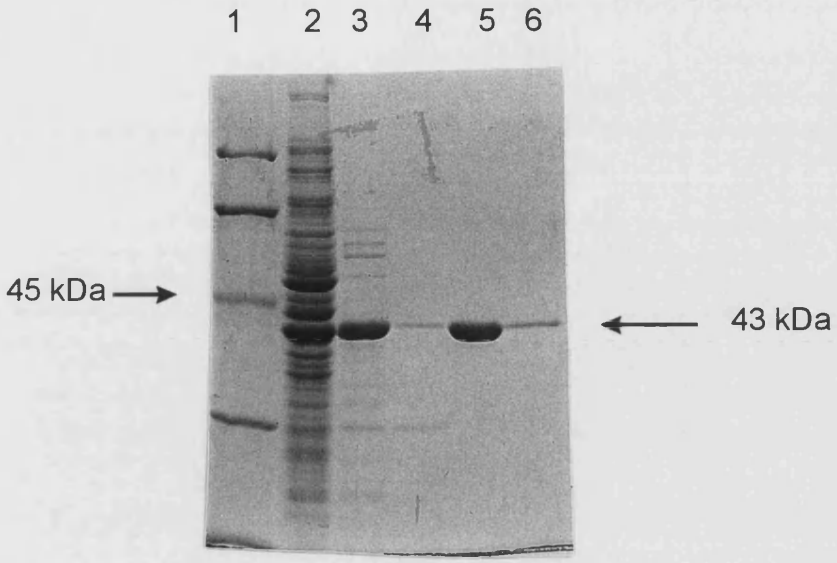
The amino acid sequence showed that the CS produced from the TP::pyr construct had the N-terminal methionine cleaved, while the PYR::th CS did not . The amino acid sequences obtained corresponded to the expected sequences, and no N-terminal cleavage of the proteins occurred. Both domain swapped CS mutants have now been confirmed by restriction digestion, DNA, and N terminal protein sequencing.

3.6 Purification of CS Enzymes

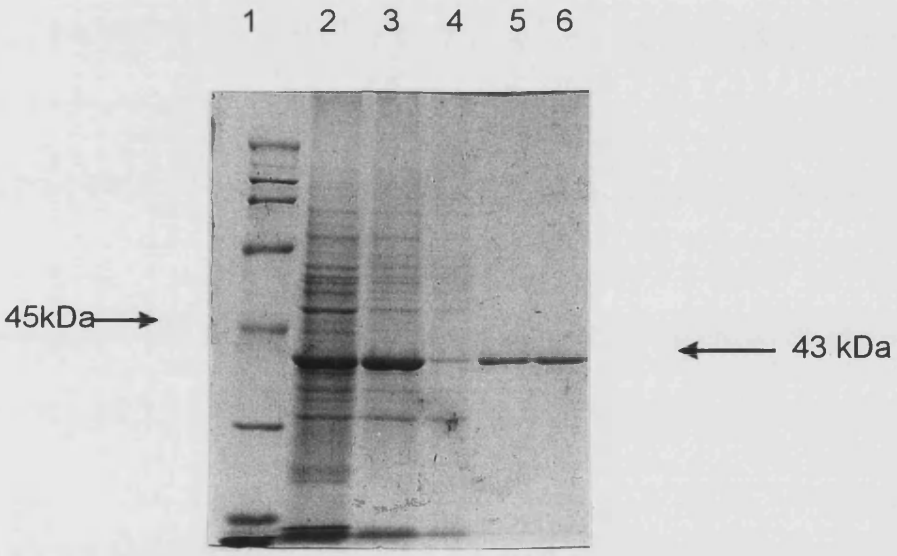
Large scale cultures of the domain swap mutants, and of the wild type enzymes *Tp*CS and *Pf*CS, were set up for expression of the CS's (see section 2.3.2). The innate thermostability of the CS enzymes from *P.furiosus* and *T.acidophilum* permitted a heat step to be used as a preliminary purification step. Previous experiments showed that heating at 80°C for 15 min was suitable as a preliminary purification step for *Pf*CS (Muir et al., 1995), and heating at 65°C for 15 min was suitable as a preliminary purification step for *Tp*CS (James et al., 1994). Preliminary denaturation experiments with the two domain swap CS mutants showed that the PYR::tp CS construct (pBAP3110), could be partially purified using a heat step of 80°C for 15min, and the TP::pyr CS construct (pBAP3008) could be partially purified using a heat step of 60°C

for 15 min.

Following the heat step, the CS's were purified by Matrex Red gel A chromatography (see section 2.3.4). Protein determination was carried out using the Bradford method (section 2.3.1). Figs 3.16 and 3.17 shows some of the SDS-PAGE gels run during the purification of the mutant and wildtype CS's, and Table 3.1 shows the yield and specific activity for each step of the purification.

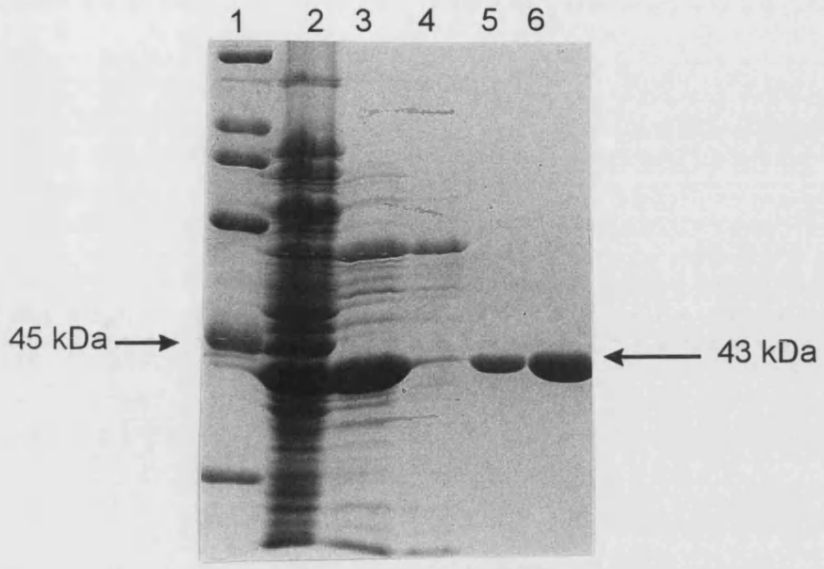


TP::pyr CS purification

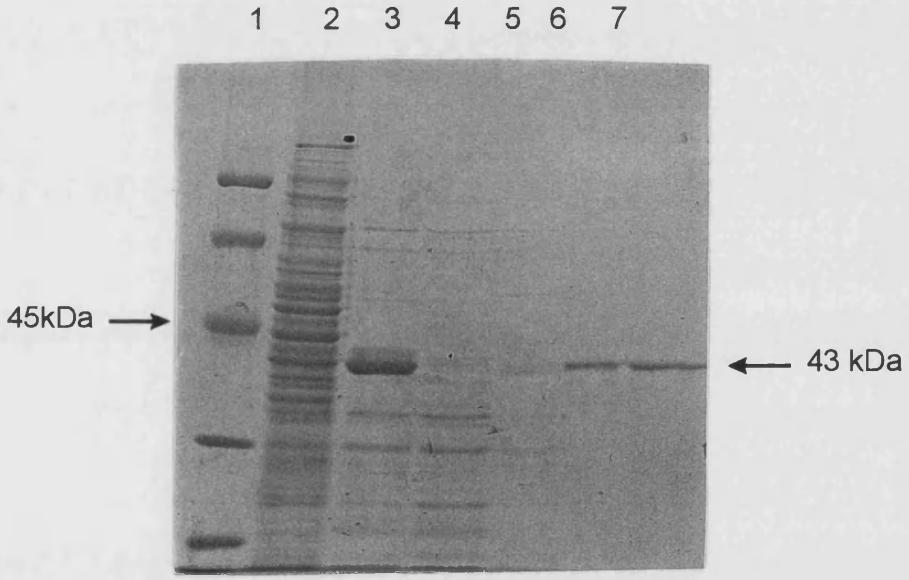


*Tp*CS purification

Figure 3.16 SDS-PAGE analysis of fractions from the purification of TpCS and TP::pyr CS. 1) Low range molecular markers. 2) Cell free extract. 3) Heat treated supernatant. 4) Red gel A unbound fraction. 5-7) Red gel A eluted fractions.



PYR::tp CS purification



*Pf*CS purification

Figure 3.17 SDS-PAGE analysis of fractions from the purification of *Pf*CS, and PYR::tp CS. 1) Low range molecular markers. 2) Cell free extract. 3) Heat treated supernatant. 4) Red gel A unbound fraction. 5-7) Red gel A eluted fractions

Table 3.1 Purification of wildtype *Tp*CS and *Pf*CS and domain swapped mutants. *Indicates that the specific activity figures are low because of substrate inhibition. NB all activity assays were carried out at 55°C.

CS Construct / step	Volume (ml)	Activity /ml (Units/ml)	Total Activity (Units)	Protein (mg/ml)	Total Protein (mg)	% Yield	S.activity (Units/mg)
PYR::tp							
Cell extract	5.0	188	940	55.2	276	100	3.4
Heat step	3.5	164	574	17.8	61	62	9.4
Matrex red gel A	2.4	137	329	14.6	11.5	35	28.6
TP::pyr							
Cell extract	10	11.3	113	41	410	100	0.3 *
Heat step	7	10.6	74	5	34.7	65.6	2.1 *
Matrex red gel A	6	7.3	44	2.2	13.5	38.7	3.3 *
***Tp*CS**							
Cell extract	8	28	222	43	344	100	0.6
Heat step	5.5	27	149	14	78	67	1.9
Matrex red gel A	4.5	18	81	0.3	1.3	37	62
***Pf*CS**							
Cell extract	8	44	351	49	390	100	0.9
Heat step	6	37	225	12.4	74.4	64	3
Matrex red gel A	4	31	123	2.8	11.1	35	11

3.7 Characterisation of the purified CS enzymes

3.7.1 Determination of the kinetic parameters of the CS's

To determine values of K_m and V_{max} for an enzyme that uses two substrates, one should ideally do a '5x5 matrix' of substrate concentrations experiment (Henderson, 1992). However a '5x5' could not be carried out in the case of these four CS's because their K_m values were so low that when both substrates were at a low concentration the rates of reaction could not be accurately measured by the spectrophotometer. Instead the concentration of one substrate was held in excess while the concentration of the second substrate was altered, and vice versa. The high concentrations of substrates used were 150μM and 200μM of Ac-CoA and OAA respectively unless stated otherwise.

Dilutions of each of the four CS's were carried out in 50mM phosphate pH7.2, 2mM EDTA to determine an appropriate enzyme concentration, that would produce a measurable rate on the spectrophotometer, but that did not use up all the substrate too quickly at low substrate concentrations. The CS's were assayed in 20mM EPPS pH 8.0, 2mM EDTA, 100mM KCl, at 55°C. Each CS was assayed 3 times and the direct linear plot of Eisenthal and Cornish-Bowden (1974) was used to determine the K_m and V_{max} values, except in the case of the TP::pyr CS construct which was substrate inhibited by Ac-CoA. In this case, K_m and V_{max} values for the substrate Ac-CoA were determined using the Scientist plot to fit the rates to a curve given by the equation:

$$v = \frac{Vmax}{1+\frac{Km}{S} + \frac{S}{Ki}}$$

The inhibitor constant (Ki) determined for the substrate Ac-CoA from the

Scientist plot was 61(±24). The K_m and V_{max} determination for the substrate oxaloacetate for TP::pyr CS was carried out at an Ac-CoA concentration of 5μM, as this was the concentration of Ac-CoA which gave the highest rate. Mean K_m and specific activity values (± standard errors), and turnover numbers/monomer, measured at 55°C for the four CS's are shown in Table 3.2. Note the specific activity quoted for the TP::pyr CS construct is determined using the mean V_{max} values calculated for the substrate Ac-CoA.

Citrate Synthase	Km OAA μM	Km Ac-CoA μM	Specific Activity (Units/mg)	Turnover number (sec^{-1})
P.furiosus	10 ± 1.3	3.0 ± 0.7	17 ±2	12
T.acidophilum	2 ± 0.2	4.8 ± 0.5	68 ± 4	49
PYR::tp	11 ± 0.5	8.4 ± 1.0	30 ±1	22
TP::pyr	2.7± 1.0	1.7 ± 0.3 *	13 ± 2	9

Table 3.2 Comparison of turnover numbers, Km values for the substrates oxaloacetate and Ac-CoA, and the specific activities of the four CS's at 55°C. * Indicates that this enzyme is substrate inhibited by Ac-CoA.

V against S plots of the data used to determine the K_m and V_{max} values of TP::pyr CS and PYR::tp CS were drawn using the Microcal Origin graphics package (Fig 3.18a-d). A simulated line was plotted through the experimentally determined V versus S points using simulated V and S values calculated from the equation;

$$V = \frac{Vmax\ S}{S + Km}$$

using the determined K_m and V_{max} values. Where substrate inhibition was occurring the simulated V and S values were determined from the substrate inhibition equation;

$$v = \frac{Vmax}{1+\frac{Km}{S} + \frac{S}{Ki}}$$

using the determined K_m, K_i, and V_{max} values. Note each plot was produced from the data from one K_m and V_{max} determination for each enzyme with the substrates oxaloacetate and Ac-CoA. These plots show in the case of figs 3.18a, 3.18b, and 3.18c that the data fit a Michaelis-Menton kinetic model. Whereas in the V against S plot in Fig 3.18d it can be seen that TP::pyr CS is not obeying Michaelis-Menton kinetics with respect to Ac-Coa, rather it is showing a substrate inhibition effect.

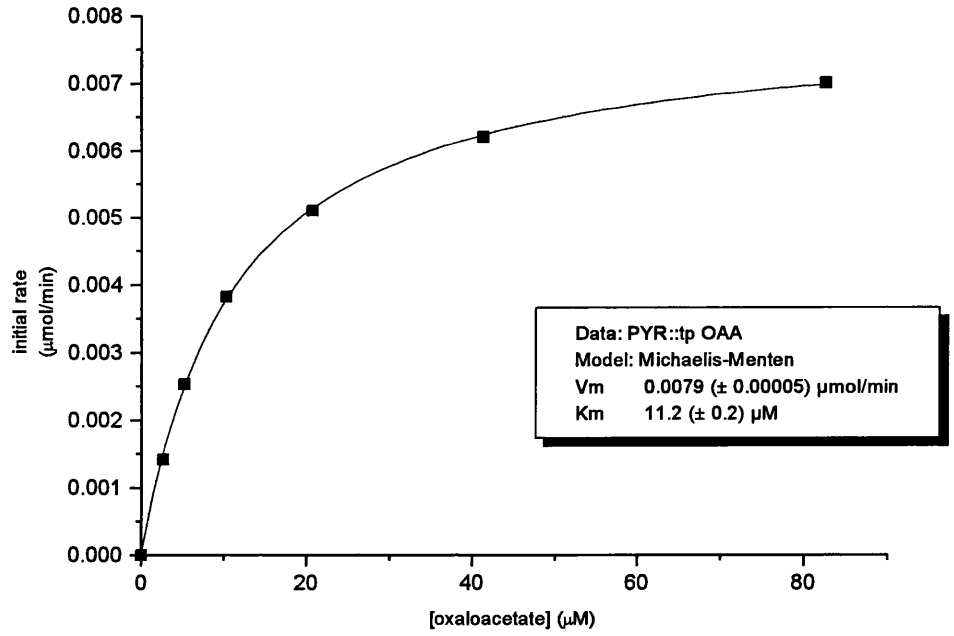


Figure 3.18a V against S plot of the data used to determine the K_m and V_{max} values for PYR::tp CS for the substrate oxaloacetate.

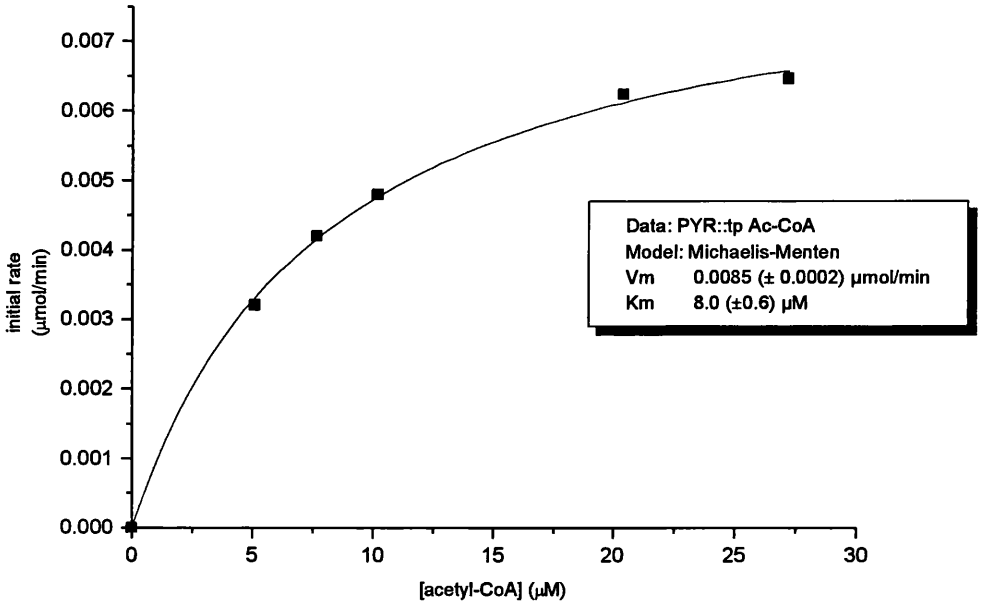


Figure 3.18b V against S plot of the data used to determine the K_m and V_{max} values for PYR::tp CS for the substrate Ac-CoA.

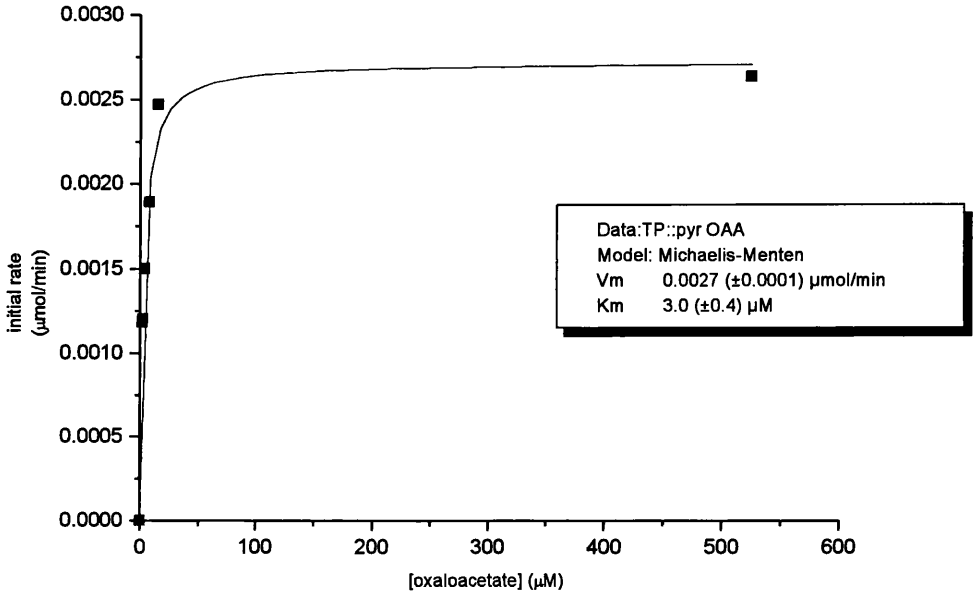


Figure 3.18c V against S plot of the data used to determine the K_m and V_{max} values for TP::pyr CS for the substrate oxaloacetate.

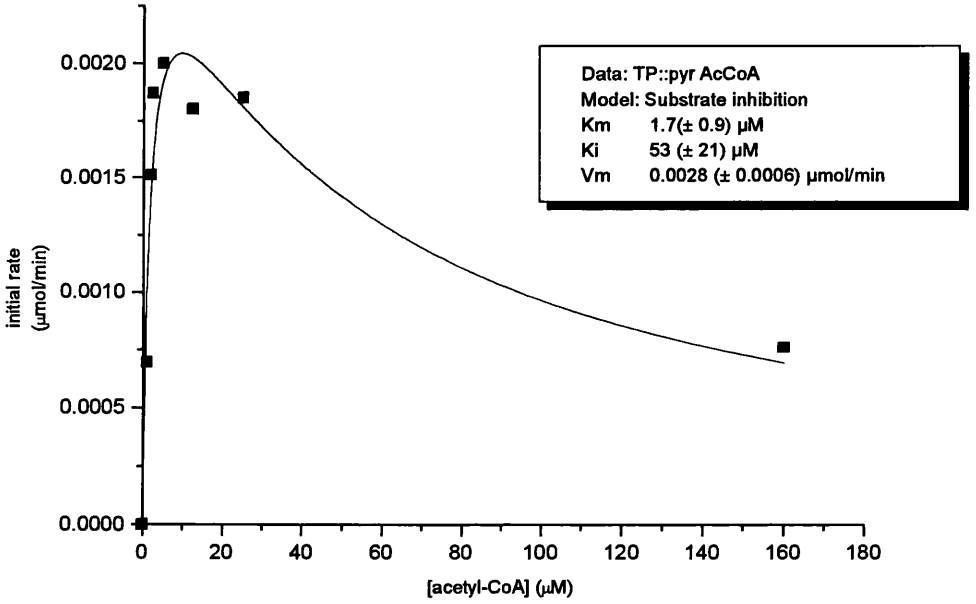


Figure 3.18d V against S plot of the data used to determine the K_m and V_{max} values for TP::pyr CS for the substrate acetyl-CoA.

3.7.2 Thermal stability of the four CS's

The aim of this set of experiments was to compare the mutant CS enzymes *PYR::tp* and *TP::pyr,* with *Pf*CS and *Tp*CS in terms of their abilities to withstand thermal denaturation.

All thermal stability studies were carried out in the presence of 50mM phosphate pH 7.2, 2 mM EDTA using an enzyme concentration of 0.1mg/ml. The buffer was heated to the required temperature in an Eppendorf tube in a heating block, and the temperature inside a second Eppendorf containing buffer only was measured using an electronic temperature probe. When the desired buffer temperature was reached in the control tube, enzyme was added to the first Eppendorf tube, mixed and the timer started. Aliquots were removed at timed intervals and placed into thin walled glass tubes on ice, where they were stored prior to assay. A time zero value was obtained by assaying an 0.1mg/ml concentration of enzyme in the same buffer kept at 4°C. Samples were assayed in triplicate.

It was difficult to choose a single temperature at which to compare the thermal inactivation of all four enzymes, since preliminary experiments showed that the thermostabilities of the four enzymes varied widely. However, a temperature of 85°C was initially chosen to compare the thermal stabilities of the 4 CS's. A graph showing the relative thermal stability of the four CS's at this temperature is shown in Fig 3.19. From this graph it can be seen that the TP::pyr CS was rapidly inactivated at 85°C, and that PfCS and PYR::tp CS were not inactivated at this temperature. For this reason the inactivation of PYR::tp and *Pf*CS was followed at a higher temperature of 95°C (note *Tp*CS was inactivated very rapidly at 95°C). Fig 3.20 shows the relative thermostability of

*Pf*CS and *PYR::tp* CS at 95°C. The inactivation rate constants calculated for the chimaeras and wildtype enzymes are shown in Table 3.3.

Citrate Synthase	Temperature of inactivation (°C)	Inactivation rate constant (min^{-1})
TP::pyr	85°C	0.28± 0.004
T.acidophilum	85°C	0.048± 0.002
PYR::tp	95°C	0.12±0.004
P.furiosus	95°C	No detectable inactivation

Table 3.3 Inactivation rate constants for the chimaeric and wildtype CS's at 85°C and 95°C.

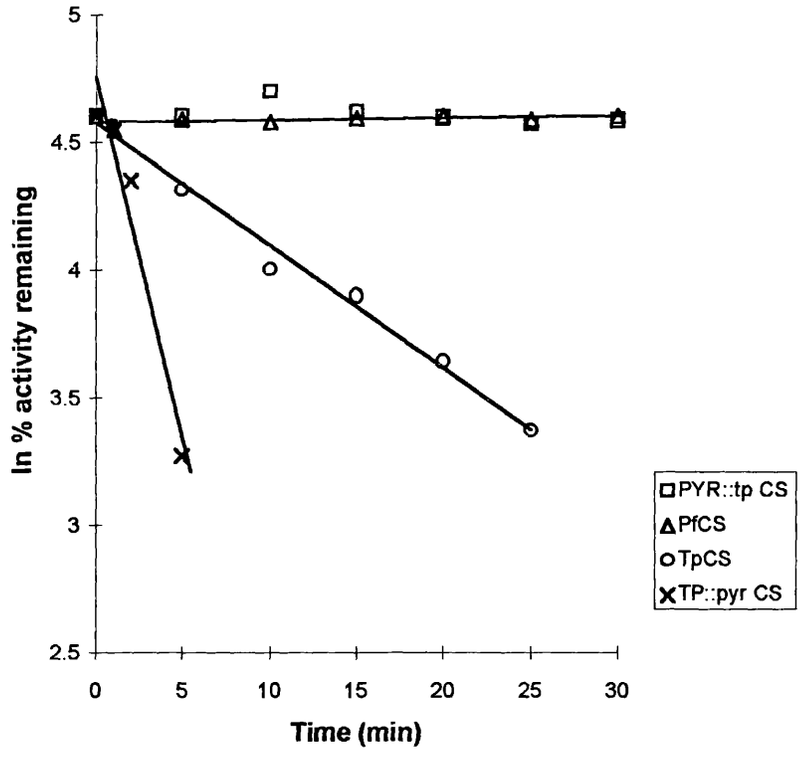


Figure 3.19. Relative thermostability of *Pf*CS, *Tp*CS, and the domain swapped mutants at 85°C. The slopes are drawn by linear regression, (the linear regression line for PYR::tp CS has been omitted for clarity).

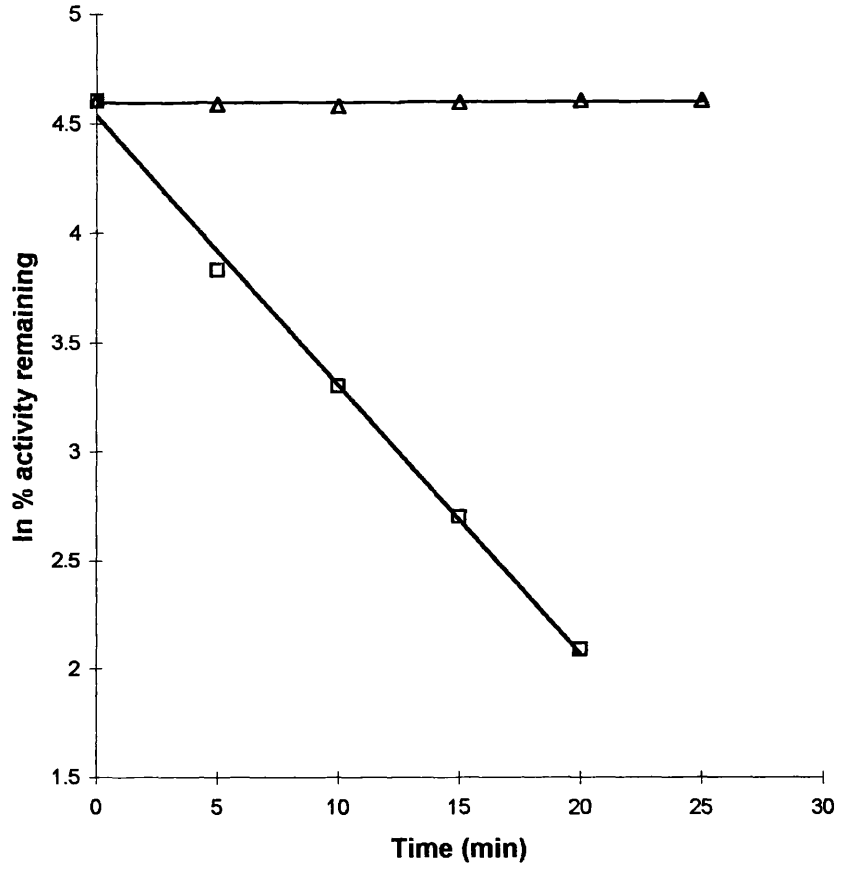


Figure 3.20 Relative thermostability of PYR::tp CS to *Pf*CS at 95°C. The lines are drawn by linear regression.

3.8 Discussion

3.8.1 Use of pAlter-1 kit in site directed mutagenesis

Many problems were encountered when using the pAlter-1 kit for site directed mutagenesis to create Apa1 and BspE1 sites in *Pf*CS and *Tp*CS, respectively. Firstly, it was found to be very difficult to recognise pAlter-1 transformants because often white colonies tended to be pAlter-1 vector only clones. This meant that a very large number of clones had to be screened before pAlter-1 clones containing a citrate synthase gene were identified. Secondly, it proved very difficult to make the *E.coli* repair-minus cells supplied with the kit (strain ES1301) competent, which meant that another repair-minus strain, DH5α, had to be used. Finally, because of the high G:C content of the DNA coding for the CS's, no mutants could be made starting with double stranded DNA.

The difficulties encountered making mutants using the pAlter-1 kit contrasted with the ease with which the Nde1-Kpn1 PYR::tp, and TP::pyr CS mutants were made using a PCR approach. For this reason in future a PCR mutagenesis method would be used in preference to the p-Alter method.

3.8.2. Kinetic studies of the mutants

It was fortunate that both domain swapped mutants within the pREC7-Nde1 vector (pBAP3008 and pBAP3010) expressed high levels of soluble proteins. It can be seen from Table 3.2 (showing a comparison of the specific activity and K_m values for the 4 CS's) that both domain-swapped mutants were active. This, coupled with the fact that the affinities for both substrates have not been greatly altered in the mutants, indicates that the mutant enzymes are

adopting a similar conformation to the wild types. However, the fact that the TP::pyr CS enzyme is substrate inhibited with Ac-CoA indicates that the folding of this enzyme differs subtly from that of the wildtype *Tp*CS and *Pf*CS. This substrate inhibition effect is a direct consequence of the changes caused by the incorporation of the *Pf*CS small domain into the *Tp*CS large domain. This Ac-CoA substrate inhibition effect is discussed further in chapter 6.

There appears to be a correlation between the K_m value for OAA and the large domain partner in *Pf* CS and *PYR::tp* CS, and in *Tp* CS and *TP::pyr* CS. This may be because two of the three OAA binding residues are located within the large domain of *Pf* CS and *Tp* CS.

From the specific activity values shown in table 3.2 it can be seen that the specific activity of the chimaeric mutant is related to the small domain partner at 55°C. This is shown by the fact that *Pf*CS and TP::pyr CS have a similar specific activity at 55°C, and the specific activity of the PYR::tp CS mutant is approximately half of that of the *Tp*CS at 55°C but nearly twice that of *Pf*CS. This finding is consistent with the idea that the small domain determines the catalytic activity of the CS enzyme. This conclusion is perhaps not so surprising, because when one looks at the positions of the catalytic residues that have been predicted to play a central role in the mechanism of action of *Pf*CS and *Tp*CS (from the crystal structure), two of the three are located within the small domain, while the third lies in the hinge region between the large and small domains (Russell et al., 1994, 1997). This conclusion supports the findings of Molgat et al. (1992) who swapped the domains between *E.coli* CS and *A.anitratum* CS, and concluded from these mutants that the small domains of the CS enzymes are involved in catalytic activity.

3.8.3 Thermal inactivation of mutant citrate synthases

From Fig 3.19 it can be seen that in terms of thermal stability at 85°C, *Pf*CS, and PYR::tp CS>*Tp*CS>TP::pyr CS. There is no observable difference in the thermal stability of *Pf*CS and PYR::tp CS at 85°C. However at 85°C *Tp*CS and TP::pyr CS show significant rates of denaturation (*Tp*CS is nearly six times more thermostable than TP::pyr CS at this temperature). Thus by incorporating the region coding for the small domain of *Pf*CS into the region coding for the large domain of *Tp*CS, the resulting chimaeric CS is much more unstable than either of the wildtype CS's. One possible reason for this is that the *Pf*CS small domain may lack residues which help stabilise the *Tp*CS (or change their spatial position within the molecule) thus disrupting the packing of *Tp*CS and creating strain within the CS molecule, or it might form unfavourable interactions with the *Tp*CS large domain. Perhaps one should not be surprised that one of the chimaeric mutants is unstable as each small domain region is being coupled with a large domain, which shows only 39% sequence identity at the amino acid level, to its wild type large domain partner. Other groups have found the thermal stability of chimaeric enzymes greatly reduced. Lebbink et al., (1995) found that one of their domain swap glutamate dehydrogenase mutants was less stable than either of the wild types from which the mutants were constructed. This led them to conclude that the presence of a domain from a hyperthermostable enzyme may not significantly enhance the temperature at which an enzyme becomes susceptible towards thermal inactivation. While this is true in the case of the TP::pyr CS, in PYR::tp CS the presence of the large domain from the hyperthermophilic *Pf*CS significantly enhances the thermostability of the *Tp*CS small domain with respect to *Tp*CS.

The fact that PYR::tp CS shows greater thermostability than *Tp*CS at 85°C indicates that the *Pf*CS large domain is stabilising the *Tp*CS small domain. This shows that the large domain of *Pf*CS has a greater thermostabilising effect than the *Tp*CS large domain. From this one could draw a conclusion that the structural features which make *Pf*CS more thermostable than *Tp*CS reside in the large domain, and that the major determinants of thermostability lie within the large domain. It is difficult to quantify how much more thermostable PYR::tp is compared with *Tp*CS, because at 95°C when PYR::tp CS is showing a significant rate of denaturation, the activity of *Tp*CS is destroyed in under 1 minute. Similarly it is difficult to quantify just how much more thermostable *Pf*CS is than PYR::tp CS, since at 95°C *Pf*CS does not show a significant rate of denaturation (Fig 3.20), while at 103°C when *Pf*CS shows a significant rate of denaturation (see Fig 5.16) all PYR::tp CS activity is destroyed in under 1 minute. These results confirm the findings of Molgat et al., (1992), that it is the large domain region which plays the main stabilising role.

While these data point to the importance of the small domains to the catalytic activity of the CS's under investigation, and the importance of the large domains to thermostability, it is more difficult to draw a conclusion about the importance of the small domains to thermostability, since in both cases changing the nature of the small domain reduces the thermostability properties of the large domain partner.

3.8.4 Concluding Comments

We have created soluble, active chimaeric enzymes which can be used to investigate the contributions of the large and small domains to thermostability. Given the structural homology between the psychrophilic DS23r CS and *Pf*CS it

would have been interesting to have made chimaeras with these two enzymes. However the structure of DS23r came too late in my project too do this. It was therefore decided to investigate the conclusion that the structural determinants of hyperthermostability in CS lie within the large domain, by making more citrate synthase mutants that would possess disrupted structural elements within the large domain of *Pf*CS, which were predicted to play a role in thermostability from the crystal structure of *Pf*CS.

CHAPTER 4

Production of *Pf*CS C-Terminal Mutants

4.1 Introduction

This chapter describes the creation and characterisation of two C-terminal mutants of *Pf*CS, a -2 C-terminal amino acid deletion mutant (Pf-2), and a -13 C-terminal amino acid deletion mutant (Pf-13).

The C-terminal region of *Pf*CS is located within the large domain. It was decided to investigate the importance of the C-terminal region of *Pf*CS to thermostability because the crystal structure of the closed form of *Pf*CS (Russell et al., 1997) has revealed that the C-terminal arm of one monomer appears to be wrapped around the N-terminus of the other monomer in an 'intimate embrace'. It was predicted that this interaction might be playing a thermostabilising role by forming a subunit interaction between the two monomers at the C-terminus of the protein. This C-terminal interaction has been observed in the closed form of *Pf*CS but is not observed in low resolution studies of the open form of *Pf*CS (R. J.M. Russell, personal communication) nor in the open form of *Tp*CS. This has led to the idea that the C-terminal arm only becomes ordered upon the binding of the substrates oxaloacetate and Ac-CoA. Fig 4.1 shows the structure of the closed form of *Pf*CS with each monomer shaded differently; the C-terminal arm of one monomer can be seen to be wrapped around the N-terminus of the other monomer.

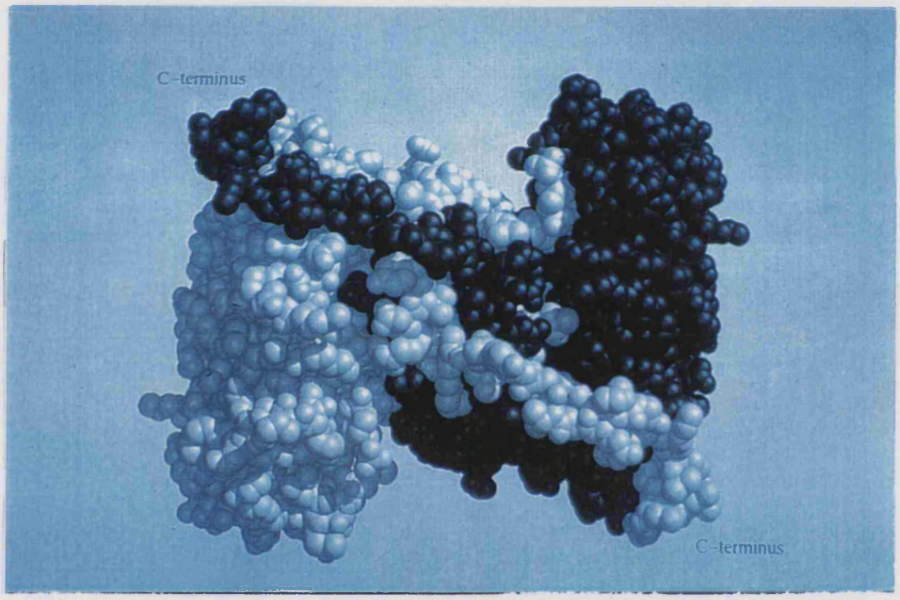


Figure 4.1 Space filling representation of the *Pf*CS dimer, showing the C-terminal residues of each monomer wrapping around the other monomer.

One can see by looking at an amino acid alignment of the C-terminal region of citrate synthases, from a number of organisms, that this region varies between different organisms (Fig 4.2). From Fig 4.2 it can be seen that the CS from the psychrophilic organism DS23r, *B.subtilis* R, and from pig are shorter at the C-terminus than *Pf*CS, and are relatively thermolabile compared with *Pf*CS, DS23r and pig CS being rapidly denatured at temperatures of 45°C and 60°C respectively.. One might suppose that the truncation of the C-terminal region relative to *Pf*CS may be one reason for the relative thermolability of the CS's from DS23r, *B.subtilis*, and pig, compared with *Pf*CS. However it is very difficult to draw conclusions from amino acid alignments alone, because although pig CS and *Pf*CS are slightly different in length at the C-terminus, the C-terminal region of *Pf*CS is in an extended conformation with respect to pig CS. On the other hand DS23r CS with a 6 residue C-terminal deletion (relative to *Pf*CS), forms a similar C-terminal interaction to *Pf*CS, although the loss of the 6 residues reduces the extent of the subunit interface (Russell et al., 1998). From Fig 4.2 it can also be seen that only the thermophilic CS's have two positively charged amino acid residues at the C-terminus. Glutamate residue 48 to which Arg375 forms an ionic interaction in *Pf*CS is conserved in the other CS's shown.

		Number of end residue
DS23R	RPLSEYNGPEQRQVP	379
E.coli	RPRQLYIGYEKRDFKSDIK	426
Pig	RPKSMSTDGLIKLVDSK	437
Bacillus subtilis R	RPSAQYTGAIPEEVLS	365
Thermoplasma	RPRAVYVGPAERKYVPIAERRK	384
Sulfolobus	RPRALYVGPEYQEYVSIDKR	377
Pyrococcus	RPRLQYVG*EIGKKYLPIELRR*	376

Figure 4.2 Amino acid alignment of CS C-terminal regions from Archaea, a Eukaryote, Gram negative and Gram positive bacteria.

To investigate the importance of the C-terminal region to the thermostability of *Pf*CS it was decided to produce mutants of the enzyme where the C-terminal arm was deleted. The first C-terminal mutant constructed involved the removal of the 2 terminal arginine residues. This was to determine the importance of the ionic bond between the penultimate arginine residue 375 and glutamate residue (48). This ionic bond is thought to anchor the whole C-terminal arm.

The second C-terminal mutant to be constructed involved removing the C-terminal arm from where it emerges from the inside of the *Pf*CS molecule. This is thought to be after valine 362 residue which aligns with valine 370 from *Tp*CS which is the last residue to be clearly defined from the open structure of *Tp*CS. Rather than put stop codons after the valine 362 residue it was decided to put the stop codons after glycine 363 as one did not want to disrupt the QYVG hydrophobic sequence. The amino acid residues removed by this mutation are shown in Fig4.2 in italic.

Several groups have investigated the importance of the C-terminal regions to thermal stability. Chen et al. (1995) investigated the importance of the C-terminal region of glucoamylase in terms of its thermostability and ability to bind starch by the deletion of 8 and 25 C-terminal residues. They found that the thermostability of the enzyme was unaffected but the enzyme progressively lost the ability to bind and hydrolyse starch. Meinnel et al. (1996) progressively deleted the C-terminal residues of *E.coli* peptide deformylase and found that the last 21 amino acid residues may be omitted without loss of activity, and that the truncated form was in fact more thermostable than the wild type.

4.2 Creation of the Pf-2 C-terminal deleted *Pf*CS mutant

A decision was made to use PCR to create this mutant. The 5' sense primer used was MAA-7, which bound to and amplified from the 5' end of the *Pf*CS gene (see section 3.5.1) adding an Nde1 site. The antisense primer used was MAA-15 (see Fig 4.3) which added a stop codon after the DNA sequence coding for the third amino acid residue (residue 374) from the end of the C-terminal arm, and contained a Kpn1 site to allow cloning of the CS into the pRec7-Nde1 vector.

5'TTTTTT*GGTACC***CTA**TAATTCTATGGGTAG 3'

Figure 4.3 MAA-15 Primer. The bases shown in bold represent the stop codon, the bases shown in italics represent the Kpn1 site, the underlined bases code for the amino acids preceding the 2 penultimate amino acid residues.

0.5μg template DNA [pBAP3031(*Pf*CS in pRec7Nde1)], 5mM $MgCl_2$, and 18 pmoles of each primer were used in each reaction; for the other PCR conditions see section 2.2.11. Several identical PCR reactions were set up in order to obtain large quantities of the PCR product. Samples of two of the PCR reactions were run on a 1% agarose minigel, along with pRec7-Nde1 digested singly with Nde1 and Kpn1 (Fig 4.4). From this gel it can be seen that the PCR product corresponded to 1.3 kb in size.

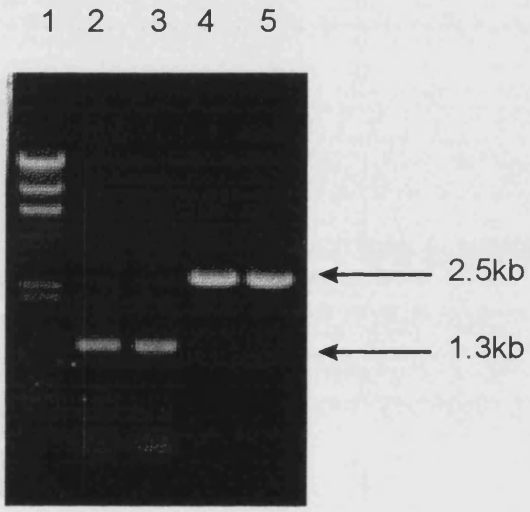


Figure 4.4 PCR products and Nde1 and Kpn1 single digests of pRec7-Nde1 vector. 1) kilobase ladder DNA. 2) -2 PCR product. 3)-2 PCR product. 4) pRec7-Nde1 digested with Nde1 enzyme. 5) pRec7-Nde1 digested with Kpn1 enzyme.

The PCR product was purified using the Gene Clean method before being double digested with Nde1, and Kpn1. The Nde1-Kpn1 digested PCR product was again purified using the Gene Clean method and an overnight ligation was set up at 15°C between digested vector and insert (see sections 2.2.7, and 2.2.8). The following day the ligation reactions were transformed into the CS-minus *E.coli* strain MOB154 and plated out onto LA plates containing ampicillin at a concentration of 125μg/ml. No transformants appeared on the

vector only control so several colonies were picked from the vector +insert plates and were restreaked and 'miniprep' DNA prepared from them. This DNA was digested with Nde1 and the digests run on a 1% agarose gel; bands of 3.8 Kb could be seen (gel not shown) which corresponded to the *Pf*CS gene in the pRec7-Nde1 vector. A recombinant was sequenced using primer MAA-10 [which binds downstream from the Kpn1 site of pRec7-Nde1 (see Appendix A) allowing the antisense strand of any gene cloned into the pRec7-Nde1 vector to be read]. Sequencing with the MAA-10 primer identified a Pf-2 mutant. This 2 C-terminally deleted *Pf*CS clone within the vector pRec7-Nde1 was designated pBAP3011. This clone was sequenced completely using primers MAA-9 and MAA-11 (see Appendix A) to determine whether any extra mutations had been introduced by the PCR procedure. No further mutations were found.

4.3 Creation of the Pf-13 C-terminal deleted *Pf*CS mutant.

A similar strategy was used to produce the Pf-13 mutant as was used to create the Pf-2 mutant, the only difference being that a different C-terminal antisense primer, MAA-16, was used. This primer incorporated two stop codons upstream of the bases coding for the terminal 13 amino acids of *Pf*CS (Fig 4.5).

5' TTTTTT*GGTACC***CTACTA**TCCTACATACTGCAACCTGGG3'

Figure 4.5 MAA-16 Primer. The Kpn1 site is shown in italics, the two stop codons are shown in bold, and the underlined bases code for the amino acids upstream of the 13 C-terminal amino acids.

0.5μg template DNA [pBAP3031(*Pf*CS in pRec7Nde1)], 5mM $MgCl_2$, and 18 pmoles of each primer were used in each reaction; for the other PCR conditions see section 2.2.11. The PCR products were purified using the Gene

Clean method before being double digested with Nde1 and Kpn1. The Nde1-Kpn1 digested PCR product was again purified using the Gene Clean method and an overnight ligation was set up at 15°C between digested vector and insert. The following day the ligation reactions were transformed into the CS-minus *E.coli* strain MOB154 and plated out onto LA plates containing ampicillin at a concentration of 125μg/ml (see sections 2.2.6, and 2.2.7). No colonies were obtained on the vector only control plates so several colonies were picked from the vector + insert plates and 'miniprep' DNA was prepared from these, which was then digested with Nde1 enzyme. A 1% (w/v) agarose minigel was run of these digests (Fig 4.6). From this gel it can be seen that most of the recombinants gave a 3.8 Kb band with Nde1 enzyme, showing that the -13 *Pf*CS PCR product had been cloned into pRec7-Nde1. One of these clones was sequenced initially with primer MAA-10 [which binds downstream from the Kpn1 site of pRec7-Nde1 (see Appendix A) allowing the antisense strand of any gene cloned into the pRec7-Nde1 vector to be read]. The 13 C-terminally deleted *Pf*CS clone identified by this sequencing was designated pBAP3013. The pBAP3013 construct was sequenced completely using primers MAA-9 and MAA-11 (see appendix A) to determine whether any extra mutations had been introduced by the PCR procedure. No further mutations were found.

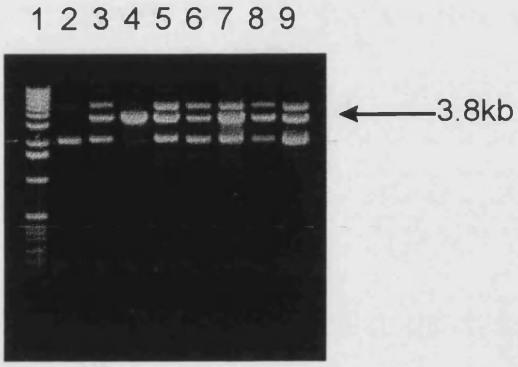


Figure 4.6 Nde1 digest of vector + insert recombinants from the cloning of the -13 *Pf*CS PCR product into pRec7-Nde1 vector. 1) kilobase ladder DNA. 2) Uncut miniprep 1 DNA. 3-9) Nde1 digested miniprep DNA samples.

Now that both C-terminal mutants had been confirmed by sequencing, the next step was to over express them and obtain pure protein for the purpose of characterisation.

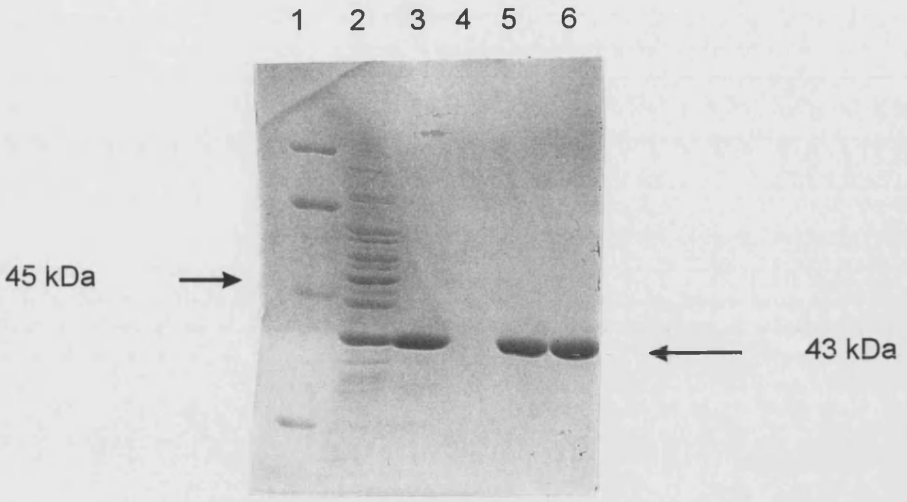
4.5 Purification of the two *Pf*CS C-terminal mutants

Cells were grown, induced, and cell extracts prepared (see sections 2.3.2, and 2.3.3). Experiments were done to determine a suitable heat step for removing most of the *E.coli* host proteins from the cell extract without inactivating the mutant CS's. A heat step of 80°C for 15 min was found to be adequate. Hence the 2 mutants were purified in the same way as the wild type *Pf*CS. Table 4.1 shows the yields and specific activities obtained from each step of the purification, starting with 0.2g of cells. Fig 4.7 shows SDS-PAGE gels of samples taken during the purification of the Pf-2 and Pf-13 mutants. All the specific activity figures for the Pf-13 mutant are low because these assays

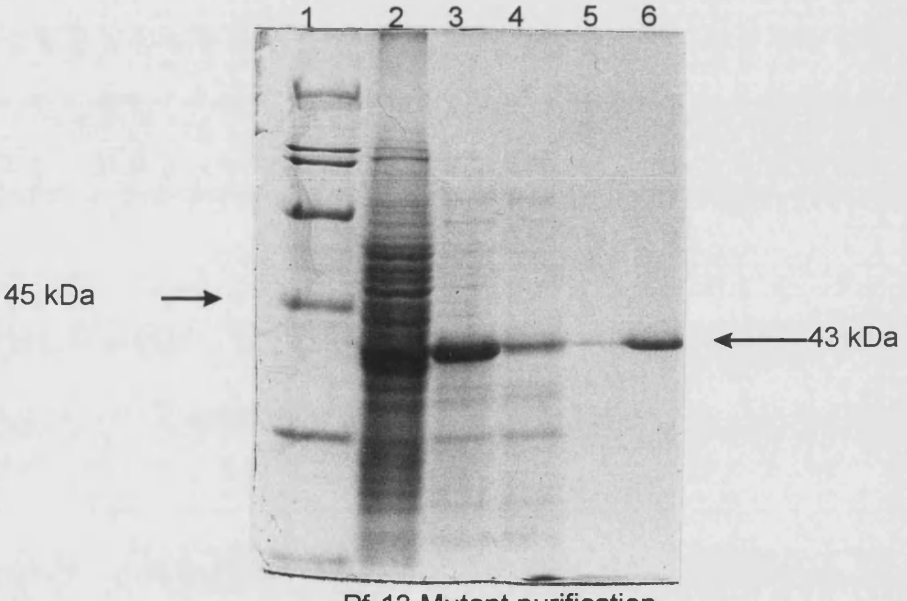
were carried out using standard oxaloacetate and Ac-CoA concentrations (see section 2.3.6) which turned out to be approximately equal to the K_m values for this enzyme, rather than 10 x K_m as for *Pf*CS and the Pf-2 mutant. All the specific activity figures for the Pf-2 mutant are low because it is substrate inhibited by Ac-CoA

CS Construct/ step	Volume (ml)	Activity /ml (Units)	Total Activity (Units)	[Protein] mg/ml	Total Protein (mg)	% Yield	Specific Activity Units/mg
-2 Mutant							
Cell Extract	1	11.8	11.8	9.35	9.35	100	1.3*
Heat Step	0.9	11	9.9	7.75	7	84	1.4*
Red gel A	2.5	3	7.4	0.3	0.7	62	10.7*
-13 Mutant							
Cell Extract	1	1.8	1.8	5.4	5.4	100	0.3
Heat step	0.7	2.3	1.6	2.5	1.85	87	0.8
Red gel A	3.6	0.25	0.9	0.16	0.39	50	2.3

Table 4.1 Purification of Pf-2 and Pf-13 *Pf*CS mutants. All assays were carried out at 55°C. * Indicates that the specific activity values are low because of substrate inhibition by Ac-CoA.



Pf-2 purification



Pf-13 Mutant purification

Figure 4.7 SDS-PAGE analysis of fractions from the purification of the Pf-2 and the Pf-13 mutants. 1)Low range molecular weight markers. 2) Cell extract. 3) Heat step. 4) Red gel A Unbound fraction. 5-6) Red gel A eluted fractions.

4.6 Characterisation of the C-terminally deleted mutants.

4.6.1 Determination of the kinetic perameters

The K_m and V_{max} values for both C-terminal mutants were determined in the same way as the domain swapped mutants (see section 3.7.1.). All measurements were repeated three times at 55°C, and the mean K_m and specific activity values are shown in Table 4.2. The K_m and V_{max} values for wild type *Pf*CS were determined at the same time for comparative purposes. The direct linear plot of Eisenthal and Cornish-Bowden (1974) was used to determine the K_m and V_{max} values, except in the case of the Pf-2 mutant which showed substrate inhibition by Ac-CoA. Here the K_m and V_{max} values were determined by using the Scientist plot to fit the rates to a curve given by the equation

$$v = \frac{Vmax}{1+\frac{Km}{S} + \frac{S}{Ki}}$$

Citrate Synthase	Km OAA (μM)	Km Ac-CoA (μM)	Specific Activity (μmol/min/mg)
P.furiosus	10 ±1.3	3.0 ± 0.7	17 ±2
Pf-2	16 ±1	5.0 ± 0.8 *	35 ± 3
Pf-13	125 ±16	96 ± 8	8 ±1

Table 4.2 K_m and V_{max} values for *Pf*CS and the C-terminally deleted *Pf*CS mutants. *Indicates enzyme substrate inhibited by Ac-CoA. The mean K_i value for Ac-CoA for the Pf-2 mutant was 100 (± 4) μM.

The V against S plots of the data used to determine the K_m and V_{max} values for the Pf-2 and Pf-13 CS mutants are shown in Fig 4.8a-d. These graphs were plotted in the same way as those in section 3.7.1 using the Microcal Origin graphics package. Note the K_m value of the Pf-2 mutant for OAA was determined at an Ac-CoA concentration of 12μM.

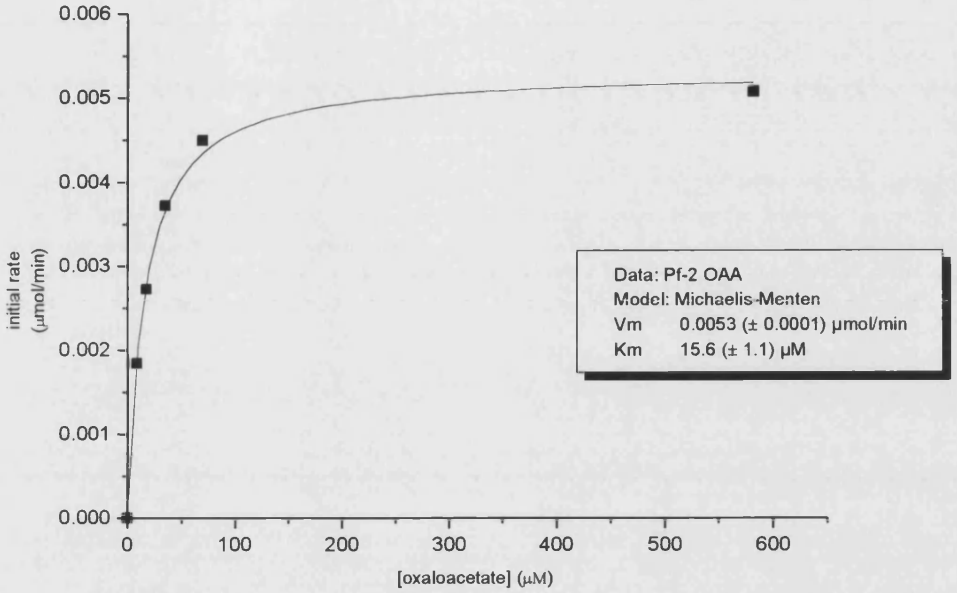


Figure 4.8a V against S plot of data used to determine the K_m and V_{max} values for the Pf-2 mutant for the substrate oxaloacetate.

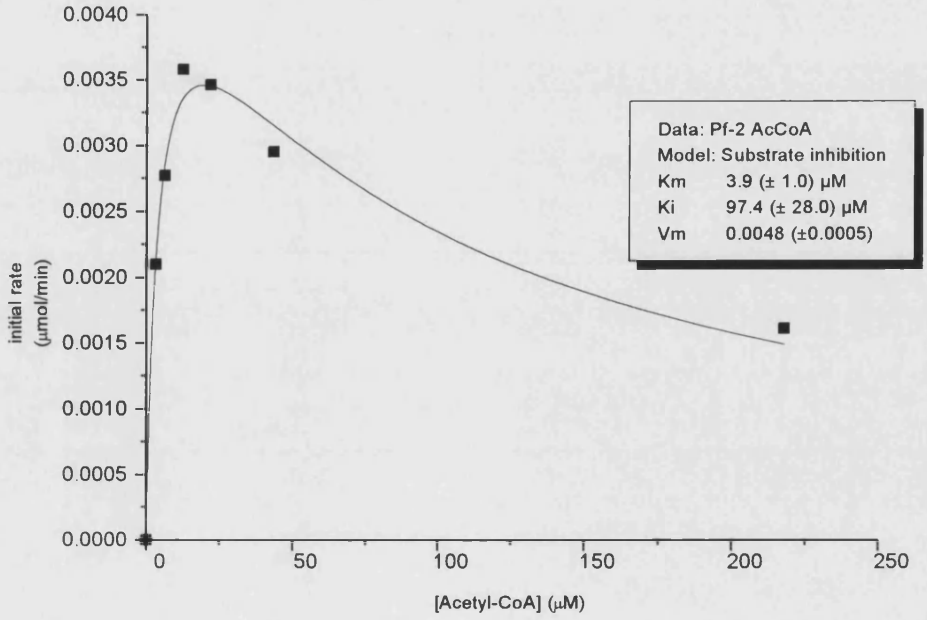


Figure 4.8b V against S plot for the substrate acetyl-CoA of data used to calculate the K_m and V_{max} values of the Pf-2 mutant.

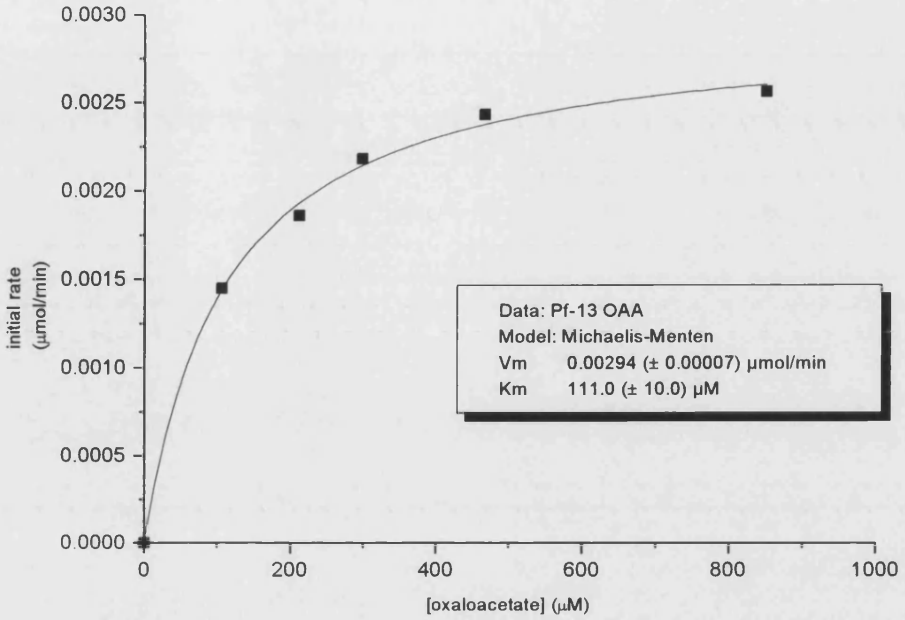


Figure 4.8c V against S plot of data used to determine the K_m and V_{max} values for the Pf-13 mutant for the substrate oxaloacetate.

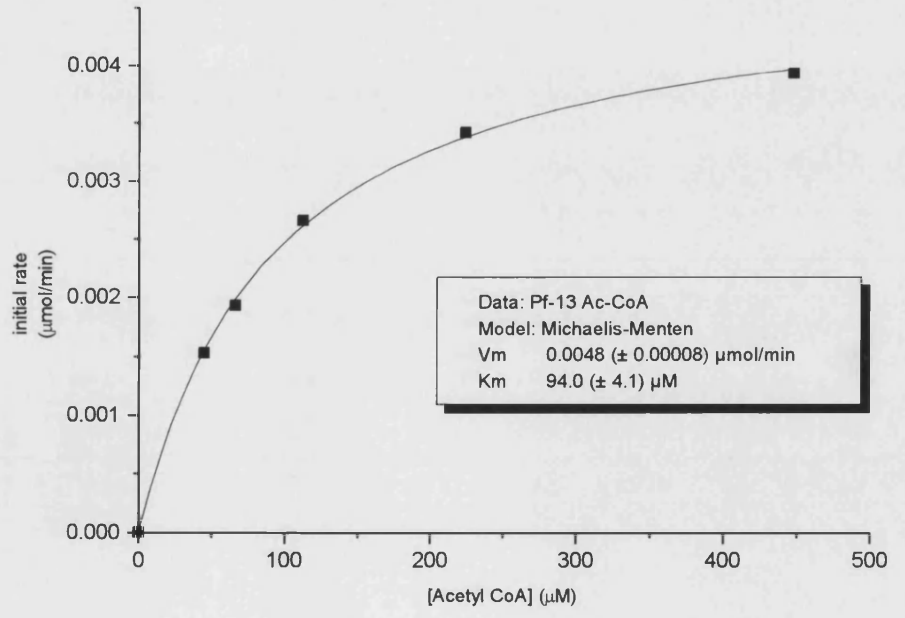


Figure 4.8d V against S plot of data used to calculate the K_m and V_{max} values of the Pf-13 mutant for the substrate Ac-CoA.

4.6.2 Thermal stability of the C-terminally deleted mutated CS's

The aim of these experiments was to compare the C-terminally deleted mutant CS enzymes with *Pf*CS in terms of their abilities to withstand thermal inactivation.

All thermal stability studies were carried out at 103°C in the presence of 50mM phosphate buffer, pH 7.2, containing 2mM EDTA, using an enzyme concentration of 0.1mg/ml. The buffer was heated to the required temperature in an Eppendorf tube in a bath containing polyethylene glycol grade 400 (PEG bath), and the temperature was measured using an electronic temperature probe inside a second Eppendorf tube containing PEG grade 400 solution. When the desired buffer temperature was reached, enzyme was added to the Eppendorf tube, mixed with the buffer and the timer started. Aliquots were removed at timed intervals and placed into thin walled glass tubes on ice, where they were stored prior to assay. A time zero value was obtained by assaying a 0.1mg/ml concentration of enzyme in the same buffer kept at 4°C. Samples were assayed in triplicate, and the results are presented as a plot of ln (mean % activity remaining) against time. Fig 4.9 shows the relative thermal stability of the 2 C-terminally deleted mutants and *Pf*CS at 103°C. From this graph it can be seen that in terms of thermostability *Pf*CS >Pf-13 mutant >Pf-2 mutant. The rate constants for denaturation, and the times required to inactivate 50% of the enzyme's activity at 103°C are shown in Table 4.3.

Citrate Synthase construct	Rate constant for inactivation (min^{-1})	Time required to inactivate 50% of enzyme (min)
***Pf*CS**	0.10 ± 0.006	8.5
Pf-13	0.17 ± 0.01	4.7
Pf-2	0.25 ± 0.02	2.6

Table 4.3 Rate constants for inactivation of *Pf*CS, and the C-terminal mutants, and times at which 50% of the enzyme's activity remains, at 103°C.

From the rate constants for inactivation at 103°C, the change in the free energy of activation of thermal unfolding for both mutants compared to the wildtype *Pf*CS was determined (using the method given in section 2.4). It was found that there was a decrease in the free energy of activation for thermal inactivation of about 1.6 and 2.8 kJ/mol at 103°C for the Pf-13 and Pf-2 mutant respectively.

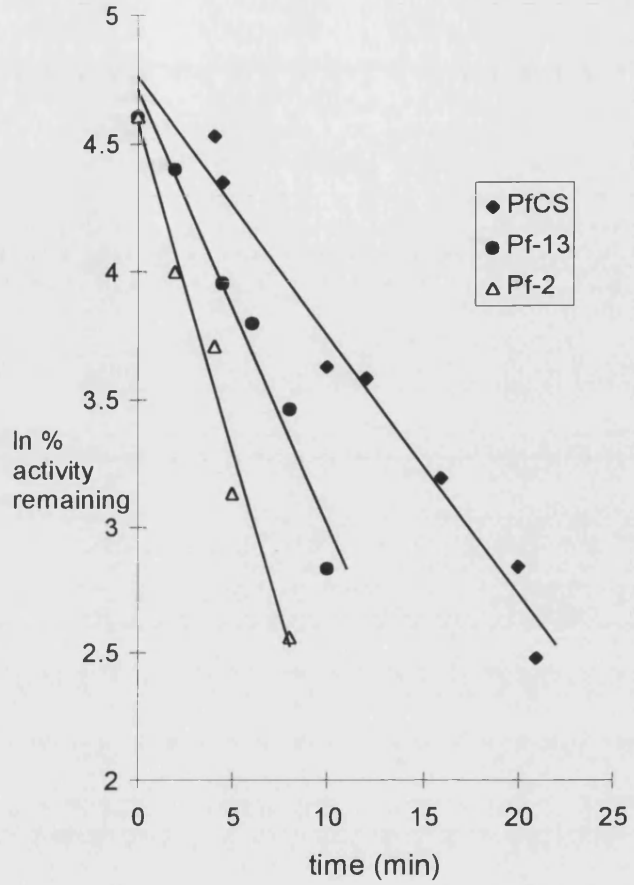


Figure 4.9 Relative thermostability of *Pf*CS and the C-terminally deleted mutants at 103°C.

4.7 Discussion

4.7.1 Kinetic Studies of the mutants

It can be seen from Table 4.2 (showing a comparison of the specific activity and K_m values for the 2 mutants and *Pf*CS) that both C-terminal *Pf*CS mutants were active. In the case of the Pf-2 mutant, where the K_m values for the substrates oxaloacetate and Ac-CoA were relatively unchanged relative to *Pf*CS, the enzyme can be assumed to be folding in a similar way to *Pf*CS. However there must be a subtle difference in the enzyme's folding, as this enzyme is substrate inhibited with Ac-CoA, a phenomenon also observed with the TP::pyr CS (see section 3.8.2). This substrate inhibition is a consequence of removing the 2 terminal arginine residues. One could argue that this is occurring either as a result of the destruction of the ionic interaction between Glu 48 and Arg 375, or because of the physical shortening of this C-terminal arm. One way to investigate whether or not the destruction of the ionic interaction was the cause of this effect would be to mutate Glu 48 or Arg 375 to an amino acid with an uncharged side chain such as serine thus destroying the ionic interaction but leaving the C-terminus of the protein unchanged in terms of its length.

The fact that the Pf-2 mutant has a specific activity at 55°C of double that of *Pf*CS indicates that this mutant is more efficient at producing citrate than the wildtype *Pf*CS. Again this could be a consequence of the destruction of the Arg 375-Glu 48 ionic bond, or as a result of the enzyme being shortened by 2 amino acid residues. Perhaps the destruction of the Arg 375-Glu 48 bond changes the dynamic properties of the enzyme so that the enzyme becomes

less rigid and more flexible, enabling it to convert substrate into product more rapidly at 55°C.

In the case of the Pf-13 mutant, the K_m values are changed significantly as a result of the deletion, which may indicate that the enzyme is adopting a subtly different folded conformation at the C-terminal subunit interface (compared to *Pf*CS) which affects the binding of the substrates and the specific activity of the mutant. It is interesting that the Pf-13 mutant is not substrate inhibited by Ac-CoA, in contrast to the Pf-2 mutant. Perhaps once the 2 terminal amino acids are removed from *Pf*CS, the remainder of the C-terminal arm contributes to the Ac-CoA substrate inhibition conformation, so that when this region is removed the *Pf*CS is no longer Ac-CoA inhibited.

4.7.2 Thermal inactivation of the C-terminal mutants

From Fig 4.9 it can be seen that in terms of thermostability *Pf*CS> Pf-13 >Pf-2 mutant. The Pf-2 mutant had a half life at 103°C of almost half that of the Pf-13 mutant, and a third of the half life of *Pf*CS. This indicates that the resistance towards thermal unfolding is strongly diminished both by the removal of the 2 C-terminal and 13 C-terminal amino acid residues. This suggests that both the ionic interaction between Glu 48 and Arg 375, and the C-terminal arm as a whole, play a thermostabilising role in *Pf*CS. The decrease in free energy of activation of thermal inactivation calculated for both mutants is similar to that calculated by Pappenberger et al. (1997). They found that the free energy of activation of thermal unfolding of glyceraldehyde-3-phosphate dehydrogenase was decreased by 4kJ/mol when the central residue of an ionic network was replaced by an uncharged alanine residue.

One question that these results raise is why the more C-terminally deleted Pf-13 CS should be more thermostable than the Pf-2 CS, because the Pf-13 mutant will also have had its Glu 48-Arg 375 ionic interaction destroyed. Meinnel et al. (1996) have reported similar findings from their work on C-terminal mutants constructed from peptide deformylase. They found that when the 21 and 23 terminal amino acids were removed from the peptide deformylase, the truncated mutant was more stable than the wildtype enzyme. Their explanation for the fact that the truncated form of the enzyme was more stable than the wildtype, was that in the wildtype enzyme the terminal 3 amino acids are disordered in the structure thereby causing relative instability to the catalytic core. Thus, when they are removed the enzyme is more stable. Therefore one explanation for the Pf-2 mutant being less thermostable than the Pf-13 mutant is that on loss of the ionic interaction at the C-terminus in the Pf-2 mutant the C-terminal arm becomes destabilising, and this destabilising effect is reduced when the extra 11 amino acid residues are removed in the Pf-13 mutant.

Hamana & Shinozawa (1999) also observed a decrease in thermostability in the C-terminal mutants that they constructed of orotate phosphoribosyl transferase (OPRTase) from *Thermus thermophilus*. In these mutants the terminal 1, 2, 3, and 5 amino acid residues were removed; all the mutants showed a significant decrease in thermostability compared to the wildtype enzyme, which led them to suggest that the terminal five amino acids at the C-terminus contribute to the stabilisation of the enzyme. All the mutants showed lower specific activities compared with the wildtype, and the K_m value for one substrate was found to have increased 4 fold in all the mutants while the K_m value for the other substrate was unchanged compared to the wild type enzyme. These results suggested that three of the deleted C-terminal residues

were important for orotate binding to OPRTase and the deletion of a few amino acids from the C-terminus caused structural changes to the active site. These results are similar to our findings in that removal of C-terminal residues from *Pf*CS caused changes to the enzyme's thermostability and catalytic activity.

4.7.3 Concluding Comments

These results show that the whole C-terminal arm, from where it leaves the core of the *Pf*CS molecule, is dispensable for activity, in that the enzyme still retains activity without it. However they show that the whole C-terminal arm, and in particular the 2 terminal amino acid residues (and the ionic interaction between Glu 48 and Arg 375), play a role in the thermostability of *Pf*CS, as predicted from crystal structural comparisons.

CHAPTER 5

Production of *Pf*CS subunit interface ionic network mutants

5.1 Introduction

In *Pf*CS the monomer-monomer association is made up of two parts: the interaction of the C-terminus of one monomer with the N-terminus of the other monomer (as described in chapter4), and interactions centred on an eight α-helical sandwich, comprising four antiparallel pairs of helices (F, G, L, and M). For the purposes of this chapter the term subunit interface refers to the interactions between the 8 α-helical sandwich and not the C and N terminal interactions.

In comparison with CS's isolated from organisms that live at lower temperatures than *P.furiosus, Pf*CS has the greatest number of intersubunit ionic interactions. In addition there are a greater number of ionic interactions within the 8 α-helical sandwich region at the subunit interface of *Pf*CS than found in any other CS. *Pf*CS possesses two five-membered ionic networks in this region (Russell et al., 1997). The five-membered ionic *Pf*CS ionic network is shown in Fig 5.1. Fig 5.2 shows a diagrammatic representation of the ionic ring network present in *Pf*CS at each end of the subunit interface region. Fig 5.3 shows the ionic interactions present at the 8 α-helical sandwich interface region of CS (calculated at a cut-off distance of 4Å) for pig, DS23r, *T.acidophilum*, *S solfataricus*, and *P.furiosus.* It can be seen that pig CS has none, *Tp*CS has a single interaction, *S.solfataricus* CS has two which are not networked, *Pf*CS has

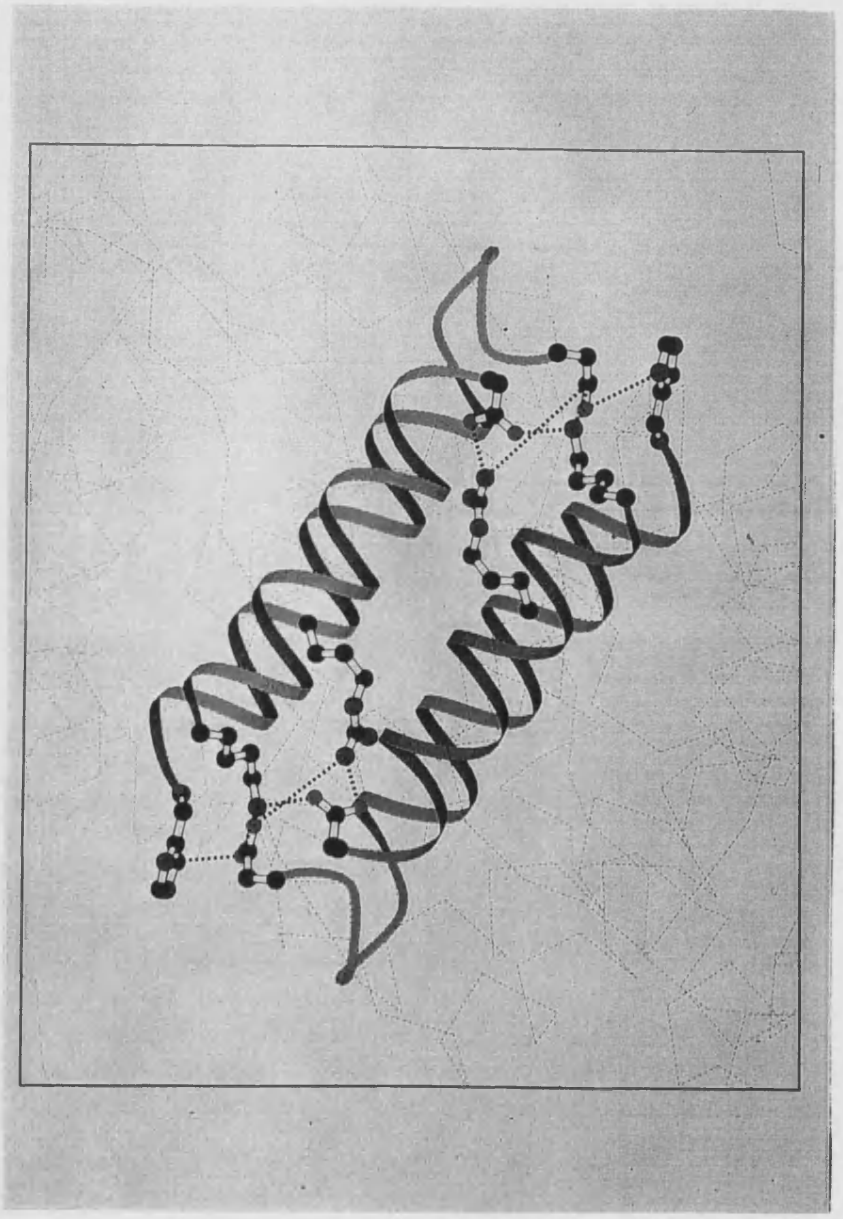


Figure 5.1 Five membered intersubunit ion pair networks in *Pf*CS, viewed down the two fold axis of the dimer. Equivalent helices (G -brown, and M -green) from each monomer are coloured similarly.

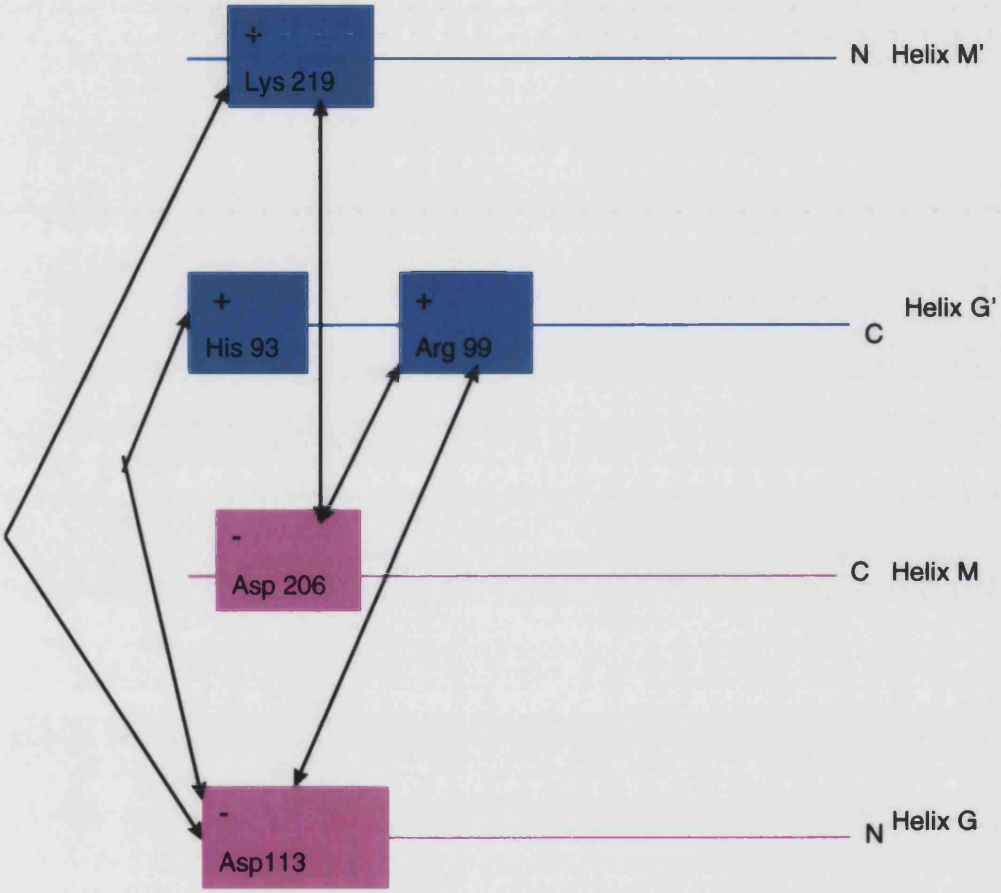


Figure 5.2 Representation of the 5 membered ionic network at the subunit interface of *Pf*CS, one monomer is shaded in purple while the other is shaded in light blue. The ionic network is not drawn to scale.

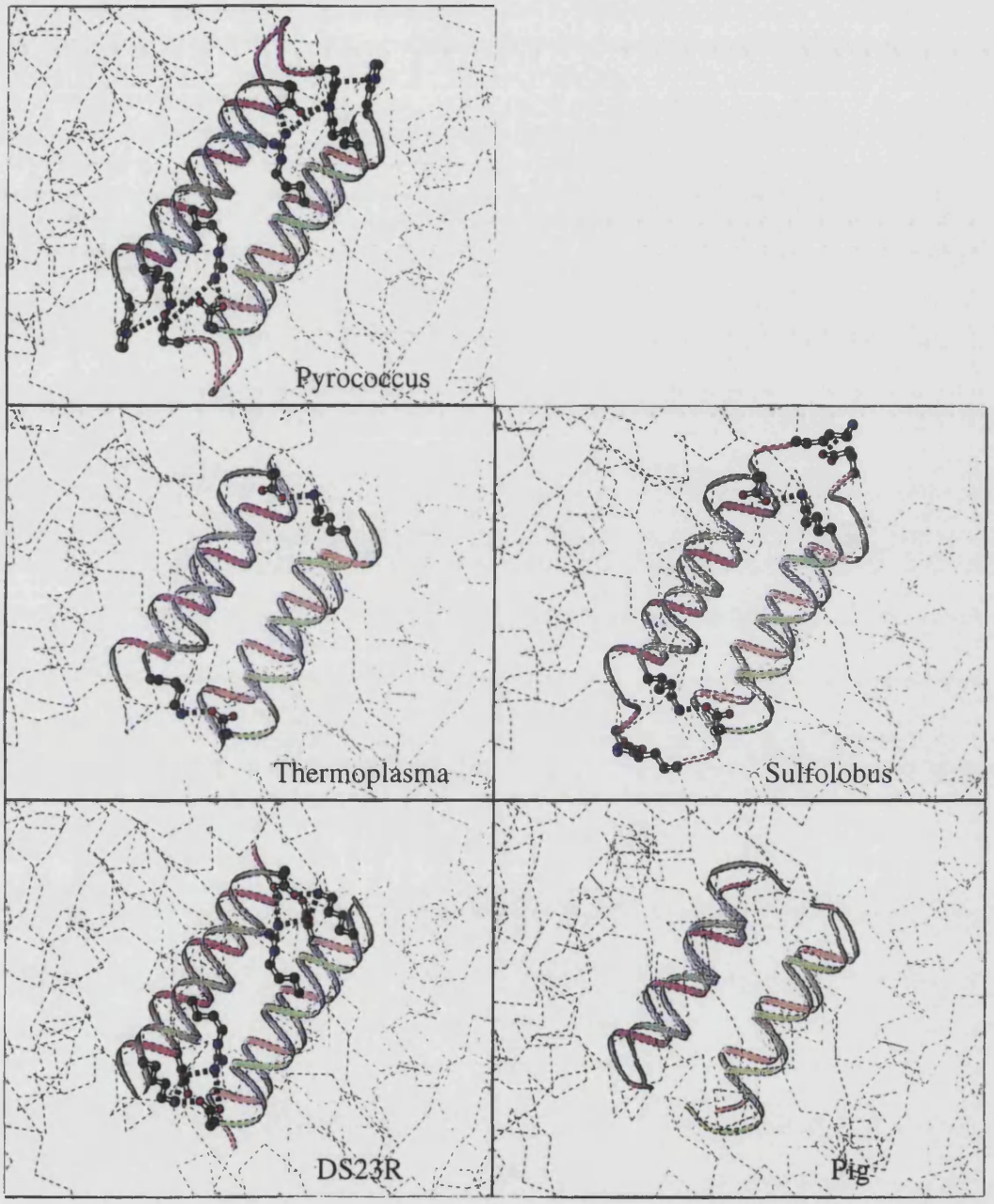


Figure 5.3 Ion pairs present at the 8 α-helical sandwich subunit interface region of CS (calculated at a cut-off distance of 4Å) for pig, DS23R, *S.solfataricus*, *T.acidophilum*, and *P.furiosus*. Helix G -brown. Helix M - green.

5 which are networked, while DS23r CS has two intersubunit ionic interactions and two intrasubunit interactions at the subunit interface. The presence of these complex ion pair networks has been observed in the structure of other hyperthermophilic enzymes such as glutamate dehydrogenase, and they are thought to be essential for stability at high temperatures (Yip et al., 1995, 1998).

The extensive ionic network at the subunit interface of *Pf*CS has been predicted to play a key role in the thermostability of the enzyme by maintaining the enzyme's dimeric integrity at raised temperatures (Russell et al., 1997) and may be one of the reasons why *Pf*CS is more thermostable than *Tp*CS.

It was proposed to test the prediction of the importance of the intersubunit ionic network to thermostability in *Pf*CS by using site directed mutagenesis to change residues involved in the network. Asp 113 was chosen as a target for mutagenesis as it is thought to participate in 60% of the ionic interactions at the *Pf*CS subunit interface, and because there is no equivalent residue to Asp113 found in the amino acid sequence alignment between *Pf*CS and *Tp*CS. From Fig 5.1c it can be observed that Asp113 forms ionic bonds with Lys 219, His 93, and Arg 99. Thus, by substituting Asp113 with another residue, the ionic bonding present at the subunit interface of *Pf*CS can be considerably disrupted.

This chapter describes the disruption of the ionic network at the subunit interface of *Pf*CS by the substitution of Asp113 with an alanine and a serine residue. Alanine was chosen as a replacement for Asp113 because its side chain cannot form any ionic interactions. Serine was chosen as a replacement for Asp113 because its side chain cannot form ionic interactions but it contains a

hydroxyl group which is capable of forming hydrogen bonds. The characterisation of both mutants with respect to *Pf*CS is then described.

Several other groups have investigated the importance of subunit interactions to protein thermostability. Pozo et al. (1996) investigated the importance of a Trp 139 in bacterial D-amino acid transferase. This residue is located in a hydrophobic pocket at the subunit interface region, and when substituted for alanine, proline or histidine, the thermostability of the enzyme was decreased significantly. It was concluded from this that interactions of the Trp-139 residue at the subunit interface are important for the catalytic activity and thermostability of the enzyme.

Erduran & Kocabiyik (1998) have investigated the role of hydrophobic interactions at the dimer interface of *Tp*CS. They did this by amino acid substitutions in the interface helices of *Tp*CS; Ala 97-Ser, Ala 104-Thr (these substituted amino acids are found at the equivalent position in the more hydrophilic pig CS subunit interface), Gly209-Ala (the alanine residue being found in the equivalent position at the subunit interface of *Pf*CS). They found that amino acid substitutions increasing or decreasing interface hydrophobicity led to an increase in the thermostability of *Tp*CS. They concluded from this that if one wants to increase the thermostability of an enzyme by engineering the subunit interface region, then one must take into account the packing effect of the residues as well as their degree of hydrophobicity.

Mandelman et al. (1998) made point mutations at the subunit interface of ascorbate peroxidase to change an intersubunit ion pair to determine the effects of a neutral and repulsive mutation on dimerization and stability. Gel filtration analysis indicated that the ratio of monomer to dimer increased as the

interactions went from attractive to repulsive. The thermal stability of the mutants was also decreased with respect to the wildtype.

Scopes et al. (1998) made mutations to reduce charge repulsion at the subunit interface of human glucose-6-phosphate dehydrogenase, and found that these mutations increased the thermostability of the enzyme.

Vetriani et al. (1998) noticed from a comparison of the structures of glutamate dehydrogenases from *Thermococcus litoralis*, and *P.furiosus*, that an intersubunit ion pair network was substantially reduced in *T.litoralis.* When two residues were mutated to restore this ionic network, the thermostability of the *T.litoralis* glutamate dehydrogenase enzyme was substantially increased.

Pappenberger et al., (1997) have investigated the importance of a four membered ionic network (intra subunit) in glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima* . They did this by replacing the central member of this ionic network (Arg 20) by replacing it with either alanine or asparagine. They found that resistance towards thermal denaturation was strongly diminished in the mutant enzymes. They concluded from this that the ionic network around the Arg 20 residue is very important for the thermal stability of this enzyme.

5.2 Methods

5.2.1.Nested PCR

The technique used to mutate the Asp113 residue was nested PCR. Nested PCR is a two step process. In the first step, PCR is carried out using a primer containing the mutation to be introduced. This produces a DNA fragment which is then purified and used as one of the primers in the second PCR reaction. Because the PCR 1 product is relatively large compared to a standard primer designed for PCR, the PCR conditions in the second PCR are different from the first to take account of problems associated with annealing of the larger primer. An outline view of the nested PCR process is shown in Fig 5.4.

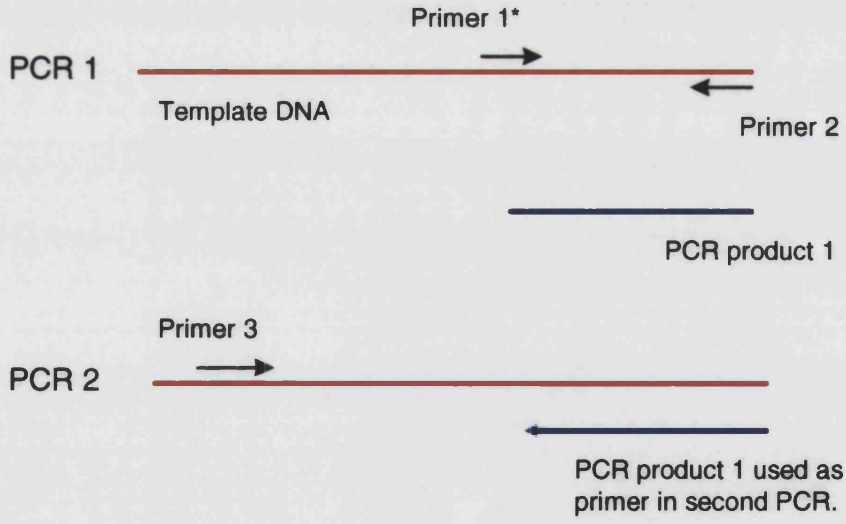


Figure 5.4 Steps in the nested PCR procedure. *Denotes the primer containing the mutation to be introduced.

5.3 Production of the Alanine and Serine *Pf*CS subunit interface mutants

5.3.1 Production of the first PCR product

The nested PCR technique was used to create the subunit interface mutants because there were no suitable restriction sites adjacent to the Asp 113 position that would have allowed a standard PCR mutagenesis approach (as used in chapter 4) to be used. The primers used in the PCR 1 reactions are shown in Fig 5.5. MAA17 was the sense primer designed to change the Asp113 amino acid residue to an alanine residue, and MAA18 was the sense primer designed to change the Asp113 residue to a serine residue. These primers bind to the DNA between positions 413 and 443 of the *Pf*CS DNA sequence. The sequences shown in bold code for alanine and serine respectively. MAA19 is the antisense primer used in each of the PCR 1 reactions and binds to the Apa1 site of *Pf*CS, between bases 759 and 736 (5'-3').

MAA17 5'GACGATAGTGGA**GCG**ATTCCAGTAACCCCTG 3'

MAA18 5'GACGATAGTGGA**TCC**ATTCCAGTAACCCCTG 3'

MAA19 5'CCATGGATT*GGGCCC*TTTAAAGC 3'

Figure 5.5 Primers used to create the PCR 1 products. The sequences shown in bold code for alanine and serine respectively. The nucleotides shown in italics represent the Apa1 site of *Pf*CS.

The PCR thermocycler programme used to create the PCR 1 product was the same as that described in section 2.2.11. The reagents used in the PCR reactions are shown below.

12pmoles of each primer

50μmol of each dNTP

50 ng pBAP 3031 template DNA

2 Units of Taq polymerase

x10 Taq polymerase buffer

5mM $MgCl_2$

Products from the PCR 1 reactions were run on a 1% (w/v) agarose minigel and were purified by the Gene Clean method. A photograph of samples of the purified PCR1 products run on a 1% agarose minigel is shown in Fig 5.6. It can be seen from this gel that the PCR 1 products were about 350 bp in length which was the expected size of the fragments.

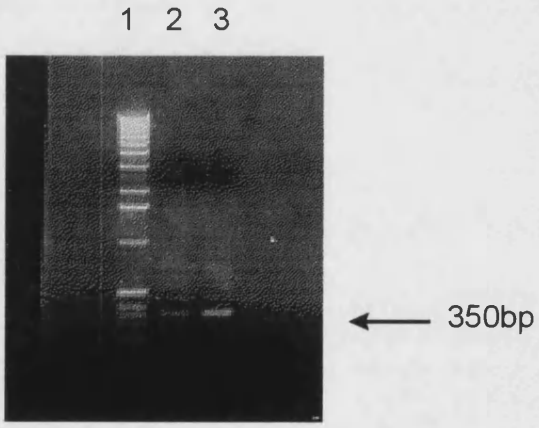


Figure 5.6 PCR 1 products. 1) Kilobase ladder DNA. 2) Serine mutant PCR 1 product. 3) Alanine mutant PCR 1 product.

5.3.2 Production of the second PCR product

While the purpose of the first PCR reaction was to create mutant DNA fragments for use as primers in the second PCR, the purpose of the second PCR reaction was to add an Nde1 site to allow the fragments to be cloned into the pBAP3031 construct (*Pf*CS in pRec7-Nde1).

The PCR 1 products produced and purified in section 5.3.1 were used as the antisense primers for the second PCR reaction, while primer MAA-7 (see section 3.5.1) was used as the sense primer. The PCR 2 conditions were as follows.

Thermocycler PCR programme and reagent concentrations used are those described by Landt et al., (1990). This was to take account of the difficulties involved in using a long primer in PCR.

Programme:

95°C For 1 min

55°C For 5 min

72°C For 2 min

72°C For 10 min after 30 cycles.

This was followed by a cooling step of 4°C until the samples were retrieved for analysis

Reagents:

50 ng pBAP3031 template DNA

12 pmoles MAA-7 primer

0.2μg PCR 1 product

2 units of Taq polymerase.

X10 Taq buffer

200 μM of each dNTP

6mM $MgCl_2$

Following the PCR, the PCR2 reactions were run on a 1% agarose minigel, (Fig 5.7). Gel bands of 700 bp could clearly be seen, which was the expected size of the PCR 2 products. The 700 bp bands were cut out of the gel and purified by the Gene Clean method.

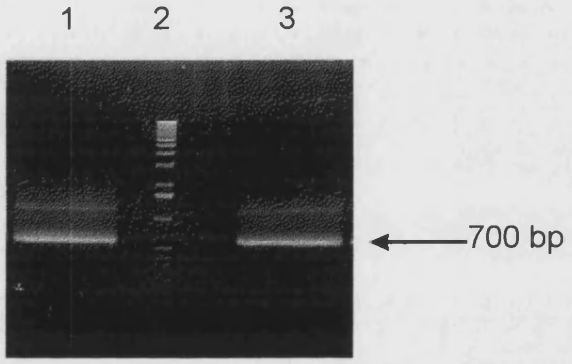


Figure 5.7 Products of PCR 2 reaction. 1) Serine mutant PCR 2 product. 2) Kilobase ladder DNA. 3) Alanine mutant PCR 2 product

5.3.3 Cloning of the PCR 2 products into pBAP3031

To clone the PCR 2 products into pBAP3031, they were double digested with Apa1 and Nde1; pBAP3031 DNA (pRec7-Nde1 vector containing the *Pf*CS gene) was also digested at the same time. This acted as a control that the Apa1 and Nde1 enzymes were actually cutting. The pBAP3031 Nde1-Apa1 digest was run out on a 1% agarose minigel and the 3.1 kb band was excised from the gel and purified using the Gene Clean procedure. The Apa1-Nde1 digest of the PCR 2 products was also run on a 1% agarose minigel and the 700bp bands excised from the gel and purified using the Gene Clean procedure. An overnight ligation was set up to ligate the band purified Nde1-Apa1 digested PCR 2 fragments into the band purified Apa1-Nde1digested *Pf*CS pRec7-Nde1 construct. The following day these ligations were transformed into the CS-minus

MOB154 *E.coli* strain, and were plated onto LA plates containing ampicillin at a concentration of 125μg/ml. A number of colonies were chosen from the vector + insert plates and were grown up and 'miniprep' DNA prepared from them.

It was possible to identify whether the recombinants were pBAP3031 or contained the subunit interface mutated Nde1-Apa1 fragments by restriction digestion, because by changing the DNA coding for the Asp 113 residue to that coding for a serine and an alanine residue, a BamH1and Hinf1 site, were created respectively. There is only one BamH1 site in the pBAP3031 construct and thus a serine subunit interface mutant (Asp113/Ser) would give 2 bands when digested with BamH1. The pBAP3031 construct contains 5 Hinf1 sites, so an alanine subunit interface mutant (Asp113/Ala) would give an altered band pattern when digested with Hinf1.

'Miniprep' DNA from recombinants (from the cloning of both subunit interface mutants into the Apa1-Nde1 digested *Pf*CS pRec7-Nde1 construct) was digested with BamH1 and Hinf1 and run on a gel (Fig 5.8). It can be seen from this gel that clone 6 gave two bands with BamH1, which identified it as an Asp113/Ser mutant (designated pBAP3040), and that clone 2 gave an altered band pattern when digested with Hinf1 thus identifying it as an Asp113/Ala mutant (designated pBAP3041). Note the digests run in lanes 2 and 7 were for the purpose of checking DNA stocks of the pRec7-Nde1 vector and pBAP3031.

Figure 5.8 Restriction digest of potential subunit interface mutants. 1) kilobase ladder DNA. 2) pRec7-Nde1 vector digested with Nde1. 3) *Pf*CS in pRec7-Nde1 (pBAP3031) digested with Hinf1. 4) Clone 2, Asp113/Ala mutant digested with Hinf1. 5) *Pf*CS in pRec7-Nde1 (pBAP3031) digested with BamH1. 6) Clone 6, Asp113/Ser mutant digested with BamH1. 7) pBAP3031 Nde1-KpN1 digest 8) Uncut pBAP3031

5.3.4 Sequencing of subunit interface mutants pBAP3040 and 3041

Although the number of bands obtained from the BamH1 and Hinf1 digest were as predicted, the sizes of the bands were smaller than expected. For this reason it was decided to sequence completely the 1.3 kb CS inserts of the pBAP3040 and pBAP3041 clones. Primers MAA-9, and MAA-10 ,were used to sequence in from either end of the insert and MAA-11 was used to sequence from base position 518 (on the sense strand of *Pf*CS).

The pBAP3041 sequence was found to have a 119bp deletion from base 625 to base 744, and an insertion of one T base at position 626 (a deletion from the BstX1 site to the Apa1 site), and occurred after the intended Asp113/ala mutation (see Fig 5.9). The remainder of the sequence of the 1.3 Kb insert was as expected. The insertion of the extra T base at position 626 alters the reading frame of the CS which resulted in a truncation of the protein at

amino acid residue 187. The reason why this deletion occurred is hard to explain, although there is a region adjacent to the BstX1 site which resembles an Apa1 site, and it is possible that the Apa1 enzyme cut at this position as well as the Apa1 site at position 745.

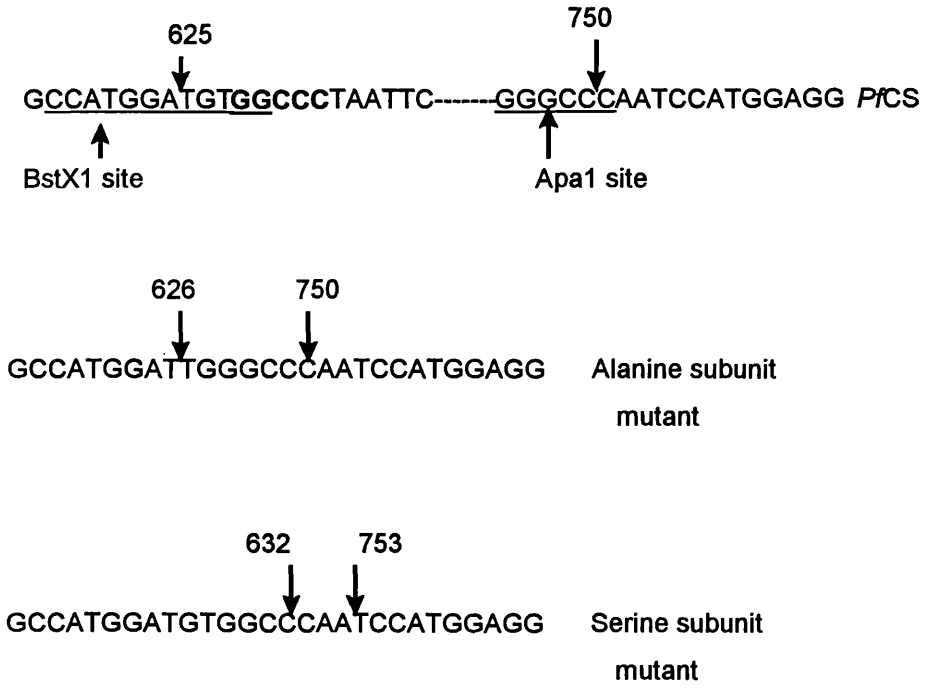


Figure 5.9 DNA base sequence of the subunit interface mutants spanning the deleted region. The *Pf*CS sequence is shown at the top, the dashed lines represent bases between position 638 to 745, and the sequence adjacent to the BStX1 1 site which was cut by Apa1 is shown in bold. The natural Apa1 and BstX1 sites are underlined on the *Pf*CS sequence.

pBAP3040 (Asp113/Ser mutant) was found to have a similar deletion to pBAP3041 in that there was a deletion from base number 632 to base 750. Again this deletion threw out the reading frame of this mutant resulting in truncation of the protein at amino acid residue 188. The sequence obtained at

this region is shown in Fig 5.9; the sequence of the remainder of the 1.3 kb CS insert was as expected.

5.3.5 Remaking the two subunit interface mutants

The large deletions in both subunit interface mutants made it necessary to remake both subunit interface mutants. However, because it was not clear how the extra mutations arose in the subunit interface mutants, pBAP3040 and pBAP3041, it was thought best not to repeat the same cloning procedure, in case the same deletions occurred again. As the rest of the sequence of both of the subunit interface mutants was correct, it was decided to cut out a fragment containing the subunit interface mutation from pBAP3040 and pBAP3041 (but none of the extra mutations), and clone each fragment into pBAP3031 (*Pf*CS in pRec7-Nde1). Fig.5.10 shows the restriction sites used for remaking the 2 subunit interface mutants.

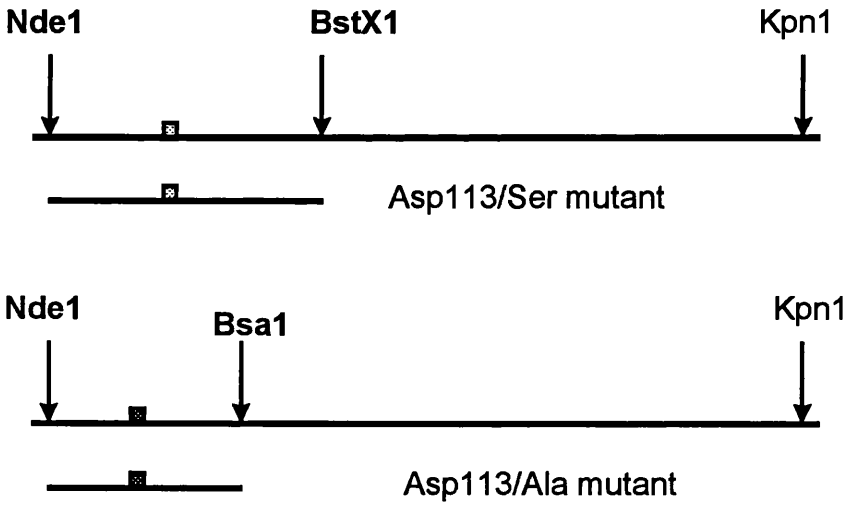


Figure 5.10 Restriction sites used to subclone the subunit interface mutants into *Pf*CS in pRec7-Nde1 (pBAP3031). The coloured boxes represent the mutated Asp113 residue. The restriction sites used to subclone the subunit interface mutants are shown in bold.

5.3.5.1 Remaking the serine subunit interface mutant

The strategy used to remake the Asp113/ser mutant was to cut out the region containing the Asp113/ser mutation on an Nde1-BstX1 fragment, and clone this into pBAP3031 (*Pf*CS in pRec7-Nde1), which had the corresponding Nde1-BstX1 fragment previously removed. A large scale digest of pBAP3040 and pBAP3031 was set up with the enzymes Nde1 and BstX1. The digests were run on a 1% agarose minigel and the 540 bp Asp113/ser mutant Nde1-BstX1 fragment, as well as the 3.3 kb *Pf*CS Nde1-BstX1 fragment, were cut out of the gel and purified using the Gene Clean method (Fig 5.11).

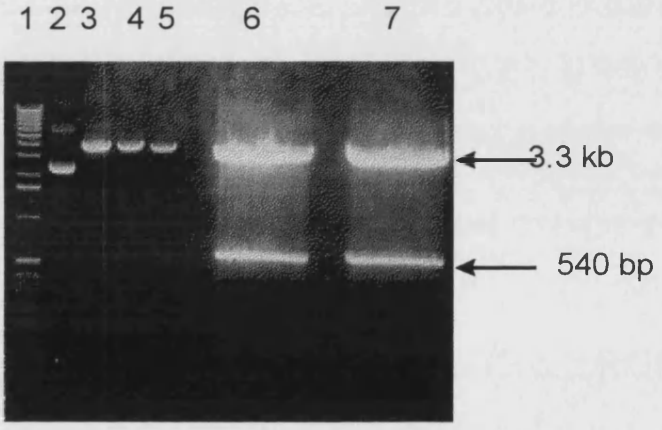


Figure 5.11 Restriction digest of pBAP3040, and pBAP3031 with Nde1 and BstX1. 1) Kilobase ladder. 2) Uncut pBAP3031. 3) pBAP3031 digested with BstX1. 4) pBAP3040 digested with BstX1. 5) pBAP3040 digested with Nde1. 6) pBAP3031 digested with Nde1 and BstX1. 7) pBAP3040 digested with Nde1 and BstX1.

An overnight ligation was set up at 15°C of the purified 540 bp Nde1-BstX1 subunit interface fragment and the purified 3.3 kb *Pf*CS Nde1-BstX1 fragment. The following day the ligations were transformed into the CS-minus

E.coli strain MOB154, colonies were picked and 'miniprep' DNA was prepared from them. The 'miniprep' DNA was digested with BamH1 (Fig 5.12). It can be seen that all the clones, except for number 3 (lane 4), contained the serine subunit interface mutation. The reason that there is such a variation in the banding pattern between clones in Fig 5.12 is that some of the 'miniprep' DNA samples have cut better than others.

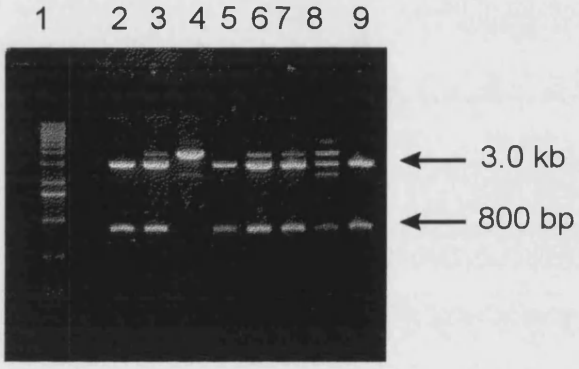


Figure 5.12 BamH1 digest of potential serine subunit interface mutants. Key. 1) Kilobase ladder. 2-9) Clones 1-8 inclusive.

Clone 2 had its insert sequenced completely using the sequencing primers MAA-9, MAA-10, and MAA-21. The sequence showed that this clone contained the serine subunit interface mutation but did not contain any other mutations. The repaired Asp113/Ser mutant within the pRec7-Nde1 vector was designated pBAP3014.

5.3.5.2 Remaking the alanine subunit interface mutant

The strategy for remaking the Asp113/Ala mutant was to cut out the region containing the Asp113/ala mutant on an Nde1-Bsa1 fragment and clone this into pBAP3031 (*Pf*CS in pRec7-Nde1), which had the corresponding Nde1-

Bsa1 fragment previously removed. The BstX1 site could not be used for remaking this mutant as the BstX1 site was destroyed in pBAP3041by the 119bp deletion and T base insertion (see section 5.3.4). The Bsa1 site was not chosen for repairing the serine subunit interface mutant as the pBAP3031 construct has two Bsa1 sites (which complicates any cloning using a Bsa1 site).

A large scale digest of pBAP3041 and pBAP3031 was set up with the enzyme Bsa1. After 3 hours a sample was taken from this digest and run on a 1% agarose minigel. This gave two bands of 2.0 kb and 1.8 kb in size (gel not shown); Nde1 restriction enzyme was added to the remainder of the Bsa1 digest and left for a further 3 hours. The doubly-digested DNA was run on a 1% agarose minigel (Fig 5.13), showing that three bands were produced from each clone. The 450 bp Nde1-Bsa1 fragment from pBAP3041, and the 1.3 kb and 2.0 kb fragments were excised from the gel and purified using the Gene Clean method. An overnight ligation was set up of the three purified Nde1-Bsa1 fragments at 15°C. The following day the ligations were transformed into the CS-minus *E.coli* strain MOB154; no colonies appeared on the vector only plates but several colonies appeared on the vector + insert plates, these were picked and 'miniprep' DNA was prepared from them. The 'miniprep' DNA was digested with Nde1 and run on a 1% agarose gel to check that the recombinants were 3.8 kb in size (gel not shown). The 'miniprep' DNA of one recombinant was sequenced using primer MAA-11 to determine whether the 119 bp deletion had been repaired, and the rest of the insert was sequenced using the primers MAA-9 and MAA-10 to check if any other mutations were present. None were found. The 'repaired' alanine subunit interface mutant was designated pBAP3015.

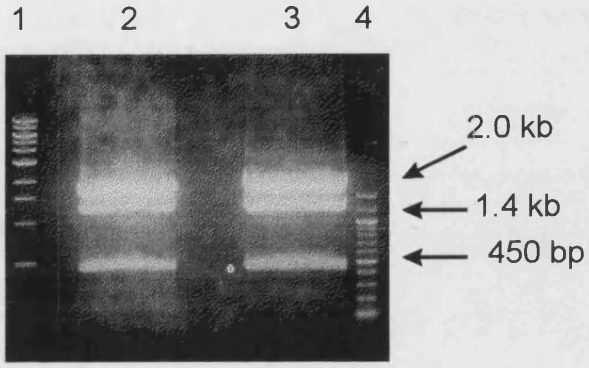


Figure 5.13 Restriction digestion of pBAP3041, and pBAP3031 with Nde1 and Bsa1. 1) Kilobase ladder. 2) pBAP3041 doubly-digested with Nde1, and Bsa1. 3) pBAP3031 doubly-digested with Nde1 and Bsa1. 4) 100 bp marker.

Now that both the serine and alanine subunit interface mutants had been confirmed by sequencing, the next step was to over express them and obtain pure protein for the purpose of characterisation.

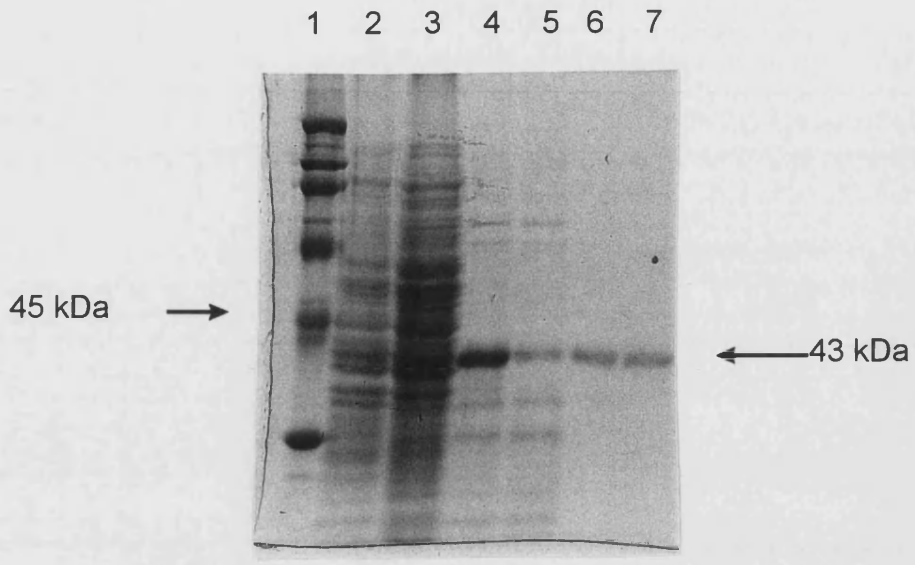
5.4 Production and purification of the 2 subunit interface *Pf*CS mutants

50ml of LB containing ampicillin at a concentration of 125μg/ml was inoculated with a single colony of *E.coli* MOB154 (CS-minus strain) transformed with the plasmid pBAP3015; this was grown up overnight at 37°C in a shaking incubator. The following day this culture was used to inoculate 500ml of LB containing ampicillin at a concentration of 125μg/ml in a 1 litre flask; this was incubated overnight on a shaking incubator at 37°C (at 200rpm) and the

following day nalidixic acid was added to the culture to give a final concentration of 50μg/ml. The cells were then incubated overnight at 37°C in the shaking incubator and were harvested the following morning. The same method was used to grow up cells of *E.coli* MOB154 transformed with pBAP3014. 0.2g of cells were used to purify the CS (see sections 2.3.2 and 2.3.3). Experiments were done to determine a suitable heat step for removing most of the *E.coli* host proteins without losing too much of the CS activity. A heat step of 80°C for 15 min was found to be adequate. Hence the two subunit interface mutants were purified in the same way as the wild type *Pf*CS. Table 5.1 shows the yields and specific activities obtained from each step of the purification, starting with 0.2g of cells. Fig 5.14 shows SDS-PAGE gels of samples taken during the purification of the serine and alanine subunit interface mutants.

CS Construct / step	Volume (ml)	Activity /ml (Units/ml)	Total Activity (Units)	Protein (mg/ml)	Total Protein (mg)	% Yield	Specific Activity (Units/mg)
Asp113/Ser mutant							
Cell extract	1.0	8.4	8.4	8.1	8.1	100	1*
Heat Step	0.8	6.2	5	0.9	0.7	59	7*
Red gel A	3.0	1.2	3.5	0.086	0.25	42	14*
Asp113/Ala mutant							
Cell extract	1	7.6	7.6	7.85	7.85	100	1.0*
Heat step	0.75	4.3	3.2	0.8	0.6	43	5.6*
Red gel A	3.0	0.8	2.3	0.06	0.17	30	13.5*

Table 5.1 Purification of the Asp113/Ser and Asp113/Ala subunit interface mutants. All assays were carried out at 55°C. * indicates that the specific activity values are low because of substrate inhibition by Ac-CoA.



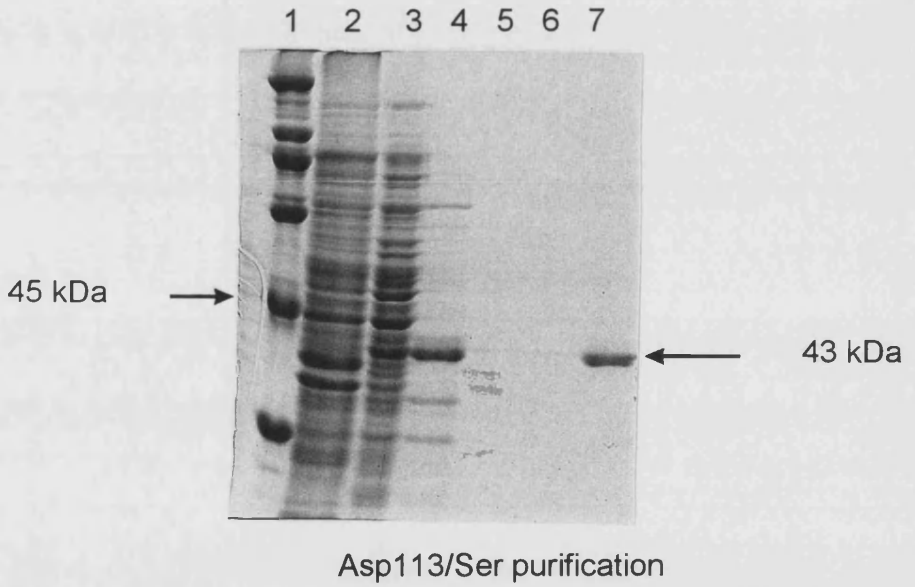


Figure 5.14 SDS-PAGE analysis of fractions from the purification of the Asp113/Ser and Asp113/ala mutants. 1) Low range molecular weight markers. 2) Cell extract. 3) Heat step. 4) Red gel A unbound fraction. 5-7) Red gel A eluted fractions.

5.5 Characterisation of the subunit interface *Pf*CS mutants.

5.5.1 Determination of the kinetic perameters

The K_m and V_{max} values for both subunit interface mutants were determined in the same way as the domain swap mutants (see section 3.7.1). All measurements were repeated three times at 55°C; the mean, and standard errors for the K_m and V_{max} values are shown in Table 5.5.1.. The direct linear plot of Eisenthal and Cornish-Bowden (1974) was used to determine the K_m and V_{max} values for the substrate oxaloacetate. However, because both subunit interface mutants were substrate inhibited by Ac-CoA, the K_m and V_{max} values for the substrate Ac-CoA were determined by using the Scientist plot to fit the rates to a curve given by the equation:

$$v = \frac{Vmax}{1+\frac{Km}{S} + \frac{S}{Ki}}$$

The specific activity values shown in table 5.2 are calculated from the means of the V_{max} values calculated for the substrate Ac-CoA, since the V_{max} values calculated for the substrate oxaloacetate by the direct linear plot do not give a true value of V_{max} because of substrate inhibition by Ac-CoA. The K_m and V_{max} values for OAA for the Asp113/Ala, and Asp113/Ser mutants were determined at an Ac-CoA concentration of 24μM and 9μM respectively, as these were the concentrations of Ac-CoA which gave the greatest rates of activity at the enzyme concentrations used.

Citrate Synthase	Km OAA (μM)	Km Ac-CoA (μM)	K_i Ac-CoA (μM)	Specific Activity (μmol/min/mg)
P.furiosus	10 ±1.3	3.0 ± 0.7		17 ±2
Serine subunit mutant	11 ±1	3.5 ± 0.8 *	76 ± 2	27 ± 2
Alanine subunit mutant	13 ±3	3.4 ± 0.2 *	122 ± 28	28 ± 2

Table 5.2 K_m and specific activity values determined for *Pf*CS, and the subunit interface mutants at 55°C. *indicates enzyme substrate inhibited by Ac-CoA.

V against S plots of the data used to determine the K_m and V_{max} values for the serine and alanine subunit interface mutants are shown in Figs 5.15a-d. These graphs were plotted in the same way as those in section 3.7.1 using the Microcal Origin graphics package.

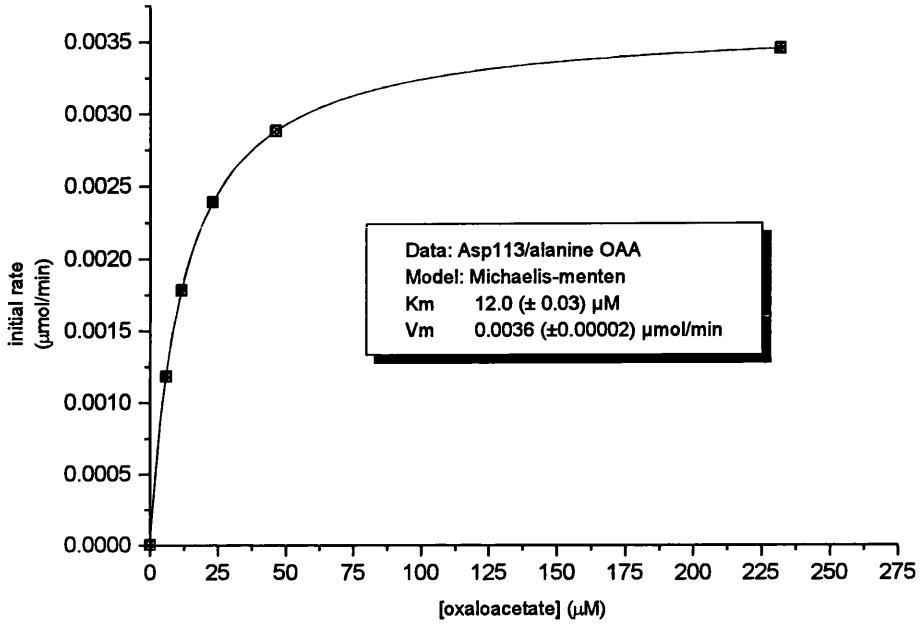


Figure 5.15a V against S plot for the Asp113/Ala mutant for the substrate oxaloacetate.

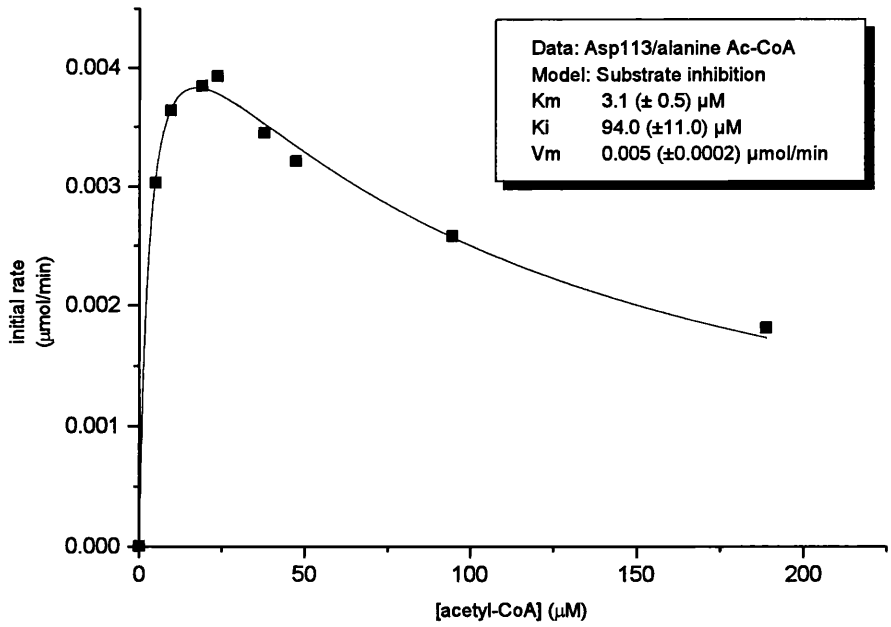


Figure 5.15b V against S plot of data used to determine the K_m and V_{max} values for the Asp113/Ala mutant for the substrate Ac-CoA.

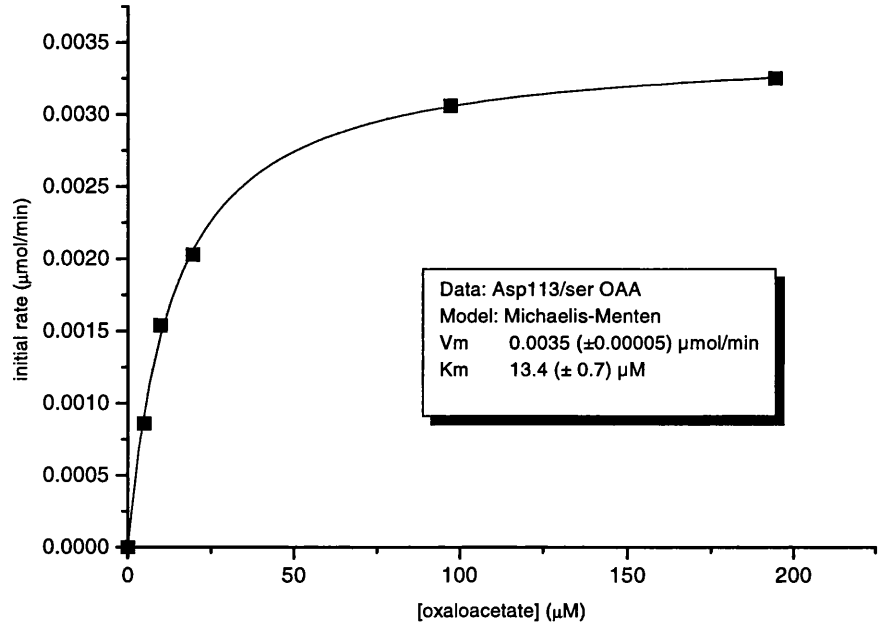


Figure 5.15c V against S plot for the Asp113/Ser mutant for the substrate oxaloacetate.

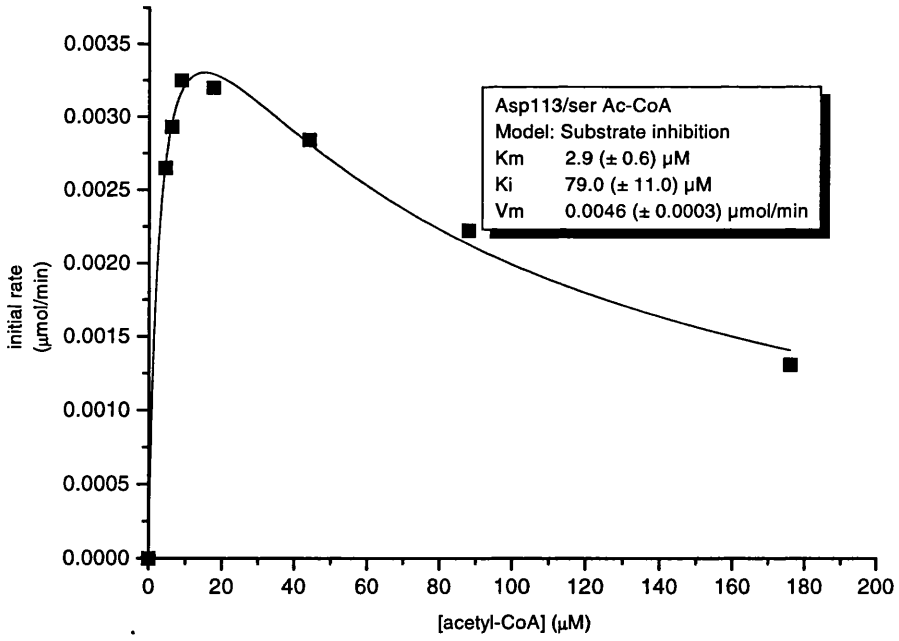


Figure 5.15d V against S plot of data used to determine the K_m and V_{max} values for the Asp113/Ser mutant for the substrate Ac-CoA.

5.5.2 Thermal stability of the subunit interface CS's

The aim of these experiments was to compare the subunit interface mutants with *Pf*CS in terms of their abilities to withstand thermal denaturation

All thermal stability studies were carried out in the presence of 50mM phosphate, pH 7.2, 2mM EDTA using an enzyme concentration of 0.1mg/ml. The buffer was heated to the required temperature in an Eppendorf tube in a bath containing polyethylene glycol grade 400 (PEG bath). The emperature was measured using an electronic temperature probe inside an Eppendorf tube containing PEG grade 400 solution, placed in the bath. When the desired buffer temperature was reached enzyme was added to the Eppendorf tube, mixed with the buffer and the timer started. Aliquots were removed at timed intervals and placed into thin walled glass tubes on ice, where they were stored prior to assay. A time zero value was obtained by assaying an 0.1mg/ml concentration of enzyme in the same buffer at 4°C. Samples were assayed in triplicate and a mean ln (% activity remaining) value was calculated at each time point for each sample. A graph was plotted of the ln% (activity remaining) versus time to show the relative thermal stability of the two subunit interface mutants and *Pf*CS at 103°C (Fig 5.16). From this graph it can be seen that in terms of thermostability *Pf*CS >serine subunit interface mutant >alanine subunit interface mutant. The rate constants for inactivation, and the times at which 50% of the enzymes activity remain at 103°C for *Pf*CS and the two subunit interface mutants are shown in Table 5.3.

Citrate Synthase construct	Rate constant for inactivation (min^{-1})	Time at which 50% of activity remains (min)
***Pf*CS**	0.057 ± 0.002	12
Asp113/Ser	0.16 ± 0.005	4.3
Asp113/Ala	0.23 ± 0.012	2.6

Table 5.3 Rate constants for denaturation of the subunit interface mutants and *Pf*CS, and times at which 50% of the enzymes activity remains at 103°C.

From the rate constants for denaturation at 103°C calculated for the wildtype *Pf*CS and the subunit interface mutants, the change in the free energy of activation of thermal inactivation for both mutants compared to the wildtype *Pf*CS was determined (using the method given in section 2.4). It was found that there was a decrease in the free energy of activation for thermal unfolding of about 3 and 4.2 kJ/mol at 103°C for the Asp113/Ser and Asp113/Ala mutants respectively.

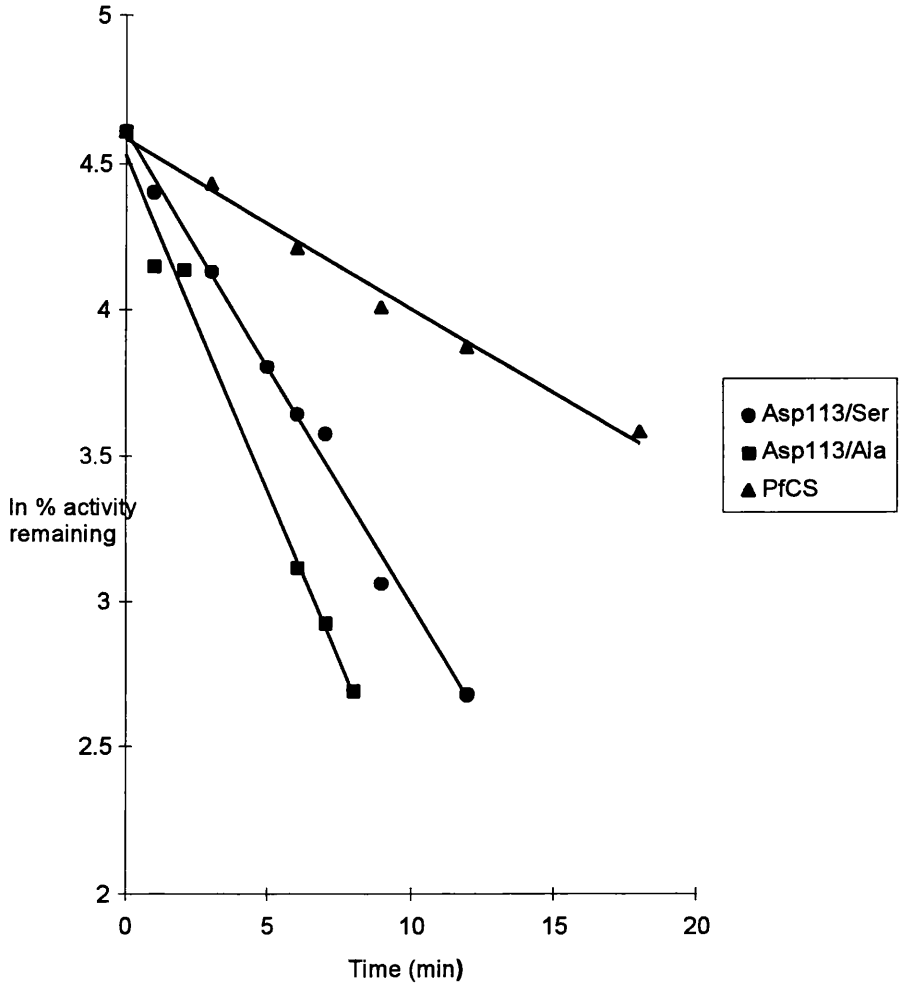


Figure 5.16 Relative thermostability of the subunit interface mutants to *Pf*CS at 103°C. The lines are drawn by linear regression.

5.6 Discussion

5.6.1 Kinetic Studies of the mutants

It can be seen from Table 5.2 that both the subunit interface mutants were active, and showed very similar K_m values to *Pf*CS with the substrates oxaloacetate and Ac-CoA (with the exception being that both mutants were substrate inhibited by Ac-CoA). Thus both mutants can be assumed to be folding in a similar way to *Pf*CS. However there must be a subtle difference in the folding as both mutants were substrate inhibited with Ac-CoA, a phenomenon also observed with the TP::pyr CS domain swapped mutant (see section 3.8.2), and the C-terminal Pf-2 mutant (see section 4.7.1). This substrate inhibition is an unexpected consequence of mutating the Asp113 residue, and it is difficult to explain why the mutation of the Asp113 residue should cause substrate inhibition with Ac-CoA. Both mutants showed a greater specific activity than *Pf*CS; this could be a consequence of the partial destruction of the ionic network at the subunit interface, giving rise to an enzyme which is more flexible than *Pf*CS at 55°C.

5.6.2 Thermal inactivation of the subunit interface mutants

From Fig 5.3 (showing the relative thermostabilities of the two subunit interface mutant CS's to *Pf*CS at 103°C) it can be seen that in terms of thermostability *Pf*CS >Asp113/Ser mutant >Asp113/Ala mutant. The Asp113/Ala mutant had a half life at 103°C of almost half that of the Asp113/Ser mutant, and a quarter of *Pf*CS. This indicates that the resistance towards thermal unfolding is strongly diminished in both subunit interface mutants. This suggests that the aspartate 113 residue plays an important role in the 5

membered ionic network, and the thermostability of *Pf*CS (as predicted from the crystal structure by Russell et al. [1997]). The decrease in the free energy of activation of thermal inactivation values, calculated for both mutants show good agreement with that determined by Pappenberger et al.(1997). These authors found that when they substituted the central residue in a 4 membered ionic network with non charged residues (in glyceraldehyde-3-phosphate dehydrogenase), the free energy of activation of thermal inactivation was decreased by about 4 kJ/mol, as compared to 3 and 4.2 kJ/mol for the Asp113/Ser and Asp113/Ala mutants, respectively, in the current study. However, it must be noted that Pappenberger et al (1997) were investigating an intramolecular ionic ring network rather than an intermolecular ionic ring network.

If one accepts that there is a difference in the thermostabilities of the two mutants one explanation why the Asp113/Ser mutant is more thermostable than the Asp113/Ala mutant is that the hydroxyl group from the side chain of serine at position 113 may form a hydrogen bond with His 93 (as this is within hydrogen bonding distance from Ser 113), thus creating an extra 2 hydrogen bonds at the subunit interface compared with the Asp113/Ala mutant, whose side chain cannot form any hydrogen bonds.

Although these results strongly suggest that the Asp113 residue, and the 5 membered ionic ring network at the subunit interface of *Pf*CS, plays an important role in the thermostability of *Pf*CS, they also support a model which proposes that the thermostability of *Pf*CS is made up of a number of thermostabilising interactions, since when this residue is substituted destroying 60% of the ionic interactions at the subunit interface the thermostability of the enzyme is reduced but not destroyed.

To extend this work, further mutants could be made where the *Pf*CS subunit interface ionic network could be disrupted in a different position, or in addition to the mutation of the Asp113 residue, Asp206 could be mutated to destroy all of the ionic interactions at the subunit interface. Another strategy to investigate the importance of this ionic network to thermostability would be to create an ionic network at the subunit interface of *Tp*CS similar to that found in *Pf*CS by site directed mutagenesis to see if this increases the thermostability of *Tp*CS.

5.7.3 Concluding Comments

These results show that the subunit interface ionic network in *Pf*CS makes a contribution to the enzyme's thermostability at high temperatures as predicted from the crystal structure. The fact that the enzyme's thermostability is not drastically changed when 60% of the interactions are destroyed suggests that it is not just the 5 membered ionic network at the subunit interface of *Pf*CS that is responsible for making *Pf*CS more thermostable than *Tp*CS (which lacks a subunit interface ionic network).

CHAPTER 6

Discussion

A number of predictions have been made from a comparison of the crystal structures of *Pf*CS, *Tp*CS, and pig CS, as to what structural features could lead to the enhanced thermostability of the *Pf*CS relative to *Tp*CS, and pigCS (Russell et al., 1997). These include: an increased compactness of the enzyme, a more intimate association of the subunits, an increase in the number of ion pairs and a reduction in the number of thermolabile residues.

Mutagenesis has been used to test some of these predictions. As described in chapter 3, domain swap mutants between the small domain of *Pf*CS and *Tp*CS were generated in order to determine where the principle determinants of thermostability and catalytic activity lie in *Pf*CS and *Tp*CS. These mutants were confirmed by N-terminal and DNA sequencing, and showed that in *Pf*CS and *Tp*CS the small domains are largely responsible for the catalytic activity of the enzyme, and that the structural determinants of thermostability were principally located within the large domains. Note the subunit interface comprises interactions between the large domains.

As a result of these findings it was decided to investigate the large domain of *Pf*CS further, to try to determine which structural features within the large domain contributed to the enhanced thermostability of the *Pf*CS. To this end, structural features which were predicted to play a role in the thermostability of *Pf*CS (and which lie within the large domain), namely the C-terminal arm and the 5-membered ionic network at the subunit interface of *Pf*CS, were mutated.

As described in chapter 4, truncation mutants were made of *Pf*CS. These were designed to test the importance of the 2 C-terminal residues (one of which is involved in a ionic interaction) and the whole C terminal arm (from where it exits the core of the molecule) to thermostability. These mutants were made by a PCR-based mutagenesis method and were confirmed by DNA sequencing. Both the Pf-2 and the Pf-13 mutants were significantly less thermostable than the wildtype *Pf*CS, which indicates the importance of the C-terminal region to thermostability within *Pf*CS. The Pf-2 mutant showed a two-fold increase in specific activity compared with *Pf*CS, which was attributed to an increase in the flexibility of the enzyme at 55°C, while the Pf-13 mutant showed greatly increased K_m values for the substrates oxaloacetate and Ac-CoA, and a two-fold decrease in the specific activity compared with *Pf*CS. One suggestion to explain the change in specific activity and in K_m values in the Pf-13 mutant, is that the removal of the whole external part of the C-terminal arm (which forms part of the subunit interaction) affects the subunit interaction between the monomers, causing a small conformational change, leading to an increase in the K_m values for the two substrates and a decrease in specific activity.

As described in chapter 5, subunit interface mutants were made where the Asp113 residue, which is the central residue in the subunit interface 5-membered ionic network, was substituted with an alanine and a serine residue, destroying 60% of the ionic interactions at the subunit interface. Both mutants were less stable than the wildtype *Pf*CS, indicating the importance of the ionic interactions at the subunit interface to thermostability in *Pf*CS. Both mutants displayed a greater specific activity at 55°C compared to *Pf*CS, and this was attributed to an increase in the mutant enzyme's intrinsic flexibility as a result of the disruption of this ionic network.

These results confirm the predictions made from the structural comparison of pig, *Tp*, and *Pf*CS that the 5-membered subunit interface ionic network, and C-terminal interactions play a role in the thermostability of *Pf*CS. The changes in thermostability observed in the C-terminal and subunit interface mutants (chapters 4 and 5) suggest a model of protein thermostability of *Pf*CS being composed of a number of interactions rather than just a few, since the changes in thermostability resulting from these mutations are minor ones. These results also show for the first time that in *Pf*CS the principal determinants of thermal stability lie within the large domain of *Pf*CS, and the principal determinants of catalytic activity reside in the small domain. The relationship of these results to other mutagenesis/thermostability studies is discussed in the discussion sections of chapters 3, 4, and 5.

A number of problems were encountered in the course of this study. These include the speed and reliability of generating mutants, experimental difficulties in assessing thermostability at around 100°C, and unexpected changes in enzyme properties. Both the pAlter and nested PCR methods used for carrying out mutations had inherent problems. The pAlter method was very slow (see chapter 3) and the nested PCR method gave rise to unwanted extra mutations (see chapter 5).

It was difficult to assess enzyme thermostability at temperatures around 100°C by following the denaturation of the mutant and wild type enzymes, because of the problem of replicating the same conditions of temperature when working on different days . This may explain why there is such a difference in the half life times of *Pf*CS at 103°C, as described in chapters 4 and 5. Therefore, the inactivation rate constants determined for the subunit interface

mutants and the C-terminal mutants should be considered separately. For this reason one cannot really comment on whether the C-terminal mutants are more thermolabile than the subunit interface mutants. One way to get around this problem would be to ensure that one was in a position to assay all of the mutants in triplicate on the same day, or one could try carrying out the denaturations at lower temperatures for much longer periods of time. Alternatively the results could be presented as an Arrhenius plot and in this way one would not have to replicate the temperatures from day to day. Other groups such as Pappenberger et al., (1997), have used CD spectroscopy to follow denaturation, but the problem with this approach is that it assumes that the enzyme is denatured when it unfolds whereas a local change in structure (as opposed to unfolding) may be responsible for the loss of enzyme activity.

The global nature of interactions within a protein have made it difficult to explain some of the observations made during this project. An example of this is the unexpected change in some of the mutants' properties, i.e. the substrate inhibition effect with Ac-CoA that has been observed with the TP::pyr, Pf-2, and both of the subunit interface mutants. What is difficult to understand is why such differing structural changes should give rise to the same effect, and why in the case of the C-terminal mutants should the Pf-2 mutant, but not the Pf-13 mutant, be inhibited by Ac-CoA . One explanation for the Ac-CoA substrate inhibition in the affected mutants is that in *Pf*CS and *Tp*CS, oxaloacetate binds first causing a conformational change in the enzyme enabling it to bind Ac-CoA. In the substrate inhibited mutants however the Ac-CoA may be binding before the oxaloacetate, preventing the binding of the oxaloacetate.

This structural comparison/mutagenesis strategy for investigating the structural determinants of thermostability is hampered in this case by the lack of

a crystal structure of a CS from a mesophilic archaeon, and lack of a crystal structure of the closed form of *Tp*CS. If there was a crystal structure of a mesophilic archaeon then one could be a little more certain that differences in structure that are observed in moving from a mesophile to a thermophile and hyperthermophile are not of a phylogenetic origin. Also the crystal structure of the closed form of *Tp*CS may reveal that there is very little difference in the C-terminal interactions between *Pf*CS and *Tp*CS. Hence one suggestion for future work would be to obtain the crystal structure of the closed form of a mesophilic archaeon, and the closed form of *Tp*CS.

Another suggestion for future work would be the generation of further mutants. As well as the additional mutants described in the discussion sections in chapters 4 and 5, one could try mutating other regions within *Pf*CS. For example one could mutate the intersubunit ion pair located at the N-terminus Lys 8-Asp16' *Pf*CS, to see if this ionic interaction played a role in the enzyme's thermostability. One could extend the domain swap study by swapping the small domain of the psychrophilic organism DS23r with that of *Pf*CS to see what the thermostability of the resultant chimaeric enzymes would be, given the vastly different thermostabilities of the wildtype enzymes. This would also test the validity of our conclusions from the present domain swap work.

An alternative approach for investigating what makes *Pf*CS more thermostable than its counterpart from *T.acidophilum* would be to use an evolutionary technique with *Tp*CS within *P.furiosus,* accumulating stabilising mutations by recombination, and random mutagenesis, an approach used by Akanuma et al. (1998). The aim of using this approach would be to determine what mutations evolve naturally to increase the thermostability of the *Tp*CS. To

produce mutants using the *in vivo* evolutionary technique one would first have to generate a CS-minus strain of *P.furiosus* and then integrate the CS gene from *T.acidophilum* into its chromosome. The resulting *Pf* strain (containing the chromosomally incorporated foreign CS) would then be cultured in conditions of increasing temperature to select for mutants of the CS which had evolved to withstand the increased temperatures. Alternatively one could use directed evolution, a method that involves random mutagenesis, but does not require *in vivo* selection in an extremophilic host (Kuchner & Arnold, 1997). The CS from these mutants would then be sequenced, the protein purified, and characterised to determine if the mutations had changed the structure of *Tp*CS in ways that were predicted from Xray crystallographic comparisons of these enzymes with *Pf*CS.

There remains a considerable scope for research into the area of protein thermostability and much work is required to unravel the complexities of the strategies of stabilisation employed in proteins from hyperthermophiles. The research carried out as part of this project has located the principal determinants of thermostability to the large domain in CS in the Archaea studied. The creation of more mutants within the large domain may shed more light on the structural determinants of protein hyperthermostability within the Archaea.

APPENDIX

Appendix A

Primers used for mutagenesis and sequencing reactions.

MAA-1 5' CGCCGGGAGTTCCCGGGCGACGTACCGCC 3'

Used to introduce an Apa1 site into *Tp*CS.

MAA-2 5' GTTGACTACTGGTCCGGATTAGTTTTCTAT 3'

Used to introduce a BspE1 site into *Pf*CS.

MAA-3 5' TGCATCTACCACGGCAGGCT 3'

Used for sequencing, binds to the sense strand of *Tp*CS between positions 842 and 862.

MAA-4 5' AGGTATATGAGATCGCCACG 3'

Used for sequencing, binds to the sense strand of *Tp*CS between positions 1144 and 1164.

MAA-5 5' ACGCATCGACTTTAGCAGTA 3'

Used for sequencing, binds to the sense strand of *Pf*CS between positions 660 and 680.

MAA-6 5' AGCTCTTCGAGATAGCTGAA 3'

Used for sequencing, binds to the sense strand of *Pf*CS between positions 945 and 965.

MAA-7 5' GGGAATTCCATATGAATACGGAAAAATACCTTGCT 3'

Used to add an Nde1 site to PYR::tp CS and *Pf*CS by PCR. Binds to the sense strand of *Pf*CS between positions 74 and 109.

MAA-8 5' CGGGGTACCAATGCTCATTTTATCTACCTCC 3'

Used to add a Kpn1 site to PYR::tp CS and *Pf*CS by PCR. Binds to the antisense strand of PfCS between positions 1243 AND 1212.

MAA-9 5' CTTCAACAGAACATATTGTT 3'

Used for sequencing, binds between positions 367 and 347 on the sense strand of pRec7Nde1 vector.

MAA-10 5' TTGCATGCCTGCAGGTCGAC 3'

Used for sequencing, binds to the antisense strand of the pRec7-Nde1 vector, between positions 264 and 284.

MAA-11 5' GCTAATTGGTATAGAATTAAG 3'

Used for sequencing, binds to the sense strand of *Pf*CS, between positions 497 and 518.

MAA-12 5' TGTAGAGGTGTACATATGCCAGAAACTG 3'

Used for adding an Nde1 site to TP::pyr CS. Binds to the sense strand of TP::pyr CS between positions 255 and 283.

MAA-13 5' TGCAGCCCGGGGGTACCGGATCCGAAAACAC 3'

Used for adding a Kpn1 site to TP::pyr CS. Binds to the antisense strand of TP::pyr CS between positions 1520 and 1508.

MAA-14 5' TGGGTAACGCCAGGGTTTTC 3'

Used for sequencing TP::pyr within the pUC19 vector. Binds to the sense strand of the pUC19 vector 30bp upstream of the EcoR1 site.

MAA-15 5' TTTTTTGGTACCCTATAATTCTATGGGTAG 3'

Used for adding a stop codon in front of the DNA sequence coding for the 2 terminal amino acids of *Pf*CS. Binds to the anti sense strand of *Pf*CS between positions 1210 and 1195.

MAA-16 5' TTTTTTGGTACCCTACTATCCTACATACTGCAACCTGGG 3'

Used for adding two stop codons in front of the DNA sequence coding for the terminal 13 amino acids of *Pf*CS. Binds to the antisense strand of PfCS between positions 1177 and 1157.

MAA-17 5' GACGATAGTGGAGCGATTCCAGTAACCCCTG 3'

Used to mutate the Asp113 residue to an alanine residue. Primer binds to the sense strand of *Pf*CS, between positions 413 and 444.

MAA-18 5' GACGATAGTGGATCCATTCCAGTAACCCCTG 3'

Used to mutate the Asp113 residue to a serine residue. Primer binds to the sense strand of *Pf*CS, between positions 413 and 444.

MAA-19 5' CCATGGATTGGGCCCTTTAAAGC 3'

Antisense primer used to make the PCR1 products (see section 5.3.1). Binds to the antisense strand of *Pf*CS between positions 759 and 737.

Appendix B

List of clones constructed and used during this project

pBAP 3004 *Tp*CS within pAlter-1 vector

pBAP 3104 *Pf*CS within pAlter-1 vector

pBAP 3005 *Tp*CS within pAlter-1 vector mutated to introduce an Apa1 site.

pBAP 3105 *Pf*CS within pAlter-1 vector mutated to introduce a BspE1 site.

pBAP 3006 TP::pyr CS within pAlter-1 vector.

pBAP 3106 PYR::tp CS within pAlter-1 vector.

pBAP3007 TP::pyr CS within pUC19.

pBAP3008 TP::pyr CS within pRec7-Nde1.

pBAP3110 PYR::tp CS within pRec7-Nde1.

pBAP3011 Pf-2 mutant within pRec7-Nde1.

pBAP3013 Pf-13 mutant within pRec7-Nde1.

pBAP3014 Asp113/ser mutant within pRec7-Nde1.

PBAP3015 Asp113/ala mutant within pAlter-1.

pBAP3031 *Pf*CS within pRec7-Nde1.

pBAP3040 Asp113/ser mutant (containing additional mutations) within pRec7-Nde1.

pBAP 3041 Asp113/ala mutant (containing additional mutations) within pRec7-Nde1.

REFERENCES

Akanuma, S., Yamagishi, A., Tanaka, N. & Oshima, T.

Protein Science. (1998) **7**, 698-705.

Akasako, A., Haruki, M., Oobatake, M., Kanaya, S.

J. Biol. Chem. (1997) **272**, 18686-18693

Alter, G. M., Casazza, J.P., Zhi, W., Nemeth, P., Srere, P.A., & Evans, C.T.

Biochemistry (1990) **29**, 7563-7571

Arfin, S. M. & Bradshaw, R.A.

Biochemistry (1988) **27**, 7979-7984

Argos, P., Rossmann, M. G., Grau, U.M., Zuber, H., Frank, G., & Tratschin, J. D.

Biochemistry, (1979) **18**, 5698-5703

Ausubel, F. M., Brent, R., Kingston, R.E, Moore, D. D., Seidman, J. G., Smith, J.A., (Editors). Current protocols in molecular biology. (1987) Wiley Interscience USA.

Athes, V., Degraeve, P., Cavaille-Lefebvre, D., Espeillac, S., Lemay, P. & Combes, D.

Biotech letts. (1997) **19**, 273-276

Bayer, E., Bauer, B., & Eggerer, H.

Febs Letters, (1981) **127**, 101-104

Bintrim, S.B., Donohue, T. J., Handelsman, J., Roberts, G. P., & Goodman, R.M.

Proc. Natl. Acad. Sci. U.S.A. (1997) **94**, 277-282

Bradford, M. M.

Anal. Biochem. (1976) **72**, 248-254

Britton, K. L., Baker, P.J., Borges, K.M.M., Engel, P.C., Pasquo, A., Rice, D.W., Robb, F., Scandurra, R., Stillman , T.J. & Yip, K.S.P.

Eur. J. Biochem. (1995) **229**, 688-695

Burley, S.K. & Petsko, G.A.

Science (1985) **229**, 23-28

Chen, L., Coutinho, P. M., Nikolov, Z. & Ford, C.
Protein Eng. (1995) **8**, 1049-1055

Ciulla, R. A., Burggraf, S., Stetter, K. O. & Roberts, M.F.
Applied and Environmental Microbiology (1994) **60**, 3660-3669

Cowan, D. A.
TIBTECH. (1996) **14**, 177-178

Danson, M.J
Adv. Microbiol. Physiol. (1988) **29**, 165-230

Danson, M. J., and Hough, D.W.
Trends In Microbiology (1998) **6**, 307-313

Darland, G., & Brock, D.T.
Science (1970) **170**, 1416-1418

Davies, G. J., Gamblin, S.J., Littlechild, J. A.& Watson, H.C.
Protein Struc. Funct. Genet. (1993) **15**, 283-289

Dixon, M. M., Nicholson, H., Shewchuk, L., Baase, W. A. & Mathews, B. W.
J. Mol. Biol. (1992) **227**, 917-933

Eidsness, M. K., Richie, K. A., Burden, A. E., Kurtz, D. M., Scott, R. A.
Biochemistry (1997) **36**, 10406-10413

Eisenthal, R. & Cornish-Bowden, A.
J. Biochem. (1974) **139**, 715-720

Erduran, I. & Kocabiyik, S.
Biochem. and Biophys. Res. Comm. (1998) **249**, 566-571

Evans, C. T., Owens, D. , Sumegi, B., Kispal, G., & Srere, P.A.
Biochemistry (1988) **27**, 4680-4686

Evans, C. T., Owens, D., Casazza, J. P., & Srere, P.A.
Biochem. Biophys. Res. Commun. (1989) **164**, 1437-1445

Evans, C. T., Kurz, L. C., Remington, S. J. & Srere, P.A.
Biochemistry (1996) **35**, 10661-10672

Eyring, H.
J. Chem Phys. (1935).**3**, 107-115.

Facchiano, A. M., Colonna, G. & Ragone, R.
Protein Eng. (1998) **11**, 753-760

Fersht, A. R. & Serrano, L.
Current Opinion in Structural Biology (1993) **3**, 75-83

Hamana, H. & Shinozawa, T.
J. Biochem. (1999) **125**, 109-114

Haney, P. J., Badger, J. H., Buldak, G. L., Reich, C. I., Woese, C. R & Olsen, G. L
Proc. Natl. Acad. Sci. U.S.A. (1999) **96**, 3578-3583

Henderson, P. J. F.
In Enzyme assays a practical approach. (Editors Eisenthal, R. & Danson, M. J. pp 277-313. Oxford University Press, Oxford, UK. (1992).

Hensel, R., Jakob, I, Scheer, H. & Lottspeich, F.
Biochem. Soc. Symph. (1993) **58**, 127-133

Hensel, R. & Konig, H.
FEMS Micro. Letts. (1988) 49, 75-79

Huber, H., & Stetter, K. O.
Journal of Biotechnology (1998) **64**, 39-52

Imada, K., Sato, M., Tanaka, N., Katsube, Y., Matsuura, Y. & Oshima, T.
J. Mol. Biol. (1991) **222,** 725-738

Ishikawa, K., Okumura, M., Katanayi, K., Kimura, S., Kanaya, S., Nakamura, H. & Morikawa, K.
J. Mol. Biol. (1993) **230**, 529-542

Ishikawa, K., Nakamura, Morikawa, K. & Kanaya, S.
Biochemistry (1993) **32**, 6171-6178

Jaenicke, R. & Bohm, G.
Current Opinion in structural biology (1998) **8**, 738-748

James, K. D., Russell, R. J. D., Parker, L., Daniel, R. M., Hough, D. W. & Danson, M. J.

FEMS Micro. Lett. (1994) **119**, 181-186

Johnson, D.B.

FEMS Microbial Ecology. (1998) **27**, 307-317

Jones , S., & Thornton, J.M.

Prog. Biophys. Mol. Biol. (1995) **63**, 31-69

Jones, T. A., Zou, J. Y., Cowan S. W. & Kjeldgaard, M.

Acta Cryst. (1991) **A47**, 110-119

Karpusas, M., Branchaud, B., & Remington, J.

Biochemistry (1990) **29**, 2213-2219

Kawamura, S., Kakuta, Y., Tanaka, I., Hikichi, K., Kuhara, S., Yamasaki, N. & Kimura, M.

Biochemistry (1996) **35**, 1195-1200

Kirino, H., Aoki, M., Aoshima, M., Hayashi, Y., Ohba, M., Yamagishi, A., Wakagi, T. & Oshima, T.

Koralev, S., Nayal, M., Barnes, W. M., Di Cera, E. & Waksman, G.

Proc. Natl. Acad. Sci. USA. (1995) **92**, 9264-9268

Korndorfer, I., Steipe, B., Huber, R., Tomschy, A. & Jaenicke, R.

J. Mol. Biol. (1995) **246**, 511-521

Eur. J. Biochem. (1994) **220**, 275-281

Krebs, H.A & Lowenstein, J.M

'Metabolic Pathways' (1960) 2nd ed., vol1, pp.129-203

Kuchner, O. & Arnold, F. H.

Trends Biotech. (1997) **15**, 523-530

Kurz, L. C., Drysdale, G.R, Riley, M. C., Evans, C. T & Srere, P.A.

Biochemistry. (1991) 31, 7908-7914

Landt, O., Grunert, H. P. &Hahn, U.

Gene (1990) **96**, 125-128

Laemmli, U. K.

Nature (1970) **227**, 680-685

Lebbink, J. H. G, Eggen, R. I. L., Geerling, A. C. M., Consalvi, C., Chiaraluce, R., Scandurra, R. & de Vos, W. M.

Protein Eng. (1995) **8**, 1287-1294

Lebbink, J. H. G., Knapp, S., Van der Oost, J., Rice, D., Ladenstein, R. & de Vos, W.M.

J. Mol Biol. (1998) **280**, 287-296

Leszczynski, J. F. & Rose, G. D.

Science (1986) **234**, 849-855

Liao, D. I., Karpusas, M., & Remington, S.J.

Biochemistry (1991) **30**, 6031-6036

Lumry, R. & Eyring, H.

J. Phys. Chem. (1954) **58**, 110-120

Mandelman, D., Schwarz, F. P., Li, H. Y. & Poulos, T.I.

Protein Science (1998) **7**, 2089-2098

Mansfeld, J., Vriend, G., Dijkstra, B. W., Veltman, O. R., Van den Burg, B., Venema, G., Ullbrich-Hofmann, R. & Eijsink, V. G. H.

J. Biol. Chem. (1997) **272**, 11152-11156

Martinez del Pozo, A., van Ophem, P. W., Ringe, D., Petsko, G., Soda, K.& Manning, J.M.

Biochemistry (1996) **35**, 2112-2116

Martins, L. O. & Santos, H.

Applied and Environmental Microbiology (1995) **61**, 3299-3303

Martins, L. O., Huber, R., Huber, H., Stetter, K. O., Da Costa, M. S. & Santos, H.

Applied and Environmental Microbiology (1997) **63**, 896-902

Mathews, B. W., Nicholson, H. & Becktel, W. J.

Proc. Natl. Acad. Sci. USA. (1987) **84**, 6663-6667

Matsumara, M., Signor, G. & Mathews, B. W.

Nature (1989) **342**, 291-293

M^{c}Cormack, M.
PhD Thesis, University of Bath, 1995

Meinnel, T., Lazennec, C., Dardel, F., Schmitter, J. M. & Blanquet, S.
FEBS Letters (1996) **385**, 91-95

Menendez-Arias, L. & Argos, P.
J. Mol. Biol. (1989) 206, 397-405

Michels, P.C. & Clark, D. S.
Applied and Environmental microbiology (1997) **63**, 3985-3991

Molgat, G. F., Donald,L.J. & Duckworth, H.W.
Archives of Biochemistry and Biophysics (1992) **298**, 238-246

Muir, J. M., Hough, D. W. & Danson, M. J.
System. Appl. Microbiol. (1994) **16**, 528-533

Muir, J. M., Russell, R. J. M., Hough, D.W., & Danson, M.J.
Protein Eng. (1995) **8**, 583-592

Munoz, V. & Serrano, L
Curr. Opin. Biotechnol. (1995) **6**, 382-386

Nicholson, H., Becktel, W.J. & Mathews, B. W.
Nature (1988) **336**, 651-656

Nojima, H., Ikai, A., Oshima, T. & Noda, H.
J. Mol. Biol. (1977) **116**, 429-442

Nojima, H., Honnami, K., Oshima, T. & Noda, H.
J. Mol. Biol. (1978) **122**, 33-42

Nosoh, Y., Sekiguchi, T.
TIBTECH. (1990) **8**, 16-20

Numata, K., Muro, M., Akutso, N., Nosoh, Y., Yamagishi, A. & Oshima, T.
Protein Eng. (1995) **8**, 39-43

Pace, C.N.
TIBS. (1990), **15**, 14-17

Pace, C.N.

J. Mol. Biol. (1992) **226**, 29-35-

Pace, C. N., Grimsley, G. R., Thomson, J. A. & Barnett, B. J.

J. Biol. Chem. (1988) **263**, 11820-11825

Pappenberger, G., Schurig, H. & Jaenicke, R.

J Mol Biol. (1997) **274**, 676-683

Pauptit, R.A., Karlsson, R., Picot, D., Jenkins, J. A., Niklaus-Reimer, A.S. &Jansonius, J.N.

J. Mol. Biol. (1988) **199**, 525-537

Pozo et al

Biochemistry (1996) **35**, 2112-2116

Purcarea, C., Herve, G., Ladjimi, M. M. & Cunin, R.

J. Bacteriol. (1997) **179**, 4143-4157

Querol, E., Perez-Pons, J.A. & Mozo-Villarias, A.

Protein Eng. (1996) **9**, 265-271.

Ramos, A., Raven, N. D. H., Sharp, R.J., Bartolucci, S., Rossi, M., Cannio, R., Lebbink, J., van der Oost, J., de Vos, W. M. & Santos, H.

Applied and Environmental Microbiology (1997) **63**, 4020-4025

Remington, S.J.

Curr. Top. Cell. Regul. (1992) **33**, 209-229

Remington, S.J, Weigand, G. & Huber, R.

J. Mol. Biol. (1982) **158**, 111-152

Rice, D. W., Yip, K. S. P., Stillman, T. J, Britton, K. L., Fuentes, A., Connerton, I., Pasquo, A, Scandurra, R., Engel, P.C.

FEMS Microbiology reviews (1996) **18**, 105-117

Robertson, J., Flory, D., Kumar, A., Desai, P. & Villafranca, J. J.

Biochemistry (1988) **27**, 3108

Rojkova, A. M., Galkin, A.G., Kulakova, L. B., Serov, A. E., Savitsky, P. A., Fedorchuk, V. V. & Tishkov, V. I.

FEBS Letters (1999) **445**, 183-188

Russell, R. J. M., Hough, D. W., Danson, M. J. & Taylor, G. L.
Structure (1994) **2**, 1157-1167

Russell, R. J. M., Ferguson, J. M. C., Hough, D. W., Danson, M., J. & Taylor, G. L.
Biochemistry (1997) **36**, 9983-9994

Russell, R. J. M., Gerike, U., Danson, M. J., Hough, D. W. & Taylor, G. L.
Structure (1998) **6**, 351-361

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989)
'Molecular Cloning A Laboratory Manual', Cold Spring Harbor Laboratory Press, Cold Spring Laboratory, New York.

Scopes, D. A., Bautista, J. M., Naylor, C. E., Adams, M. J. & Mason, P.J.
Eur. J. Biochem. (1998) **251**, 382-388

Segerer, A., Langworthy, T. A. & Stetter, K.O.
Systematic and applied microbiology (1988) **10**, 161-171

Service, R.F.
Science (1997) **275**, 1740-1742

Shih, P. & Kirsch, J.F.
Protein Sci. (1995) **4**, 2063-2072

Shirley, B.A., Stanssens, P., Hahn, U. & Pace, C.N.
Biochemistry (1992) **31**, 725-732

Singh, A. & Hiyashi, K.
J. Biol. Chem. (1995) **270**, 21928-21933

Srere, P.A., Brazil, H. & Gonen, L.
Acta Chem Scand (1963) **17,** S129-134

Sutherland, K. J., Henneke, C. M., Towner,. Hough, D. W. & Danson, M. J
Eur. J. Biochem. (1990) **194**, 839-844.

Tomazic, S. J. & Klibanov, A.M.
J. Biol. Chem. (1988) **263**, 3092-3096

Teplyakov, A. V., Kuranova, I.P., Harutyunyun, E.H. & Vainshtein, B. K.
J. Mol. Biol. (1990) **214**, 261-279

Usher, K. C., Delacruz, A. F. A., Dahlquist, F. W., Swanson, R. V., Simon, M. I. & Remington, S. J.
Protein Science (1998) **7**, 403-412

Van den Burg, B., Vriend, G., Veltman, O.R., Venema, G., & Eijsink, V.J.H.
Proc. Natl. Acad. Sci. USA. (1998) **95**, 2056-2060

Vetriani, C., Maeder, D.L., Toliday, N., Yip, K. S. P., Stillman, T.J., Britton., K.L., Rice, D.W., Klump, H.H., & Robb, F.T.
Proc. Natl. Acad. Sci. USA. (1998) **95**, 12300-12305

Vogt, G.,Woell, S. & Argos, P.
J. Mol Biol. (1997) **269**, 631-643

Volkin, D. B. & Klibanov, A. M.
J. Biol. Chem. (1987) **262**, 2945-2950

Volkin, D. B. & Middaugh, C.R.
Chemical and Physical Pathways of Protein Degradation, Edited by Ahern, T. J.& Manning, M. C. (1992) 215-247. Plenum Press. New York

Waldburger, C. D., Schildbach, J.F., & Sauer, R. T.
Nature Struc. Biol. (1995) **2**, 122-128

Wallon, G., Kryger, G., Lovett, S. T., Oshima, T., Ringe, D., & Petsko, G.A.
J. Mol. Biol. (1997) **266**, 1016-1031

Weigand, G., Remington, S.J., Deisenhofer, J., & Huber, R.
J. Mol. Biol. (1984) **174**, 205-219

Weitzman, P. D. J.
Adv. Microb. Physiol. (1981) **22**, 185-244

Weitzman, P. D. J., & Danson, M.J.
Curr. Topics Cell Regul. (1976) **10**, 161-204

Wetzel, R.
TIBS. (1987) **12**, 478-482

Wood, D.O., Atkinson, W. H., Sikorski, R. S, & Winkler, H.H

Journal of Bacteriology. (1983) **155**, 412-416

Woese, C. R., Kandler, O., & Wheelis, M. L.

Proc. Natl. Acad. Sci. USA (1990) **87**, 4576-4579

Yip, K. S. P., Stillman, T.J., Britton, K.L., Artymiuk, P.J., Baker, P.J., Sedelnikova, S.E., Engel, P.C., Pasquo, A., Chiaroluce, R., Consalvi, V., Scandurra, R. & Rice, D.W.

Structure **3**, (1995) 1147-1158

Yip, K. S. P., Britton, L., Stillman, J., Lebbink, J., DeVos, W. M., Robb, F. T., Vetriana, C., Maeder, D. & Rice D.W.

Eur. J. Biochem. (1998) **255**, 336-346

Zhukovsky, E.A., Mulkerrin, M.G. & Presta, L.G.

Biochemistry. (1994) **33**, 9856-9864