Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Structural & Functional Characterization of 3-Deoxy-D-*arabino*-Heptulosonate 7-Phosphate Synthase from *Helicobacter pylori* & *Mycobacterium tuberculosis*

Celia Jane Webby

2006

Structural & Functional Characterization of 3-Deoxy-D-*arabino*-Heptulosonate 7-Phosphate Synthase from *Helicobacter pylori* & *Mycobacterium tuberculosis*

A thesis presented in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Biochemistry

At Massey University, Turitea, Palmerston North,
New Zealand.

Celia Jane Webby

2006

ABSTRACT

The shikimate pathway, responsible for the biosynthesis of aromatic compounds, is found in microorganisms and plants but absent in higher organisms. This makes the enzymes of this pathway attractive as targets for the development of antibiotics and herbicides. Recent gene disruption studies have shown that the operation of the shikimate pathway is essential for the viability of *M. tuberculosis*, validating the choice of enzymes from this pathway as targets for the development of novel anti-TB drugs.

3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS) catalyzes the first committed step of the shikimate pathway. Two distinct classes of DAH7PS have been defined based on sequence similarity. The type I DAH7PSs are well characterized, however prior to this project there was limited mechanistic and no structural information about type II enzymes. Sequence identity between type I and type II enzymes is less than 10% raising the possibility that they represent distinct protein families, unrelated by evolution.

We have functionally characterized the type II enzyme from *Helicobacter pylori*, and have shown that type I and type II enzymes catalyze a metal-dependent ordered sequential reaction following the same stereochemical course. We have solved the structure of the type II DAH7PS from *M. tuberculosis* using single-wavelength anomalous diffraction (SAD) methods and the structure reveals a tightly associated dimer of $(\beta/\alpha)_8$ TIM barrels. The monomer fold, the arrangement of key residues in the active site, and the binding modes of PEP and Mn^{2+}, all match those of the type I enzymes. This similarity of protein fold and catalytic architecture makes it unequivocal that type I and type II enzymes are related by divergent evolution from a common ancestor. Interestingly, there are significant differences in the additional structural elements that extend from the core $(\beta/\alpha)_8$ barrel and in the quaternary structure. Further structural and functional analysis of *M. tuberculosis* DAH7PS revealed that the two major additions decorating the barrel are involved in the binding of the aromatic amino acids. Two distinct inhibitory binding sites for Trp and Phe have been identified providing an explanation for the synergistic inhibition displayed with Trp and Phe. The role of several active site residues of *Mt*-DAH7PS in enzyme catalysis has also been investigated.

ACKNOWLEDGEMENTS

I would like to thank Massey University and the Center for Molecular Biodiscovery (CMB) for the funding throughout my PhD project. I have thoroughly enjoyed my time at Massey University and the collaborations I have made with the people from CMB.

I would like to thank my chief supervisor Emily Parker for all the encouragement, advice and guidance she has given me throughout my PhD. She has shared with me her passion for research and has always been there in times of confusion, despair and of course excitement!

I would like to thank Mark Patchett for all of the advice when things were not going to plan. I would especially like to thank Geoff Jameson for all the help and the long and often 'brain-hurting' conversations about my research. Thank-you to Ted Baker, Heather Baker and Shaun Lott for the help in the structural work presented in this thesis. Thanks to Gill Norris for all the advice and the chocolate egg hunts at easter!

To the X-ray lab for all the help and entertainment throughout this project. Especially Trevor Loo who is the person I would often go to when I needed to know or find something in the lab

To the crystallography guys in Palmy and Auckland, especially Bryan Anderson, Richard Bunker and Dave Goldstone for helping me get to grips with crystallography.

To the Shikimate group for all the good times and especially Linley Schofield for all the advice and the world's greatest chocolate cakes!

Finally I would like to thank my friends and family for all their support throughout my time at university. Especially mum and dad who have always encouraged me to do whatever makes me happy, and Ben for putting up with all the bad moods when things were not going to plan.

TABLE OF CONTENTS

CHAPTER TWO

Expression, Solubilization and Biochemical Characterization of Type II DAH7PS from *H. pylori*

CHAPTER THREE

Structural and Functional Characterization of the Type II DAH7PS from *M. tuberculosis*

CHAPTER FOUR

Feedback Regulation of Type II DAH7PSs from *M. tuberculosis* and *H. pylori*

CHAPTER FIVE

CHAPTER SIX SUMMARY OF THESIS

CHAPTER SEVEN EXPERIMENTAL PROCEDURES

ABBREVIATIONS

AEC	Anion exchange chromatography
Amp	Ampicillin
A5P	D-Arabinose 5-phosphate
BTP	1,3-*bis*(tris(hydroxymethyl)methylamino)propane
CCP4	Collaborative Computational Project number 4
CEC	Cation exchange chromatogrpahy
Da	Dalton
DAH7P	3-deoxy-D-*arabino*-heptulosonate 7-phosphate
DNA	Deoxy-ribose nucleic Acid
dNTP	Deoxyribo nucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediamine tetra-acetic acid (di-sodium salt)
E4P	D-Erythrose 4-phosphate
FPLC	Fast protein liquid chromatography
G3P	D-Glyceraldehyde 3-phosphate
G6P	D-Glucose 6-phosphate
HCl	Hydrochloric acid
IEX	Ion exchange chromatography
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl-1-thio-β-D-galactopyranoside
k_{cat}	Turnover number
KDO8P	3-Deoxy-D-*manno*-octulosonate 8-phosphate
K_2HPO_4	di-Potassium hydrogen orthophosphate
K_M	Michaelis constant
LB broth	Luria-Bertani broth
MW	Molecular weight
MWCO	Molecular weight cut-off
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NMR	Nuclear magnetic resonance
OD	Optical density
ORF	Open reading frame

PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein data bank
PEP	Phosphoenolpyruvate
Phe	Phenylalanine
pI	Isoelectric point
P_i	Inorganic phosphate
ppm	Parts per million
Psi	Pounds per square inch
R5P	D-Ribose 5-phosphate
2dR5P	2deoxyR5P
Rpm	Revolutions per minute
SAD	Single wavelength anomalous dispersion
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
Se-Met	Selenomethionine
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TEV	Tobacco etch virus
Thesit	Polyethyleneglycol dodecyl ether
TIM	Triose phosphate isomerase
Trp	Tryptophan
Tyr	Tyrosine
UV	Ultraviolet
V_{max}	Maximum reaction velocity

INDEX OF FIGURES

INDEX OF TABLES

CHAPTER ONE INTRODUCTION

1.1 The Shikimate Pathway

The shikimate pathway is a series of seven enzyme catalyzed reactions responsible for the biosynthesis of chorismate (figure 1.1), a precursor for many essential aromatic compounds, including the aromatic amino acids, vitamins E and K, folic acid and various quinones. [1,2] In *Mycobacteria* chorismate is a precursor for salicylate, used in the biosynthesis of siderophores (iron-chelating compounds) through which bacteria acquire iron.[3,4]

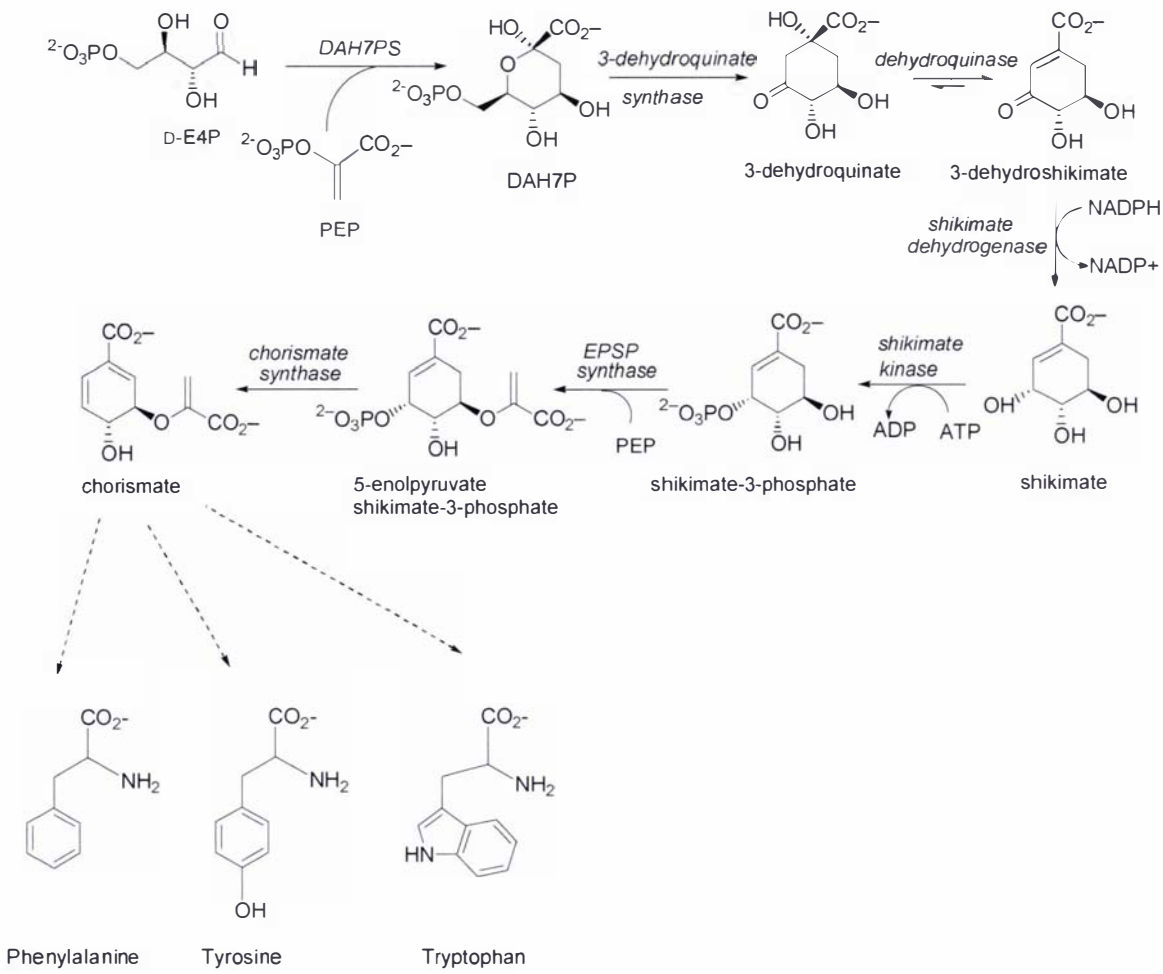


Figure 1.1 The seven enzyme-catalyzed reactions collectively known as the shikimate pathway

The shikimate pathway is responsible for up to 30 % of the total carbon flux within the cell reflecting its importance in controlling cellular carbon flux.[5] The shikimate pathway has attracted significant interest over the past thirty years as this pathway is present in plants, fungi and bacteria, but not mammals. This makes the enzymes of this pathway attractive as targets for the development of novel herbicides, fungicides and antibiotics. The shikimate pathway is the target for the successful broadspectrum herbicide glyphosate (*N*-[phosphonomethyl]-glycine), which inhibits the sixth enzyme of the pathway, 5-enolpyruvylshikimate 3-phosphate (EPSP) synthase.[6,7]

The discovery of the shikimate pathway

Shikimic acid was the first intermediate of this pathway to be identified and predictably led to the name of the pathway. Shikimic acid was first isolated in 1885 by Eykmann from the plant *Illicium religiosum*,[8] and the name was derived from the Japanese name of this plant, shikimi-no-ki.[9] Approximately fifty years later Freudenberg and Fischer established the correct structure and stereochemistry of shikimic acid.[10-13] However, it was not until the 1950s when Davis *et al* showed that a culture of an *Escherichia coli* mutant-strain, requiring aromatic amino acids for growth, accumulated shikimic acid.[14] It was this experiment that revealed shikimic acid as an intermediate of aromatic biosynthesis.[14,15] Similar experiments have shown that 5-dehydroquinic acid[16] and 5-dehydroshikimic acid[17] are successive precursors of shikimic acid. The rest of the intermediates of the pathway were isolated and identified from mutants of *E. coli* and *Aerobacter aerogenes*: chorismic acid,[18] 5-enolpyruvylshikimic acid 3-phosphate,[19] shikimic acid 3-phosphate,[20] 2-dehydroshikimic acid[17] and 3-dehydroquinic acid.[16]

Isotope tracer incorporation studies using ^{14}C-D-glucose indicated that shikimic acid is derived from a three-carbon intermediate of glycolysis and a four-carbon sugar from the pentose phosphate pathway.[21] These were later identified as D-erythrose 4-phosphate (E4P) and phosphoenolpyruvate (PEP) by the conversion of these two sugars into shikimic acid in extracts of *E. coli*.[22]

Organization of the enzymes of the pathway

The seven enzymes required for the conversion of PEP and E4P to chorismate are organized differently in various organisms. In *E. coli* the seven reactions are catalyzed

by discrete enzymes, all independently transcribed and widely scattered across the genome.[23,24] In *Saccharomyces cerevisiae*,[25] and fungi such as *Aspergillus nidulans*[26] and *Neurospora crassa*,[27] the second through to the sixth reaction of the shikimate pathway are catalyzed by a pentafunctional polypeptide referred to as the 'Arom complex', encoded by a single structural gene.[1] In *Bacillus subtilis* 3-deoxy D-*arabino* heptulosonate 7-phosphate synthase (DAH7PS) is expressed as a bifunctional enzyme possessing both DAH7PS and chorismate mutase (CM) activity (described in further detail in Section 1.3.1). There has been a long-standing debate as to whether the CM activity is genuine, as earlier studies by Llewellyn *et al* suggested DAH7PS was monofunctional.[28] More recent studies by Woodard *et al* have provided significant evidence for the bifunctional enzyme and have cleared up any confusion by showing *B. subtilis* possesses a second independent monofunctional CM.[29] In the case of the *B. subtilis* bifunctional enzyme there is debate as to whether dual activity arose from gene fusion[30] or less likely, through mutation of the allosteric binding site of a pre-existing monofunctional DAH7PS.[28] The fusion of the two enzyme activities is believed to be for the purpose of feedback-inhibition (refer to Section 1.3.1 for further detail).

In higher plants dehydroquinase-shikimate dehydrogenase activities are present as a bifunctional polypeptide and like the other multifunctional polypeptides the individual activities are catalytically independent.[31] It has been suggested that the plant bifunctional enzyme and the pentafunctional *arom* enzyme evolved by the fusion of individual genes similar to those in *E. coli*.[32] The amino acid sequence of the pentafunctional *arom* enzymes from *S. cerevisiae* and *A. nidulans* show homology with the five corresponding *E. coli* enzymes.[31]

The shikimate pathway in plants

Evidence for the operation of the shikimate pathway in higher plants has been obtained mainly from isotope tracer experiments. These have shown that not only do the pathways of biosynthesis of phenylalanine (Phe), tryptophan (Trp) and tyrosine (Tyr) involve the same intermediates as in bacteria and fungi, but that higher plants can also convert the aromatic amino acids into natural products or secondary metabolites.[9,33]

In plants the shikimate pathway enzymes are synthesized as precursors in the cytosol and then targeted to the chloroplast by a signal sequence which is subsequently cleaved

to produce mature proteins.[34] There is evidence for a separate cytosolic pathway which may be involved in secondary metabolite biosynthesis.[35]

1.2 3-Deoxy D-*arabino*-Heptulosonate 7-Phosphate Synthase (DAH7PS)

3-Deoxy D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS) (EC 2.5.1.54) catalyzes the first committed step of the shikimate pathway (figure 1.2). This is an aldol-like condensation reaction between PEP and E4P to generate a seven carbon sugar, DAH7P, with the release of inorganic phosphate. PEP is a product of the glycolytic pathway and E4P is derived from the pentose phosphate pathway.[9] One of the objectives of this project is to exploit the similarities and differences between DAH7PSs from various organisms to aid in the future design of type-specific inhibitors of this enzyme.

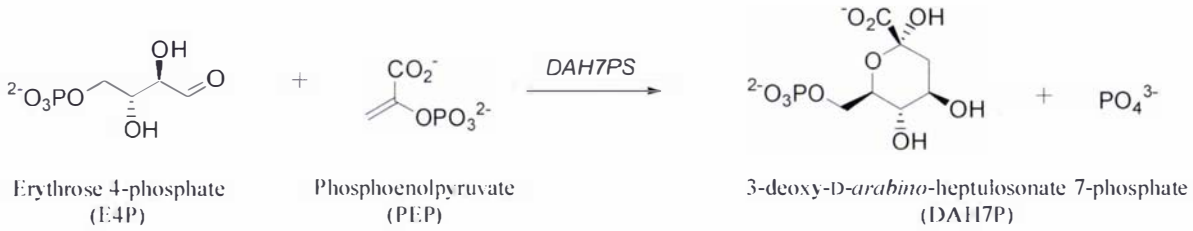


Figure 1.2 Reaction catalyzed by DAH7PS

Classification of DAH7PS

DAH7PSs can be clustered into two distinct homology families based on amino acid sequence and molecular weight.[36,37] These have been denoted type I and type II by Walker *et al*[37] and more recently they have been referred to as $AroA_I$ and $AroA_{II}$ by Gosset *et al.*[38] Type I DAH7PSs are smaller than their type II counterparts, with molecular masses less than 40 kDa. Type I DAH7PSs can be further divided on the basis of sequence similarity into two subfamilies: Iα and Iβ.[36,39] DAH7PSs of the Iβ subfamily are more closely related to 3-deoxy D-*manno*-octulosonate 8-phosphate synthases (KDO8PS) (EC 2.5.1.55) than to members of subfamily Iα.[39] KDO8PS catalyzes an analogous reaction to DAH7PS involving the condensation of PEP with D-

arabinose 5-phosphate (A5P) to form KDO8P. KDO8PS is involved in lipopolysaccharide biosynthesis in Gram-negative bacteria.[40]

Type II DAH7PSs are larger than the type I enzymes (around 54 kDa), and were originally identified in higher plants where they were believed to play an additional role in secondary metabolism.[9,33] The type II enzymes were thought to include a limited number of microbial proteins,[37] however as more microbial genomes have been sequenced it has been shown that this type of DAH7PS consists of a subset of plant enzymes clustered within a more divergent set of microbial enzymes.[38] In some organisms, such as *Amycolatopsis mediterranei*, *Amycolatopsis methanolica*, *Xanthomonas campestris*, *Pseudomonas aeruginosa* and *Stigmatella aurantiaca*, both type I and type II DAH7PSs have been identified.[41,42] Several of these type II enzymes have been shown to be required for the biosynthesis of specific secondary metabolites.[42,43] However, the presence of only type II DAH7P synthases in a number of species, including *Streptomyces*, *Corynebacterium diphtheriae*, *Campylobacter jejuni*, *Agrobacterium tumefaciens*, *Novosphingobium aromaticivorans*, *Helicobacter pylori* and several *Mycobacteria*, is consistent with type II DAH7PSs functioning in aromatic amino acid biosynthesis.

Sequence identity between the type I and type II DAH7PSs is less than 10 %, raising the possibility that they represent distinct protein families, unrelated by evolution. Jensen *et al* suggested that structural information about type II DAH7PSs may reveal the two types arose from divergent evolution with the order and spacing of active site residues essentially superimposable.[36] Interestingly, the third enzyme of the shikimate pathway, dehydroquinase (DHQase), similarly exists as two distinct protein families that are unrelated at the sequence level. The two types of DHQase exhibit altered catalytic mechanism and are distinctly different in structure indicating that despite common overall reaction chemistry, the two groups arose via convergent evolution.[44,45]

1.3 Regulation of Microbial DAH7PSs

1.3.1 Feedback-inhibition

Regulation of DAH7PS is essential for maintaining control of cellular levels of aromatic compounds in microorganisms and plants. ^{13}C-NMR studies using whole cells of *E.*

coli have shown that feedback-inhibition of DAH7PS is the main mechanism for controlling carbon flow into the shikimate pathway.[46] Different species contain from one to three isoenzymes which are inhibited by feedback control, by end-products or intermediates of the pathway, or a combination of the two. In some cases, the enzymes may be unregulated by feedback-inhibition.[37,47]

There are three isozymes of DAH7PS in *E. coli* designated DAH7PS(Tyr), DAH7PS(Phe) and DAH7PS(Trp). These proteins are encoded by the genes *aroF*, *aroG*, and *aroH* respectively, which are found scattered over the *E. coli* genome.[34] The isozymes are each subject to feedback-inhibition by their respective aromatic amino acid. DAH7PS(Phe) and DAH7PS(Tyr) account for approximately 80 % and 20 % of the total DAH7PS activity in *E. coli*, with the Trp-sensitive enzyme contributing less than 1 %.[48] *Salmonella typhimurium* and *N. crassa* also possess three isoenzymes of DAH7PS, whereas in *S. cerevisiae* there are only two isoenzymes regulated by Phe and Tyr.[49]

Slightly different patterns of control have been adopted by other organisms. In *B. subtilis* a single DAH7PS is feedback regulated through an N-terminal fusion to CM.[29,30] This is further supported by the presence of a DAH7PS-CM fusion protein in *Porphyromonas gingivalis*, in which a CM domain is fused to the C-terminus of the DAH7PS domain.[30] In both these cases CM provides the regulatory domain and is inhibited by chorismate and prephenate. The truncated DAH7PS domain is able to catalyze the condensation reaction but its activity is insensitive to the presence of feedback inhibitors. The fusion of DAH7PS and CM is believed to be for the purpose of feedback-regulation and not for the creation of a true bi-functional enzyme. It has been suggested that "feedback-regulation" may be the evolutionary link between Iα and Iβ enzymes.[30]

Dual activity of a DAH7PS and CM has been reported in *A. methanolica*.[41,50] Biochemical studies have shown that *A. methanolica* possesses two DAH7PSs, a type II and a type Iα, and a single CM enzyme. The type II DAH7PS is sensitive to feedback-inhibition by all three aromatic amino acids, whereas the type Iα enzyme is only sensitive to Tyr feedback-inhibition. The type II DAH7PS is found in complex with CM, rendering CM sensitive to feedback-inhibition by Phe and Tyr. Separation of CM

from the type II enzyme, using chromatographic techniques, yields a CM protein insensitive to feedback-inhibition.

All DAH7PSs with their crystal structures solved share a basic $(\beta/\alpha)_8$ barrel monomer-fold. However, there are significant differences in the additional structural elements that decorate the core barrel. All except the *Pyrococcus furiosus* enzyme (*Pf*-DAH7PS), which is unregulated, possess extra structural motifs (domains or subdomains) that are implicated in allosteric regulation. The type Iβ DAH7PS from *Thermotoga maritima* (*Tm*-DAH7PS), which is very similar in basic structure to *Pf*-DAH7PS, has a ferredoxin-like domain attached to the N-terminus of the barrel.[51] In contrast to *Pf*-DAH7PS, *Tm*-DAH7PS is regulated by Phe and Tyr, and this inhibition depends on the presence of this N-terminal domain.[51] The crystal structures of the type Iα Phe-regulated *E. coli* DAH7PS (*Ec*-DAH7PS(Phe)) and the Tyr-regulated *S. cerevisiae* DAH7PS (*Sc*-DAH7PS(Tyr)) in complex with Phe and Tyr, respectively, have been solved. Both amino acids bind in very similar locations in both crystal structures with the N-terminal region (from a neighboring subunit) and extended α5-β6 loop associating to form a single Phe (or Tyr) binding site.[49,52-54] At present there are no known crystal structures for *Ec*-DAH7PS(Tyr) and *Ec*-DAH7PS(Trp). However, excision of the seven N-terminal residues of *Ec*-DAH7PS(Tyr) generates a Tyr-resistant enzyme, strongly suggesting that the N-terminus of *Ec*-DAH7PS(Tyr) is involved in the feedback-inhibition mechanism.[55] Equilibrium binding studies of *Ec*-DAH7PS(Trp) with radiolabelled Trp have indicated the presence of two independent Trp binding sites per *Ec*-DAH7PS(Trp) dimer.[56] The exact location of the regulatory binding sites of *Ec*-DAH7PS(Tyr) and *Ec*-DAH7PS(Trp), and the mechanism of inhibition, await structural information.

Mutagenesis studies with the two isoenzymes from *S. cerevisiae* have shown that a single mutation of a conserved glycine (Gly226) (conserved in Tyr-regulated DAH7PSs) to a serine (conserved in Phe-regulated DAH7PSs) and vice versa, leads to a complete change in the regulation pattern without affecting enzyme kinetics.[49] The crystal structure of the *Sc*-DAH7PS(Tyr) G226S mutant in complex with Phe has been determined (structure unpublished but deposited in the Protein Data Bank). The comparison of this structure with the wild-type *S. cerevisiae* structure (in complex with Tyr) shows that the G226S mutation significantly reduces the size of the inhibitory

binding cavity favoring Phe-binding over Tyr (figure 1.3). It should be noted that Ser226 is built in as an alanine in all subunits of the G226S structure, as seen in figure 1.3. The superposition of the G226S structure onto the Phe-regulated *E. coli* enzyme (in complex with Phenylalanine, PDB code 1KFL), allows one to estimate where the serine sidechain should be placed.

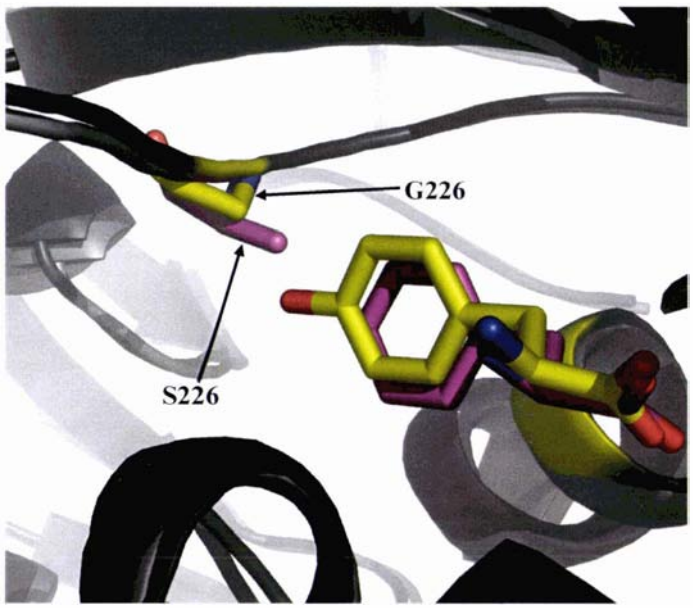


Figure 1.3 Overlay of *Sc*-DAH7PS(Tyr) G226S mutant in complex with Phe (PDB code 1OG0) (unpublished) and wild-type *Sc*-DAH7PS(Tyr) in complex with Tyr [49] (PDB code 1OF6). The Tyr molecule and Gly226 of *Sc*-DAH7PS(Tyr) are in yellow. The Phe molecule and residue Ser226 of the G226S structure are in pink. Note: the sidechain of S226 has been built in as an alanine in all subunits of the G226S mutant structure.

1.3.2 Transcription Level Control

The three DAH7PS isoenzymes from *E. coli* are also regulated, to a lesser extent, through repression at the transcriptional level. The expression of *aroF* and *aroG* is repressed by the Tyrosine repressor, *TyrR* gene product,[57,58] in complex with either Tyr or Phe, respectively.[48]

The expression of *aroH* is under the control of the Trp repressor, product of the *TrpR* gene, and Trp.[59] When intracellular levels of Trp are low, the Trp repressor exists predominately in the aporepressor form that has low affinity for operator DNA.[60] If the intracellular levels of Trp are elevated the Trp aporepressor binds its corepressor, Trp,

which subsequently binds the respective operator DNAs to repress Trp, *TrpR* and *aroH* expression.[59] The *Trp* repressor is believed to bind at the regulatory region upstream of the *aroH* gene,[59,61] and prevent transcription of the DAH7PS(Trp) gene by interfering with RNA polymerase binding.[62]

In *S. cerevisiae* the biosynthesis of aromatic amino acids is part of a complex regulatory network which regulates the derepression of over thirty structural genes involved in multiple amino acid biosynthetic pathways under amino acid starvation.[63] The final step in this general control system is the binding of GCN4 to the promoters of the target genes resulting in elevated transcription. A GCN4-binding site has been located upstream of both *S. cerevisiae* DAH7PS isoenzymes.[64,65] The expression of the amino acid biosynthetic genes is never completely shut down, allowing a basal level of transcription of these genes to be maintained at all times. In the case of *Sc*-DAH7PS(Phe) and *Sc*-DAH7PS(Tyr), the basal transcription of only the Phe-regulated enzyme is enough to maintain levels of aromatic amino acids in the cell.[66]

1.4 Regulation of Plant DAH7PSs

The regulation of plant DAH7PSs is somewhat different to that observed for bacterial DAH7PSs. Studies have shown plant DAH7PSs are regulated at the transcriptional level, although exactly how still remains unclear. None of the DAH7PSs that have been characterized to date appear to be sensitive to feedback-regulation by any of the aromatic amino acids.[67] In some cases aromatic amino acids have been reported to enhance enzyme activity, for example Trp has been found to activate DAH7PS from carrot *(Daucus carota)*[68] and potato (*Solanum tuberosum*). The differences in regulation between bacterial and plant enzymes may reflect the vital role DAH7PSs have in the synthesis of aromatic secondary metabolites in plants.[42,43]

Intracellular metabolites have been shown to regulate the levels of DAH7PS synthesis in plants. Exposure of potato cells to high levels of glyphosate, a specific inhibitor of EPSP synthase, induces an increase in DAH7PS activity *in vivo*. This increase in activity is not from the direct interaction between DAH7PS and glyphosate as *in vitro* studies with the herbicide have no effect on DAH7PS activity.[69] The nature of the signal remains unknown.

1.5 Properties of Type Iα DAH7PSs

1.5.1 Biochemical Studies

Molecular mass

The most extensively investigated DAH7PS is the Phe-regulated isoenzyme from *E. coli*. The three DAH7PS enzymes from *E. coli* have polypeptide chains of similar length with DAH7PS(Phe) having 350 residues, DAH7PS(Trp) 348 residues and DAH7PS(Tyr) 356 residues. An alignment of the amino acid sequences indicate that they are homologs, arising from gene duplications and divergent evolution of a common ancestral gene.[70] The proteins share 41 % identity in their amino acid sequences and another 15 % represent conservative substitutions.[71,72] The identical residues are located in central regions of the sequences whereas less similarity is found in the C- (46 identical residues of 157) and N-terminus (6 of 41).[73]

Ec-DAH7PS(Phe) is a tetramer in solution with a monomer mass of 38010 Da,[74] whereas *Ec*-DAH7PS(Tyr) and *Ec*-DAH7PS(Trp) are both dimeric, with subunit masses of 38804 Da and 38719 Da, respectively.[75,76] *Sc*-DAH7PS(Phe) and *Sc*-DAH7PS(Tyr) have been reported to be monomeric[66] and dimeric[65] in solution, with molecular masses of 41070 Da and 39749 Da, respectively. It should be noted that there is inconsistency between the quaternary structure observed for *Sc*-DAH7PS(Tyr) in solution compared with the crystal-form as the protein crystallizes as a tetramer (refer to Section 1.5.2 for further details). There is 76 % sequence similarity between the two *S. cerevisiae* isoenzymes and 66-73 % similarity between *Sc*-DAH7PS(Tyr) and the three *E. coli* DAH7PSs.[66]

Substrate specificity

The most extensive substrate specificity study of a type Iα enzyme has been performed with *Ec*-DAH7PS(Phe). In addition to its natural substrate, E4P, this enzyme also shows limited enzymatic activity with a range of five-carbon monosaccharides, including A5P, 2-deoxy-D-ribose 5-phosphate (2dR5P), and D-ribose 5-phosphate (R5P).[77] It has also been shown that DAH7PS from *Neisseria gonorrhoeae* can utilize A5P as a substrate as well as E4P.[78] More recent studies by members of our research group have found that 2-deoxyE4P (2dE4P), where the C2 hydroxyl group of E4P is removed, is a relatively poor substrate for *Ec*-DAH7PS(Phe).[79] D-Threose 4-phosphate

(T4P), which has the opposite stereochemistry at the C2 position of E4P, is also utilized by the *E. coli* enzyme.[80]

Metal activation

All DAH7PSs, that have been characterized to date, require a divalent metal ion for catalysis.[65,81,82] Earlier work showed that some enzyme preparations were inhibited by chelating agents and activated by divalent metals,[83,84] while others were unaffected by chelation or additional metal ions.[75,85] Recent studies by Woordard *et al* have shown that the type of chelating agent and the incubation time with the enzyme are crucial in determining the metal requirement of various DAH7PS enzymes. The majority of earlier work used EDTA as a chelating agent, however this has been shown to be ineffective in removing bound metal in *B. subtilis* DAH7PS. Woodard and co-workers found that a stronger metal chelator, dipicolinic acid (DPA), was required to remove enzyme-bound metal (refer to Section 1.6.1.1 for further details).[29]

The *in vivo* metal ion of *Ec*-DAH7PS(Phe) is still unknown. Earlier studies suggested Co^{2+}, as a significant increase in enzyme activity was observed in the presence of this divalent metal.[83] However, work by Simpson and co-workers report the *E. coli* enzyme to be insensitive to Co^{2+}, with purified enzyme containing less than 0.02 mol of Co^{2+} per mole of monomer.[85] More recent work has indicated that the *E. coli* DAH7PSs are Fe^{2+} or Zn^{2+} enzymes, with the isolated isoenzymes containing 0.2-0.3 moles of Fe^{2+} per mole of monomer.[81] The three *E. coli* enzymes were stripped of activity by treatment with EDTA and then reactivated with a variety of metal ions, giving the following order of activity: $Mn^{2+} > Cd^{2+} > Fe^{2+} > Co^{2+} > Ni^{2+} > Cu^{2+} > Zn^{2+} >> Ca^{2+}$. Although Mn^{2+} showed highest activity this divalent metal was found to be easily displaced from the enzyme by other metal ions. In contrast, Fe^{2+} and Zn^{2+} were not easily displaced from the enzyme indicating a high affinity and stable enzyme-metal complex had been formed. From these results Stephens proposed Fe^{2+} or Zn^{2+} as the preferred metal cofactor *in vivo*. Other evidence has favored Cu^{2+} as the *in vivo* metal for *Ec*-DAH7PS(Tyr).[82] Baasov and co-workers found 0.5 moles of Cu^{2+} per mole of enzyme subunit, in their native enzyme preparation. The Cu^{2+} was removed by dialysis with cyanide ion reducing the catalytic activity to 6 % of the native enzyme. The enzyme was then reactivated by Cu^{2+} with a stoichiometry of one Cu^{2+} per enzyme monomer. Zn^{2+} was also able to reactivate the enzyme, however reactivation was not seen with Fe^{2+}.

Studies with *Sc*-DAH7PS(Phe) proposed that Fe^{2+} is the *in vivo* metal for this DAH7PS with 0.6 moles of Fe^{2+} per enzyme monomer found in an enzyme preparation.[86] However, there is considerable variation in the ability of different metals to activate the *E. coli* and *S. cerevisiae* DAH7PSs. *Sc*-DAH7PS(Phe) showed a metal reactivation series as follows: $Co^{2+}>Mn^{2+}>Fe^{2+}>Zn^{2+}$ while, Ca^{2+}, Mg^{2+}, Cu^{2+} and K^{2+} were unable to restore activity.[86] Interestingly, *Sc*-DAH7PS(Tyr) has been shown to have a slightly different reactivation series of: $Co^{2+} > Zn^{2+}$, Cu^{2+}, $Fe^{2+} >> Cd^{2+}$, $Ni^{2+} >> Mn^{2+}$, Mg^{2+}.[65]

Ec-DAH7PS(Phe) was found to be prone to metal-catalyzed oxidation in the absence of PEP. Cu^{2+} and Fe^{2+} were the only divalent metals that greatly increased the rate of inactivation and subunit dissociation. Spontaneous inactivation of the enzyme is associated with the loss of two thiols, which can be completely restored upon the addition of dithiothreitol (DTT).[87] Mutation of either of the two active-site cysteines, Cys61 and Cys328, resulted in enzyme that was insensitive to the metal attack, as judged by the lack of subunit dissociation upon Cu^{2+} treatment. Peptide mapping of the inactive enzyme revealed a disulfide linkage between these two cysteine residues.[87] It has long been known that PEP stabilizes DAH7PS during purification and storage, however, E4P has the opposite effect and increases the rate of spontaneous inactivation through a Schiff-base formation between the aldehyde functionality of E4P and an active site lysine.[88] Site-directed mutagenesis, chemical modification, spectroscopic analysis and structural studies have shown that the invariant cysteine and histidine residues, Cys61[89] and His268[90] of *Ec*-DAH7PS(Phe), are involved in metal binding and are essential for enzyme catalysis.

1.5.2 Structural Characterization

Monomer

The crystal structures of *Ec*-DAH7PS(Phe)[71] and *Sc*-DAH7PS(Tyr)[53] have been determined. Both proteins fold as a $(\beta/\alpha)_8$-barrel decorated with an N-terminal extension (strand β0 followed by helices α00 and α0), and an extra β6a/β6b sheet inserted before the β6 strand of the barrel[71] (figure 1.4). In both enzymes there is an α-helix that closes off the N-terminal end of the barrel. The structures are very similar (rmsd 0.6 Å for 329 equivalent Cα atoms[53]) despite the fact that they originate from different organisms and possess different patterns of regulation.

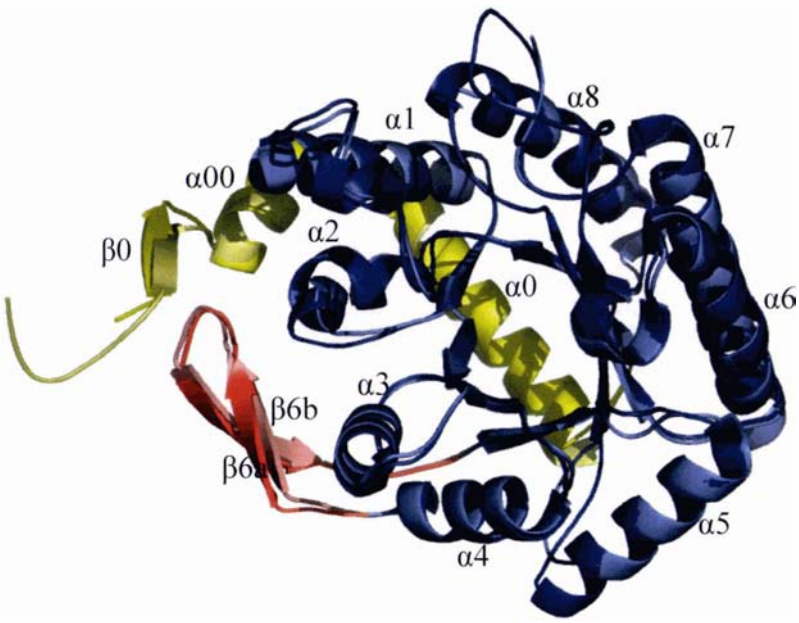


Figure 1.4 ***E. coli* (PDB code 1KFL) and *S. cerevisiae* (PDB code 1HFB) DAH7PSs superimposed, with α-helices of the basic $(\beta/\alpha)_8$ barrel labeled.** Extra structural elements are colored yellow for the N-terminal extension, and red for the respective inhibitor binding site.

Quaternary Structure

The crystal structures of both *Ec*-DAH7PS(Phe) and *Sc*-DAH7PS(Tyr) consist of two tight dimers making up a tetramer (figure 1.5). Within the dimers, the N-terminus of one monomer (e.g A) entwines the second monomer (e.g B) and forms an intermolecular β-sheet between β0 of one monomer and β6b from the neighbouring subunit. Although one of the dimers is essentially superimposable between the *E. coli* and *S. cerevisiae* enzymes, the association of the two dimers into a tetramer is significantly different. In *S. cerevisiae* the tetrameric association is purely hydrophobic, whereas in *E. coli* contact is through electrostatic and hydrogen bonding interactions. In both enzymes, the contributing residues are located at the N-terminal extension and the additional two-stranded antiparallel β-sheet, β6a and β6b strands.[53,91] It is still not understood whether the difference in quaternary structure between the type Iα enzymes corresponds to differences between species or whether it is related to the different regulatory behaviours of the Tyr- and Phe- regulated enzymes. Further studies are required to determine whether there is any communication between the two tight dimers that make up the tetramer.

It is possible that the conserved dimer found in both type Iα enzymes is a close structural analogue of the other two dimeric *E. coli* isoforms. Both *Ec*-DAH7PS(Tyr) and *Ec*-DAH7PS(Trp) have substitutions that eliminate the polar interactions contributing to the *Ec*-DAH7PS(Phe) tetramer. The mutation of a key residue (Glu24), speculated to be essential for tetrameric association of *Ec*-DAH7PS(Phe), generates a dimeric protein (in solution) with similar catalytic efficiency as the wild-type protein. It should be noted that the *Ec*-DAH7PS(Phe)-E24Q protein still crystallizes as a tetramer.[92]

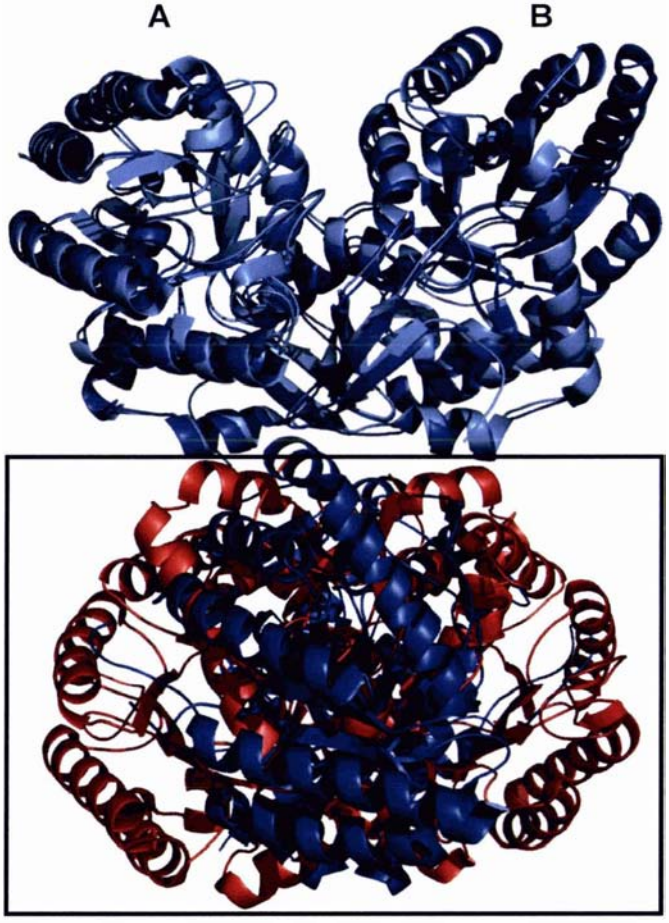


Figure 1.5 **Comparison of quaternary structures of *Ec*-DAH7PS(Phe) (PDB code 1KFL) and *Sc*-DAH7PS(Tyr) (PDB code 1HFB).** The shared dimer (A and B) is in blue, and the tetrameric association (the second dimer) of the type Iα enzymes is shown in red for *Ec*-DAH7PS(Phe) and blue for *Sc*-DAH7PS(Tyr).

Active site architecture

As is typical for $(\beta/\alpha)_8$ barrels, the active site is located at the C-terminal end of the barrel, and the residues making up the active site belong to the C-terminal ends of the β-strands and β-α loops. The β-α loops located at the C-terminal side of the barrel are more flexible than the α-β loops at the N-terminal side of the barrel. Co-crystallisation with PEP or metal ion is essential for definition of the β-α loops in the crystal structure.[53] The active sites of *Ec*-DAH7PS(Phe) and *Sc*-DAH7PS(Tyr) are very similar in terms of arrangement and conservation of protein residues involved in metal and substrate binding.

(a) PEP and metal-binding sites

In both structurally characterized type Iα DAH7PSs the metal ion is found in a trigonal bi-pyramidal coordination (figure 1.6), with a cysteine (Cys76 and Cys61 in *S. cerevisiae* and *E. coli*, respectively) and histidine (His282 and His268) occupying the axial positions and a glutamate (Glu316 and Glu302) and aspartate (Asp342 and Asp326) as the equatorial ligands. The third equatorial position is occupied by a water molecule that hydrogen bonds to a lysine (residue 112 and 97), which coordinates the carboxylate group of PEP.[53,71] The crystal structures of *Ec*-DAH7PS(Phe) and *Sc*-DAH7PS(Tyr) in complex with various metal ions and ligands have been determined.[71] The first crystal structure solved of *Ec*-DAH7PS(Phe) was in complex with PEP and Pb^{2+}. Pb^{2+} activates *Ec*-DAH7PS(Phe) to only 3 % of the activity obtained with Mn^{2+},[71] however this metal was necessary to provide adequate phase information for structure determination. This structure did not appear to give an accurate picture of where PEP is bound in a complex competent for catalysis. A higher resolution structure of *Ec*-DAH7PS(Phe) in complex with 2-phosphoglycolate (PGL), an unreactive substrate analogue of PEP, and Mn^{2+} has also been determined.[93] PEP is proposed to bind in a similar orientation to that observed for PGL in the *Ec*-DAH7PS(Phe)-Mn^{2+}-PGL structure, where the carboxylate end of PGL coordinates Mn^{2+}. More recently, higher resolution crystal structures of *Sc*-DAH7PS(Tyr) and *Ec*-DAH7PS(Phe)-E24Q (described in Section 1.5.2) in complex with PEP suggest no interaction between the metal and PEP.[53,92] Comparison of the crystal structures of *Sc*-DAH7PS(Tyr) in complex with Mn^{2+} and Co^{2+} have shown that the binding mode of PEP is essentially identical with respect to each metal, which is in contrast to that observed with Pb^{2+}.

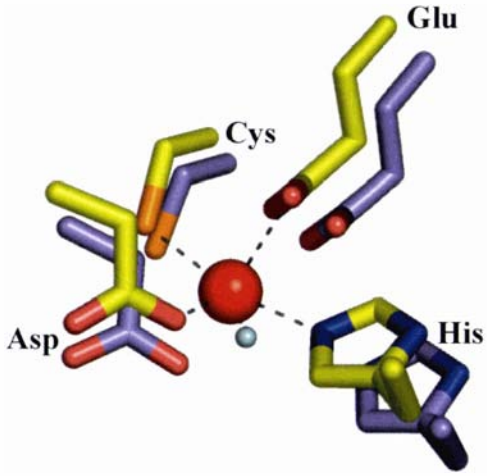


Figure 1.6 **Superposition of *Ec*-DAH7PS(Phe) (PDB code 1QR7) (yellow) and *Sc*-DAH7PS(Tyr) (PDB code 1OF6) (blue) showing respective metal binding sites**

Based on the more recent crystal structures of *Sc*-DAH7PS(Tyr) and *Ec*-DAH7PS(Phe)-E24Q, the PEP binding site can be defined by a network of hydrogen bonds between the protein and the PEP phosphate and carboxylate groups[53,92] (figure 1.7). The interactions between PEP and the protein are essentially identical for both type Iα structures. PEP is positioned in the active site by five arginine and lysine residues. Both type Iα structures indicate a deviation from planarity in the geometry of PEP, due to a 20-30° twist of the carboxylate plane relative to the enol plane. This structural distortion of PEP is speculated to be a vital step in the reaction mechanism of DAH7PS,[92] described in Section 1.9.

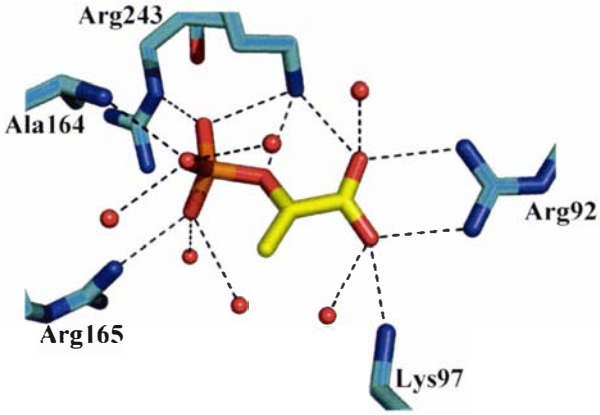


Figure 1.7 **PEP binding site of *Ec*-DAH7PS(Phe)-E24Q (PDB code 1N8F).** Dashed lines indicate interactions between PEP and surrounding water molecules and protein residues.

(b) E4P binding site

The structure of *Sc*-DAH7PS(Tyr) in complex with PEP, Co^{2+} and the three-carbon E4P analogue, glycerol 3-phosphate (figure 1.8) has been solved.[53] The phosphate moiety of G3P coincides with a bound sulfate found in all three *E. coli* structures, inferred to occupy the position of the phosphate moiety of E4P. The C2-hydroxyl group of glycerol 3-phosphate (corresponding to the C3-hydroxyl group of E4P) interacts with the metal-binding aspartate (Asp342) leading to a movement of the aspartate sidechain with respect to the metal ion. The hydroxyl group of C1 of G3P (corresponding to the C2-hydroxyl group of E4P) hydrogen bonds to the carbonyl oxygen of Pro113 (figure 1.9). E4P has been modeled into the active site of *Sc*-DAH7PS(Tyr) based on the assumption that E4P binding is essentially identical to G3P with an additional carbonyl group.[53] E4P is positioned so that C1 of E4P is 2.7Å away from the C3 atom of PEP. The oxygen atom of the carbonyl group of E4P replaces a metal coordinating water seen in both the *Ec*-DAH7PS-Pb^{2+}-PEP and the *Sc*-DAH7PS-Co^{2+}/Mn^{2+}-PEP structures. The carbonyl group of E4P coordinates the metal ion with a distance of 2.4 Å, allowing the carbonyl group to be activated by the metal ion and to be positioned correctly for catalysis. The carbonyl oxygen of E4P is also hydrogen bonded to a lysine, conserved in all DAH7PSs. It has been proposed that this lysine protonates the carbonyl group of E4P to generate the new hydroxyl group of DAH7P. The mechanism of DAH7PS is discussed in more detail in Section 1.9.

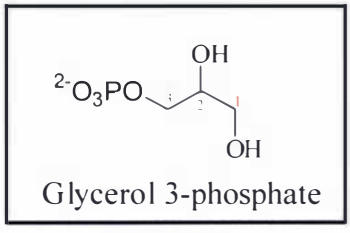


Figure 1.8

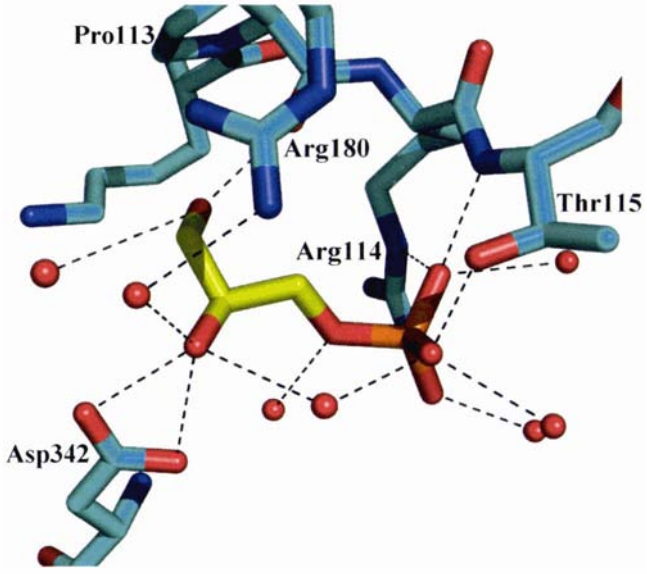


Figure 1.9 Binding of G3P in the active site of *Sc*-DAH7PS(Tyr)-Co^{2+}-PEP-G3P (PDB code 1OF8). Dashed lines indicate the interactions between G3P, and surrounding water molecules and protein residues.

1.6 Properties of Type Iβ Enzymes - DAH7PSs and KDO8PSs

There is now an extensive amount of structural and functional information on three subfamily Iβ DAH7PSs (**1**) from *P. furiosus*, *T. maritima* and *B. subtilis* (*Bs*-DAH7PS), and two subfamily Iβ KDO8PSs (**2**) from *Aquifex aeolicus* (*Aa*-KDO8PS), and *E. coli* (*Ec*-KDO8PS).

1.6.1 DAH7PSs

1.6.1.1 Biochemical Studies

Molecular mass

Size exclusion chromatography has indicated that *Pf*-DAH7PS is a dimer, whereas DAH7PS from *T. maritima* and *B. subtilis* are both tetrameric in solution. The three enzymes have monomer molecular masses of 29.2 kDa,[94] 38.0 kDa[95] and 40.0 kDa,[29] respectively. The *P. furiosus* dimer has been reported to be resistant to denaturation by SDS.

Substrate specificity

Pf-DAH7PS is able to utilize five carbon sugars, A5P, R5P and 2dR5P, as well as four carbon sugars, E4P, 2dE4P and T4P. The *P. furiosus* enzyme does not take three and

six carbon sugars.[47] DAH7PSs from *B. subtilis* and *T. maritima* appear to be more selective and do not use A5P or R5P as alternative substrates.[29,95]

Metal activation

Just like their type Iα counter-parts, type Iβ DAH7PSs are metalloenzymes. Metal reactivation studies with various divalent metal ions have been performed with the three DAH7PSs from *P. furiosus*, *T. maritima*, and *B. subtilis*. The reactivation series of *Pf*-DAH7PS and *Tm*-DAH7PS are very similar with Cd^{2+}, Mn^{2+} and Zn^{2+} having the greatest effect on activity. It was originally proposed that DAH7PS from *B. subtilis* (*Bs*-DAH7PS) was a non-metalloenzyme,[96] however more recent studies by Woodard *et al* have shown that the enzyme can be inactivated by DPA and then reactivated by Cd^{2+} and Zn^{2+}.[29] The confusion regarding the metal-dependence of *Bs*-DAH7PS has come about as EDTA, the most commonly used metal-chelator in DAH7PS studies, is not able to completely remove all enzyme-bound metal. It was not until Woodard and co-workers used a stronger metal-chelator, DPA, that metal from the active site was completely removed and enzyme activity was reduced to 1 % of that of the untreated enzyme.[29]

1.6.1.2 Structural Characterization

Monomer

Type Iα and Iβ DAH7PSs share a common monomer-fold, however there are significant differences in the structural elements that decorate the core $(\beta/\alpha)_8$ barrel. The *Pf*-DAH7PS is the smallest DAH7PS, with its monomer comprising of just the core $(\beta/\alpha)_8$ barrel[47] (figure 1.10 (a)). The *Tm*-DAH7PS adopts a similar basic structure to *Pf*-DAH7PS with an additional ferredoxin-like domain attached to the N-terminus of the barrel, which is involved in regulation[51,95] (figure 1.10). The structure of a third Iβ enzyme, from *Aeropyrum pernix*, has been solved which has a very similar monomer-fold and quaternary structure to *Pf*-DAH7PS (PDB code 1VS1, structure yet to be published). In all three of the type Iβ enzymes there is a β-ribbon, in place of the α-helix that is found in type Iα DAH7PSs, that closes off the N-terminal end of the barrel.

The monomer-fold of the *Pf*-DAH7PS and *Tm*-DAH7PS is more similar to that of the closely related enzyme, KDO8PS than to the two type Iα DAH7PSs. The superposition of *Tm*-DAH7PS (barrel domain only) and *Aa*-KDO8PS gives an rms difference of 1.3 Å for 232 Cα atoms, while superimposing *Tm*-DAH7PS (barrel domain only) with *Sc*-DAH7PS gives an r.m.s.d of 1.7Å for 221 Cα atoms.[51]

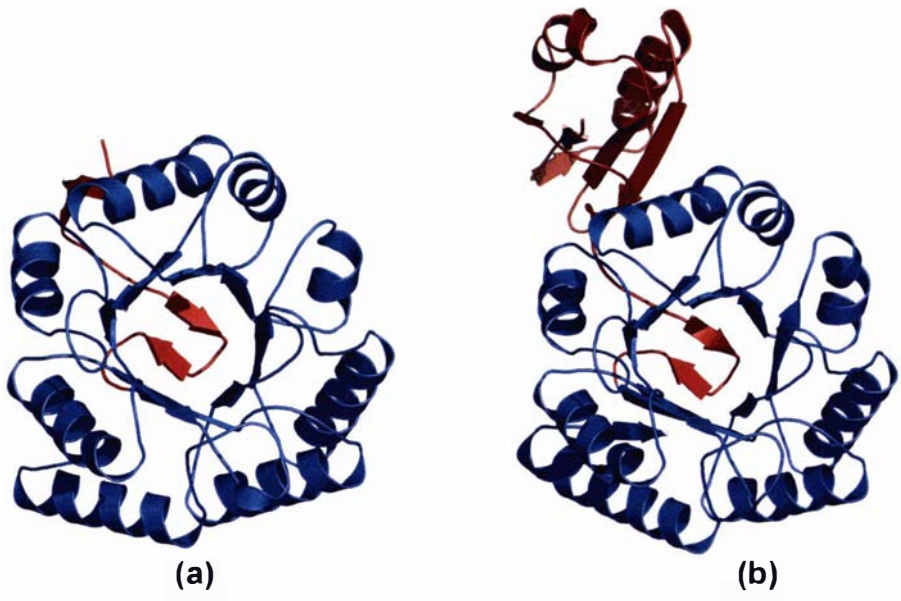


Figure 1.10 Monomer structure of (a) *Pf*-DAH7PS and (b) *Tm*-DAH7PS.

Quaternary structure

Although *Pf*-DAH7PS is a dimer in solution, the protein crystallizes as a tetramer. A homotetramer has also been reported for all other type Iα and Iβ enzymes with known crystal structures. However, only one monomer-monomer interface (monomer A and monomer B) is conserved between the two subfamilies. Despite this common dimer, the type Iα and Iβ enzymes form tetramers using completely different structural elements,[47,51,53,71] implying that while the dimers can be superimposed, the tetramers cannot (figure 1.11). The unconserved dimer interface (interface between subunit A and C, B and D is shared by *Pf*-DAH7PS and *Tm*-DAH7PS and is highly conserved (20 out of 27 residues). It has been speculated that the ancestral DAH7PS is a type Iβ enzyme comprised of this unique dimer.[47]

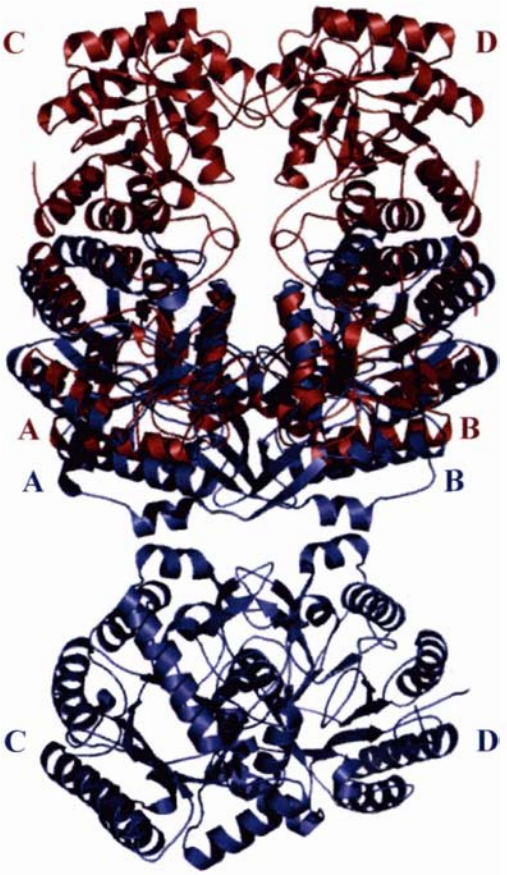


Figure 1.11 Comparison of quaternary structures of type Iα and type Iβ enzymes. Superposition of type Iα *Ec*-DAH7PS(Phe) (PDB code 1GG1), in blue, onto type Iβ *Pf*-DAH7PS (PDB code 1ZCO), in firebrick.

Active site architecture

(a) Metal and PEP binding sites

All residues that interact with the divalent metal ion and PEP are conserved in type Iα and type Iβ DAH7PSs. The only difference is that an alanine, conserved in type Iα enzymes, is substituted by a glutamine in type Iβ enzymes. This replacement allows an additional interaction to be formed between the carboxylate group of PEP and the protein.[47,51]

(b) E4P binding site

The crystal structure of *Tm*-DAH7PS was solved in complex with PEP, E4P and Cd^{2+}, i.e. containing all reaction components.[51] The presence of both PEP and E4P was observed due to the slow catalysis of the hyperthermophilic enzyme at low temperatures.

The structure showed that PEP and E4P were too far apart to allow bond formation. The orientation and conformation of PEP appeared to be appropriate for the reaction, and was consistent with what had been previously observed in other crystal structures of DAH7PS. E4P, on the other hand, appeared to be bound less tightly and the high *B*-factors of its atoms suggest some degree of rotational freedom. In the active site of *Tm*-DAH7PS E4P is positioned incorrectly for the reaction to occur with the correct stereochemistry (see Section 1.9). It appears likely that conformational changes, either in E4P or the residues interacting with the phosphorylated monosaccharide, are necessary to bring the two substrates closer together for catalysis to occur. E4P has been modeled into the active site of *Tm*-DAH7PS and is consistent with that observed for *Sc*-DAH7PS(Tyr) (described in Section 1.5.2).

1.6.2 KDO8PS

1.6.2.1 Biochemical Studies

KDO8PSs can be divided into metallo- and nonmetallo- enzymes, the best characterized being *A. aeolicus* and *E. coli*, respectively. Type 1β DAH7PSs exhibit more overall similarity to KDO8PSs than they do to the 1α DAH7PSs,[36] and it has been proposed that KDO8PSs evolved from a DAH7PS-like ancestor.[30,47] The KDO8PSs from *A. aeolicus*, *Chlamydia psittaci*, and *H. pylori* have been shown to be metalloenzymes.[36,97,98] Structural information and primary sequence analysis of these KDO8PSs indicate the presence of all four "established" metal-binding residues found in DAH7PSs (refer to Section 1.5.2). In contrast, KDO8PSs from *E. coli*, *S. typhimurium*, *N. gonorrhoeae*, and *Pisum sativum* have been reported to be non-metalloenzymes.[36] Recent studies have revealed that one of the key differences between metal-dependent and metal-independent KDO8PSs is the substitution of the metal-binding cysteine to an asparagine. The only known plant KDO8PS to be functionally characterized is from *Arabidopsis thaliana* (*At*-KDO8PS).[99] This enzyme cannot be inactivated by chelating agents, nor can it be activated by a series of divalent metal ions. Analysis of primary sequence information indicates the presence of an asparagine in place of a cysteine which is absolutely conserved in metal-dependent enzymes. The combination of this information strongly suggests that this plant enzyme is metal-independent, which is in contrast to what was originally predicted by Birck *et al* before the availability of sequence information.[97]

Molecular mass

All KDO8PSs that have been characterized to date, regardless of whether they are metal-dependent, have similar monomer molecular masses of ~30 kDa.[100-102] Size exclusion chromatography has indicated a trimeric or tetrameric quaternary structure for microbial KDO8PSs,[100-102] and a dimeric[99] association for *At*-KDO8PS, in solution. The crystal structures of *E. coli* and *A. aeolicus* KDO8PSs have been determined and both proteins crystallized as a tetramer.[103,104]

Substrate specificity

The natural substrate of KDO8PS is the five carbon sugar, A5P. This phosphorylated monosaccharide has the opposite stereochemistry at the C2 position as well as an extra carbon in comparison with DAH7PS's natural substrate, E4P. Substrate specificity studies have been performed on *E. coli* KDO8PS (*Ec*-KDO8PS). The ability of the *E. coli* enzyme to accept 4-deoxy-arabinose 5-phosphate as a substrate[105] is significant as it indicates that this enzyme can accept an acyclic form of A5P. In contrast to E4P, A5P is a cyclic structure in its most abundant form. Unlike DAH7PSs, which show relative insensitivity to changes at the C2 position of the monosaccharide (refer to Sections 1.5.1 and 1.6.1.1), the correct stereochemistry at the C2 position appears to be essential for the KDO8PS reaction. KDO8PS from *E. coli* and *A. aeolicus* are unable to utilize R5P, which has the opposite stereochemistry of A5P at the C2 position.[104,106] Interestingly, 2dR5P has been reported to act as an alternative, yet very poor substrate for *Ec*-KDO8PS.[106,107]

The two KDO8PSs that have been functionally characterized from *N. gonorrhoeae* and *E. coli* are unable to utilise four carbon monosaccharides, as E4P is not a substrate.[106,108] Further evidence comes from recent studies that show *Ec*-KDO8PS is unable to accept T4P (E4P analogue with A5P-like stereochemistry at the C2 position of the phosphorylated sugar) as an alternative substrate.[80] This is yet another difference between KDO8PSs and DAH7PSs, where DAH7PSs are able to utilize four and five carbon phosphorylated sugars (Sections 1.5.1 and 1.6.1.1), whereas KDO8PSs are only able to use five carbon monosaccharides.

Metal activation

At present, the only KDO8PSs to be characterized as metal-dependent are from *A. aeolicus*,[109] *Aquifex pyrophilus*,[110] *C. psittaci*,[97] and *H. pylori*.[111] Upon treatment with various metal-chelators all four KDO8PSs show significantly reduced activity, while the inclusion of divalent metal ions restores enzyme activity, in some cases exceeding the original activity.

Recent studies have shown that a metal-dependent KDO8PS can be converted into a metal-independent enzyme via a single amino acid change; the substitution of the metal-binding cysteine for an asparagine, C11N in *A. aeolicus*[112] and *A. pyrophilus* KDO8PSs.[113] Both C11N mutants do not bind metals and their activity is unaffected by treatment with chelating agents or divalent metal ions. The reciprocal mutation has been attempted by several groups who are in disagreement as to whether the replacement of Asn26 for a cysteine residue in the metal-independent *Ec*-KDO8PS 'restores' metal dependency. Studies by Shulami *et al* and Oliynyk *et al* provide evidence that suggests that the N26C mutant was able to bind metal, and activity was reduced upon treatment with EDTA, and activated by the inclusion of several divalent metal ions.[113,114] In contrast, work by Woodard and co-workers propose that the non-metallo KDO8PS cannot be converted to a metal-binding enzyme through this single amino acid substitution.[112] However, all authors agree that the role of the metal in *A. aeolicus* and *A. pyrophilus* KDO8PSs is similar to the role of Asn26 in *Ec*-KDO8PS. It has been suggested that the metal ion has more of a structural role in maintaining the correct orientation of substrates and is not directly involved in catalysis (refer to Section 1.9 for further detail).[107,113]

1.6.2.2 Structural Characterization

Monomer

The crystal structures of the metal-independent *Ec*-KDO8PS[103] and the metal-dependent *Aa*-KDO8PS[115] have been determined. As discussed in the previous section, the monomer-fold of KDO8PS (figure 1.12) is very similar to that of the two type Iβ DAH7PSs, with *Pf*-DAH7PS being the most closely related DAH7PS to KDO8PS yet characterized. Like *Pf*-DAH7PS, both KDO8PSs lack additional structural elements that extend from the $(\beta/\alpha)_8$-barrel and both enzymes appear to be unregulated. The

topology of the $(\beta/\alpha)_8$ barrel is very similar between the two KDO8PSs with a total of 236 Cα atoms of *Ec*-KDO8PS and *Aa*-KDO8PS superimposing with an rmsd value of 1.0 Å.[51] The only significant difference between the two KDO8PSs is that *Ec*-KDO8PS possesses a two-stranded β-hairpin that covers the N-terminal end of the barrel, similar to that observed in *Pf*-DAH7PS and *Tm*-DAH7PS (refer to Section 1.6.1.2).

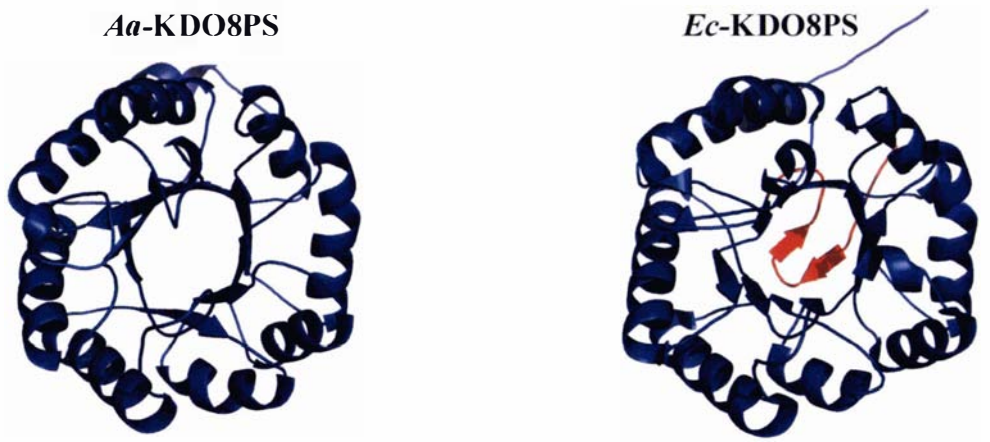


Figure 1.12 **Comparison of monomer-fold of *Aa*-KDO8PS (PDB code 1FWS) and *Ec*-KDO8PS (PDB code 1X8F).** The additional two-stranded β-hairpin that covers the N-terminal end of the barrel in *Ec*-KDO8PS is in red.

Quaternary Structure

KDO8PSs share the common dimeric arrangement of $(\beta/\alpha)_8$ barrel subunits that is observed in type Iβ DAH7PSs and type Iα DAH7PSs. The two KDO8PSs then associate into a tetramer very similar to that of the two type Iβ DAH7PSs.[47] The only significant difference between the two type Iβ enzymes is that in KDO8PS the conserved dimer interface appears to be more tightly packed, burying more surface area than that observed in the type Iβ DAH7PSs.[47]

Active Site

Several crystal structures of metal-free and Cd^{2+} forms of *Aa*-KDO8PS have been determined in complex with A5P alone and in combination with PEP.[115] Interestingly, it was observed that in the presence of metal, A5P binds to only one of the two active sites of the dimer contained in the asymmetric unit of the crystal. Using crystal symmetry to generate the tetramer it can be shown that the binding of A5P occurs at the active sites located on only one of the two faces of the enzyme. In the metal-free

enzyme, A5P binds to both active sites of the dimer. As the crystal structures were obtained by soaking the crystals in the presence of substrate/s, one would expect equal binding to the two faces of the enzyme, if the faces were equivalent. The observation that only one face binds A5P strongly suggests that catalysis cannot occur in both faces simultaneously. This hypothesis is consistent with previous pre-steady-state kinetics of phosphate release by *Ec*-KDO8PS, which indicated that only ~50 % of the active sites contribute to the release.[116]

Metal binding sphere

The metal-binding site of the metallo-KDO8PS from *A. aeolicus* (figure 1.13) is similar to that observed in the metal-dependent DAH7PSs. In the metal-independent *Ec*-KDO8PS the equivalent part of the structure is filled with an asparagine (Asn26) side chain which replaces the metal-binding cysteine (Cys11 in *Aa*-KDO8PS) in metal-dependent enzymes. The other three established metal-binding residues in metalloenzymes (His185, Asp233 and Glu222 in *Aa*-KDO8PS) are also conserved in *Ec*-KDO8PS (His202, Asp250 and Glu239). The only difference is that the β8-α8 loop bearing the aspartate ligand has moved significantly (figure 1.13). This is most likely due a proline two residues downstream from this aspartate which is found conserved in all nonmetallo-KDO8PSs.[47] It should be noted that this aspartate is disordered in all of the *Ec*-KDO8PS structures.

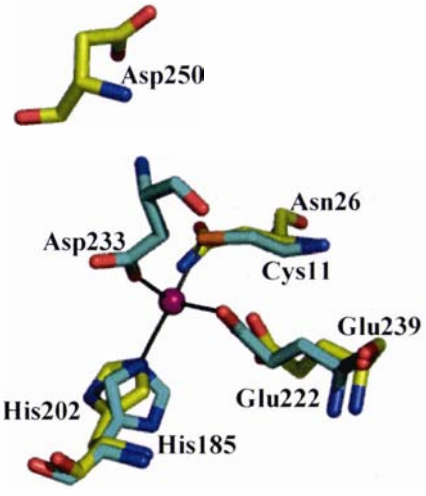


Figure 1.13 Superposition of metal binding sphere of *Aa*-KDO8PS (cyan) (PDB code 1FWW) onto the equivalent region of *Ec*-KDO8PS (yellow) (PDB code 1G7V).

PEP binding site

The majority of residues that interact with PEP in type Iβ DAH7PS and KDO8PS are similar, however several differences have been observed. An arginine (Arg115 in *Pf*-DAH7PS) that interacts with the phosphate moiety of PEP in DAH7PS is replaced by a phenylalanine (Phe117 in *Ec*-KDO8PS) in KDO8PS. A second arginine, found in DAH7PS (Arg55 in *Pf*-DAH7PS), that interacts with the carboxylate moiety of PEP is substituted for a lysine (Lys55 in *Ec*-KDO8PS) in KDO8PS.[51]

A5P binding site

There are many differences between the binding of A5P and E4P in KDO8PS and DAH7PS, respectively. The binding of A5P in the *Aa*-KDO8PS structure involves more interactions with protein residues than that observed for E4P in *Tm*-DAH7PS. There are several crystal structures of *Aa*-KDO8PS in complex with Cd^{2+} and A5P, however two different conformations of the phopshorylated sugar have been observed (figure 1.14).[115] In one conformation the C2-OH of A5P is coordinated to Cd^{2+}, and in the second conformation the C2-OH and C3-OH interact with a water molecule that coordinates Cd^{2+}. The interaction between the metal and the phosphorylated sugar in KDO8PS appears to be significantly different to that proposed for DAH7PS, where coordination to the metal is through the aldehyde moiety of the monosaccharide (Section 1.5.2).[53]

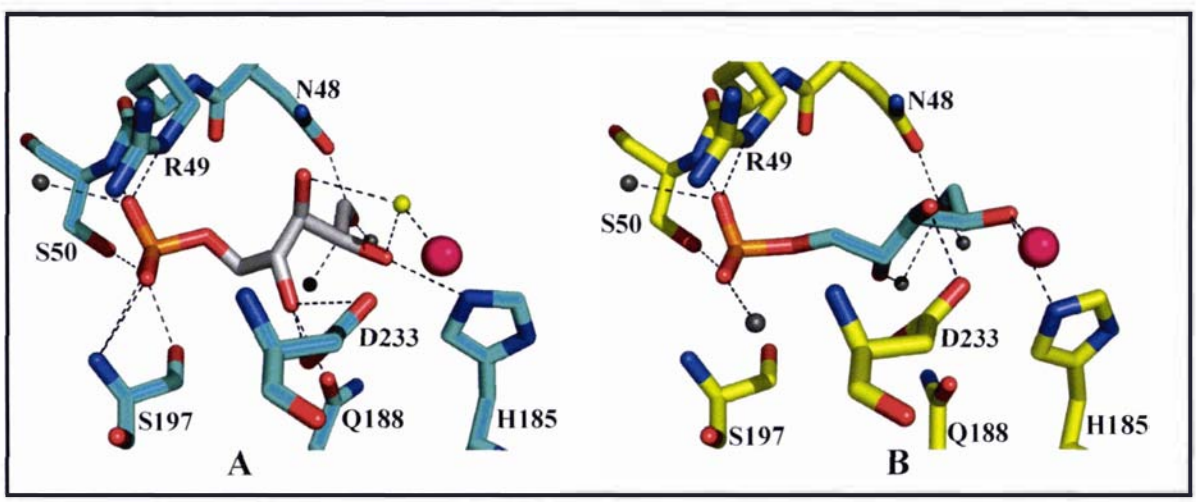


Figure 1.14 The two different conformations of A5P in *Aa*-KDO8PS. (A) shows coordination to Cd^{2+} (pink) through a water molecule (PDB code 1FWW), and (B) shows coordination of C2-OH of A5P to the Cd^{2+} ion (PDB code 1FY6).

1.7 Properties of Microbial Type II DAH7PSs

Type II DAH7PSs were first identified in higher plants, and it has only been in the last ten years that type II enzymes have been identified in microbial species.[37] Microbial type II enzymes share less than 10 % amino acid sequence identity with type I enzymes and show relatively high identity with plant DAH7PSs. The recent availability of sequence information has shown that type II DAH7PSs consist of a subset of plant enzymes clustered within a more divergent set of microbial enzymes.[38]

It was originally believed that type II DAH7PSs were involved solely in the biosynthesis of secondary metabolites,[38] however the presence of only type II DAH7PSs in the predicted proteomes of a number of species including *Streptomyces* species, *C. diphtheriae*, *C. jejuni*, *A. tumefaciens*, *N. aromaticivorans*, *H. pylori* and several *Mycobacteria* species provides evidence for a role of type II DAH7PSs in aromatic amino acid biosynthesis.

To date there has been limited characterization of type II DAH7PSs. The only microbial type II DAH7PSs reported in literature are from *Streptomyces*,[37,117,118] *N. crassa*[37,119,120] and *X. campestris*.[38] Microbial type II enzymes were examined from *Streptomyces* and in *N. crassa* prior to the availability of sequence information and classification. Recently Gosset *et al.* partially purified and characterized the type II DAH7PS from *X. campestris*.[38] There is limited functional information and no structural data about microbial type II DAH7PSs. It has been suggested that tertiary structural information about type II enzymes may provide evidence for divergent evolution of this class of DAH7PS.[36]

Molecular masses

The monomer molecular mass of type II DAH7PSs from the microbial organisms, that have been characterized, ranges from 52-59 kDa.[37,38,117-119] Several *Streptomyces* enzymes and the Trp-regulated *N. crassa* enzyme have been reported to be dimeric in solution,[37,117] while *Streptomyces caespitosus* appears to be a monomer.[118]

Metal activation

All microbial type II DAH7PSs that have been characterized to date are metal-dependent enzymes.[37,38,118,121] Metal-activation studies have been performed on

DAH7PSs from *X. campestris* and *S. caespitosus*, and both enzymes showed the highest activity with Co^{2+}, Zn^{2+} and Mn^{2+}.[38,118] While there appear to be differences between type I and type II enzymes in the ability of different metals to activate the enzyme, the variations do not appear to be type-specific as considerable differences are seen within enzymes of the type I family.

1.8 Properties of Plant DAH7PSs

Early studies suggested that plants possess two distinct DAH7PS activities, differentiated both by their intracellular location and their activation by either Co^{2+} or Mn^{2+}.[122] The Mn^{2+}-activated enzyme is located in the chloroplast and its activity is stimulated by Mn^{2+}, while the Co^{2+}-activated enzyme has been identified in the cytosol of some plant species and is activated by Co^{2+}, Mg^{2+} or Mn^{2+}.[123] Several forms of the Mn^{2+}-activated DAH7PS have been identified, which can be divided into two types, *shkA* and *shkB*. The individual proteins encoded by *shkA* and *shkB* are more similar across different species than the two *shkA* and *shkB* enzymes from a single plant.[124] The Co^{2+}-activated enzyme in *Spinacia oleracea*, and *S. tuberosum* has been shown to have broad substrate specificity with respect to its monosaccharide substrate, which has led people to question whether this is a true DAH7PS.[124,125] The lack of sequence information has made it difficult to determine how closely related this enzyme is to the known Mn^{2+}-activated DAH7PS.

Comparison of the plant sequences with the type Iα *E. coli* DAH7PSs show very low sequence identity.[124] However, expression of the potato (*shkA*) DAH7PS has been shown to complement an *E. coli* mutant devoid of the enzyme.[126] The Mn^{2+}-stimulated DAH7PSs from carrot roots (*D. carota*) and potato (*S. tuberosum*) have been purified and consist of two identical subunits of approximately 55 kDa each,[68,124] which is consistent with the molecular mass of type II DAH7PSs from microbial sources.

1.9 Mechanism of DAH7PS

The majority of mechanistic studies on DAH7PS have been performed on the type Iα enzymes from *E. coli*. Labelling studies indicate the release of inorganic phosphate from PEP occurs by the cleavage of the C-O bond rather than the usual O-P bond.[127,128]

This requires a reaction mechanism in which the anomeric oxygen is derived from a water molecule. The reaction has also been shown to proceed with a defined stereochemical course with the *si* face attack of the C3 of PEP on the C1 *re* face of E4P (figure 1.15).[127,129,130]

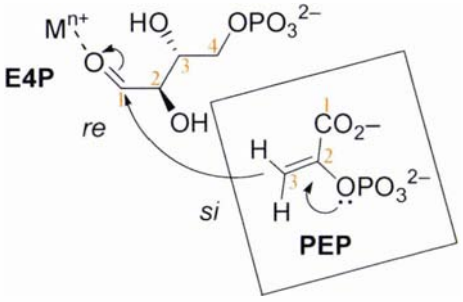


Figure 1.15 Steric course of the DAH7PS reaction

Early studies indicated a ping-pong mechanism for *Ec*-DAH7PS(Phe),[131] although more recent studies with lower concentrations of substrates and more highly purified enzyme suggest an ordered sequential mechanism in which PEP binds first and DAH7P is the last product to be released from the enzyme (figure 1.16).[132]

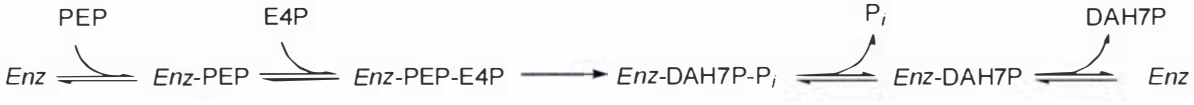


Figure 1.16 Order of substrate binding and product release of DAH7PS

Several mechanisms have been proposed for the DAH7PS-catalyzed reaction. The first mechanism involves nucleophilic attack of C3 of PEP on C1 of E4P, followed by attack of C3-OH of E4P on the carbocation formed at C2 of PEP (figure 1.17). This cyclic mechanism is not consistent with recent structural and modeling studies that suggest the C3 hydroxyl group is too far away for attack on C2 of PEP. E4P has been modeled, in its extended conformation, into the active site of *Ec*-DAH7PS(Phe),[92] which positions the monosaccharide so that C1 of E4P approaches within 2.4 Å of C3 of PEP. This is consistent with the reaction going through a linear intermediate with a water molecule attacking C2 of PEP.[92,128]

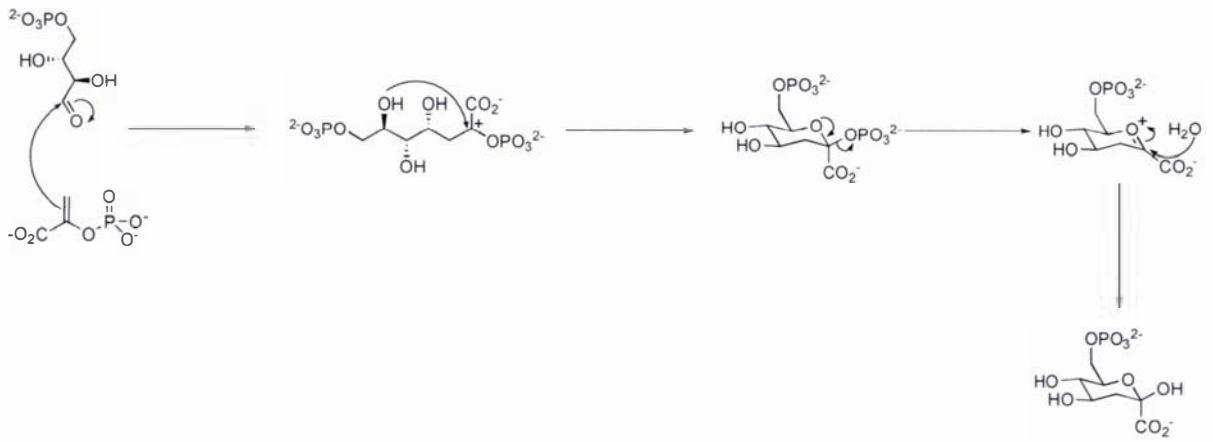


Figure 1.17 Proposed cyclic mechanism of DAH7PS

The detailed course of reaction is still under debate. The two major routes that are currently proposed differ in the initial step (figure 1.18). In path A, C3 of PEP attacks the metal-activated carbonyl of E4P, forming a linear oxocarbenium ion. A water molecule then attacks the positive centre at C2 of the acyclic intermediate.[51,53,81] In path B, initial attack of the C2 of PEP by a water molecule or hydroxide ion gives a tetrahedral carbanion, which then attacks C1 of E4P.[77,93,129,132] Both of these paths lead to the acyclic form of DAH7P, which can then cyclize to the more stable form.[133]

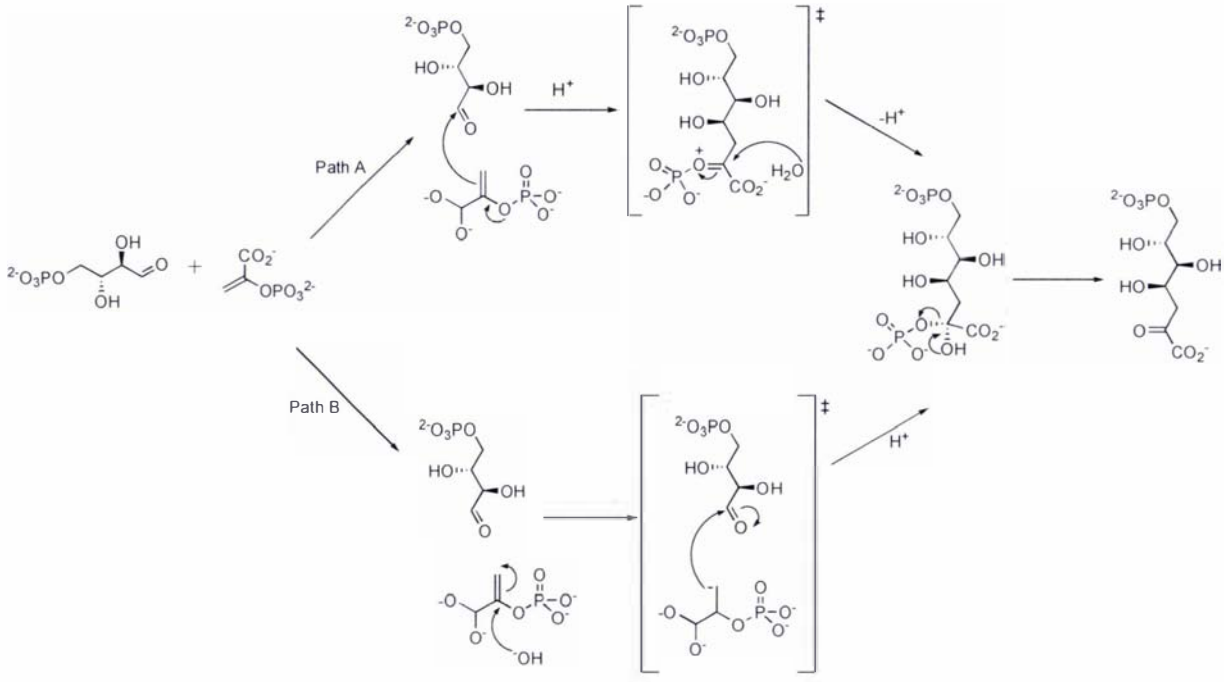


Figure 1.18 Two proposed routes for the mechanism of DAH7PS

The availability of crystal structures of DAH7PS has provided some insight into the mechanism catalyzed by DAH7PS, however these results are not without disagreement. Structural and modeling studies have allowed E4P to be positioned into the active site of several DAH7PSs, refer to Sections 1.5.2 and 1.6.1.2. The observed binding of E4P and PEP in both structures is consistent with activation of the aldehyde (E4P) by the metal, indicating that the divalent metal plays an essential and indispensable role.[51,53,107] The second step of the mechanism would then be the attack of water on C2 of PEP.

Two water molecules have been identified in both type Iα DAH7PSs that are close enough to PEP to act as the nucleophilic water required in the reaction. One is located on the *si* face and the second on the *re* face of PEP, both ~3 Å away from C2 of PEP.[53,92] It is not yet known which of these waters participates in the reaction. The water on the *si* face of PEP would present an unusual *syn* addition of water and E4P to the *si* face of PEP.[92] The second water would require a residue acting as a base to extract a proton from the water molecule, and to date no such residue has been identified.[53,92,93]

1.10 Mechanism of KDO8PS

Since the similarities in sequence, reaction mechanism and crystal structure have been known, it has been thought that KDO8PS and DAH7PS catalyze analogous condensation reactions using a common mechanism.[103,109] Both enzymes catalyze the condensation of PEP with a phosphorylated monosaccharide (figure 1.19) via an ordered sequential mechanism where PEP binds first and DAH7P or KDO8P is released last.[105] Like DAH7PS, KDO8PS catalyzes the C-O rather than the usual O-P cleavage of PEP.[134] The KDO8PS reaction parallels the DAH7PS reaction with the mechanism proceeding via a linear intermediate,[103,115] and the *si* face addition of PEP to the *re* face of A5P.[135] The two proposed mechanisms for KDO8PS are identical to the two mechanisms described in the previous section for DAH7PS. Like the DAH7PS-catalyzed reaction, the analogous reaction catalyzed by KDO8PS is yet to be clarified.[106,136]

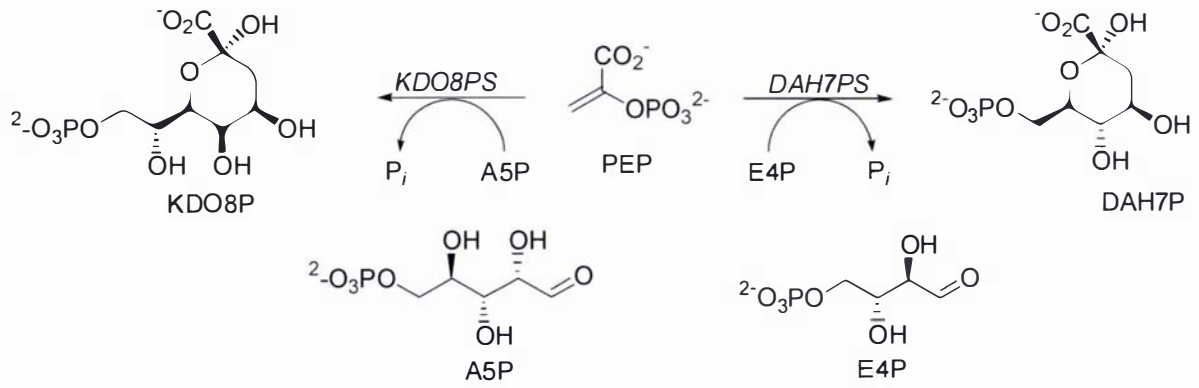


Figure 1.19 Reactions catalyzed by DAH7PS and KDO8PS

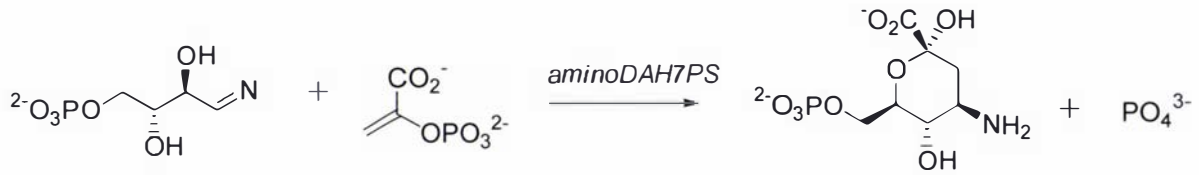
The two waters that are proposed to be involved in the reaction for DAH7PS have also been identified in several KDO8PS structures.[115] Again, it is still unclear as to whether the water molecule located on the *si* or *re* face of PEP is responsible for the attack of PEP.[115]

The discovery of various mechanistic and structural similarities between DAH7PS and KDO8PS has lead to the assumption that both enzymes share a common mechanism. However, several key differences have been neglected until recently, that suggest the two structurally-related enzymes may catalyze their reactions via different mechanistic routes.[107,136,137] Structural, kinetic and site-directed mutagenesis studies on DAH7PS have indicated the metal in DAH7PS plays an indispensable role through activating the aldehyde group of E4P.[81,107,136] In contrast, structural analysis of KDO8PS reveals the metal is likely to be too far away from the aldehyde moiety to be involved in electrophilic activation, and activation is proposed to be through protonation involving the phosphate group of PEP.[107] The ability to convert a metal-dependent KDO8PS into a metal-independent enzyme following the mutation of the metal-binding cysteine to asparagine (found absolutely conserved in metal-independent KDO8PSs)[112,113] suggests that the two types of KDO8PS catalyze the condensation reaction using similar chemistry. This is also consistent with the structural role predicted for the metal by Woodard *et al.*[112,138] The equivalent mutation in a DAH7PS results in a mutant enzyme which is unable to bind metal ions and shows no detectable enzymatic activity with or without EDTA or added metal.[107]

Another key difference between KDO8PS and DAH7PS is their tolerance to changes at the C2 position of the monosaccharide. As discussed in Section 1.6.2.1, there is a C2 configurational difference between C2 of E4P and A5P, the natural substrates of DAH7PS and KDO8PS, respectively. DAH7PS from *P. furiosus* is unselective at this position and accepts T4P,[107] which has the opposite stereochemistry at the C2 position of E4P (but the same as A5P). On the other hand, KDO8PS has a strict requirement for the correct stereochemistry at this position and does not accept R5P, which has the opposite stereochemistry of A5P at the C2 position, as a substrate.[101,106] The absolute selectivity at the C2 position of KDO8PS reflects the importance the C2-OH of A5P has in catalysis.[107] Structural studies have shown that C2-OH is coordinated to the metal either directly, or indirectly (refer to Section 1.6.2.2) to orient the monosaccharide into the correct position for attack by PEP.[115] In metal-independent KDO8PSs the C2-OH is coordinated to an asparagine sidechain (absolutely conserved in metal-independent KDO8PSs) most likely *via* an intermediate water.[107]

1.11 Involvement of Type II DAH7PSs in a Novel Shikimate Pathway

It has recently been proposed that the type II DAH7PSs from *A. mediterranei* and *Streptomyces collinus* are involved in a variant to the shikimate pathway, the amino shikimate pathway, in which a nitrogen atom is introduced in the very first step of the pathway.[139] These type II enzymes, referred to as aminoDAH7PSs, appear to be similar in sequence to type II DAH7PS enzymes involved solely in aromatic amino acid biosynthesis. AminoDAH7PSs are responsible for the condensation of PEP and iminoE4P to give a 4-amino analogue of DAH7P, 4-amino-3,4-dideoxy-D-*arabino* heptulosonate 7-phosphate (aminoDAH7P) (figure 1.20). They are also able to utilize E4P reasonably efficiently.[43,140] The source of the nitrogen atom is still unclear, although recent studies suggest kanosamine biosynthesis is the likely source of iminoE4P.[43] AminoDAH7P is an intermediate in the synthesis of 3-amino-5-hydroxybenzoic acid (AHBA),[141-143] a precursor for the formation of rifamycin (*A. mediterranei*), ansatrienin (*S. collinus*) and naphthomycin (*S. collinus*).



iminoE4P PEP aminoDAH7P

Figure 1.20 Reaction catalyzed by aminoDAH7PS

Floss and co-workers have shown that aminoDAH7P, not DAH7P, is converted into AHBA in *S. collinus* and *A. mediterranei*.[139,144] They have characterized the aminoshikimate enzymes encoded by the rifamycin (*rif*) gene cluster of *A. mediterranei* and have identified a gene encoding a type II DAH7PS within the cluster.[144] The growth of evidence supporting the involvement of type II DAH7PSs in AHBA biosynthesis indicates the diverse roles that type II enzymes play in nature, in comparison with their type I counter-parts.

1.12 OUTLINE OF THESIS

The objective of this project is to structurally and functionally characterize the type II DAH7PSs from *H. pylori* and *Mycobacterium tuberculosis*. Prior to the studies described in this thesis, there was limited functional information and no structural data on type II enzymes. Such information would allow insight into whether the two separate homology classes catalyze the condensation reaction between PEP and E4P by the same mechanism. A more complete understanding of both the functional and structural similarities and differences between type I and type II DAH7PSs are needed to determine the evolutionary relationship between these two apparently unrelated types of DAH7PSs. Characterization of the two pathogenic type II enzymes may also aid in the future development of type-specific antibiotics.

The specific goals of this project are to:

- Clone, express and solubilize type II DAH7PSs from *H. pylori* and *M. tuberculosis*.

- Functionally characterize a type II DAH7PS enzyme to investigate whether it catalyzes a similar reaction to type I enzymes.
- Investigate the substrate specificity of a type II DAH7PS using substrate analogues of E4P.
- Solve the structure of a type II DAH7PS enzyme and compare this to the known structures of type I enzymes.
- Perform mutagenesis and structural studies to probe active site residues and determine their role in catalysis.
- Investigate the feedback-regulation of both type II enzymes, using functional and structural methods.

CHAPTER TWO

EXPRESSION, SOLUBILIZATION AND BIOCHEMICAL CHARACTERIZATION OF TYPE II DAH7PS FROM *H. PYLORI*

2.1 Introduction

H. pylori strain J99 is a pathogenic Gram-negative bacterium which colonizes the gastric and upper intestinal epithelium of humans. This organism has been established as the major aetiological agent of chronic gastritis, is associated with peptic ulcers and, more recently, has been linked to the development of gastric cancer.[145,146] The genome of this human gastric pathogen encodes a single DAH7PS that belongs to the type II family. The characterization of microbial type II DAH7PSs has been limited compared to that of type I enzymes. The only reported investigations on microbial type II DAH7PSs are from *Streptomyces*,[37,117,118,121,147] *N. crassa*.[37,119,148,149] and *X. campestris*.[38] Most of these studies have been performed on partially purified protein, and were carried out prior to the availability of primary sequence information. *Hp*-DAH7PS is the first microbial type II enzyme to be characterized where complete genomic information supports its role in primary metabolism. The expression, *in vivo* solubilization, purification and biochemical characterization of *Hp*-DAH7PS will be discussed in this chapter.

2.2 Cloning and Expression of *Hp*-DAH7PS

DNA corresponding to the open reading frame (ORF) of *Hp*-DAH7PS, from *H. pylori* strain J99 genomic DNA, was ligated into pET-32a(+) and used to transform *E. coli* BL21(DE3) cells (for further details refer to Chapter Six). SDS-PAGE analysis of the whole cells showed an over-expressed protein of the predicted subunit molecular weight for *Hp*-DAH7PS (MW ~51 kDa). Following cell lysis and centrifugation this protein was evident only in the pellet fraction (figure 2.2, lane 6). No DAH7PS activity was observed in the supernatant (figure 2.2 lane 5), suggesting that the enzyme was completely insoluble. The *Hp*-DAH7PS insert in the expression plasmid was sequenced and found to be identical to the predicted sequence. Standard strategies to improve the

solubility of recombinant heterologous protein in *E. coli* were investigated without success. These included decreasing the growth temperature from 37 °C to 30 °C or 25 °C, lowering the concentration of IPTG from 1 mM to 0.2 mM, and performing growth time trials at two, four, six, and eight hours post-induction. Four hours post-induction gave the best over-expression but all conditions showed DAH7PS in the pellet fraction after lysis. The effect on solubility of several additives to the lysis buffer were tested: DTT (2 mM), KCl (200 mM), Thesit (0.005 %), EDTA (1 mM), $MnSO_4$ (100 μM), PEP (200 μM), and protease inhibitor (EDTA-free cocktail). These additives were ineffective at solubilizing recombinant *Hp*-DAH7PS.

The next step was to try transforming the plasmid into Origami B (DE3) cells which have a less reducing cytoplasmic environment that facilitates cytoplasmic disulfide bond formation. Cells were grown at 37 °C, innoculated with either 1 mM or 0.2 mM IPTG, and harvested two hours after induction. Recombinant DAH7PS was still insoluble. Cells were grown at 30 °C until OD_{600} ~0.5 and then transferred to 22 °C, induced with 0.2 mM IPTG and then harvested either three, six, or seventeen hours after induction. Cells were lysed in the presence of either DTT (2 mM), Thesit (0.005 %), or EDTA-free protease inhibitor cocktail. Results showed that none of these conditions were successful in producing soluble and active *Hp*-DAH7PS protein.

2.3 Co-expression of *Hp*-DAH7PS and *E. coli* Chaperonins

Among the numerous strategies that have been developed to overcome the problem of inclusion body formation in *E. coli*, co-expression of molecular chaperones is becoming increasingly popular.[150] Based on the observation of Houry *et al* [151] that *E. coli* DAH7PSs are substrates for *E. coli* chaperonins, we co-expressed GroEL/GroES and *Hp*-DAH7PS in *E. coli*.

GroEL is a homo-oligomer of fourteen subunits, each of molecular mass 57 kDa, which are arranged into two heptameric rings, forming a cylindrical structure with two large cavities.[152] Typically, incorrectly folded proteins display exposed hydrophobic regions, which are recognized by GroEL/GroES. The substrate protein binds in the central cavity of the cylinder, interacting with the hydrophobic surfaces exposed by the

GroEL/GroES complex (figure 2.1). The ring-shaped cofactor GroES then binds to GroEL in an ATP-dependent reaction, closing off the cavity from the crowded *E. coli* cytosol, giving the mis-folded protein a second chance to fold into its native conformation. After 6-10 seconds of folding, when GroEL-bound ATP is hydrolysed to ADP, ATP binding to the opposite ring of GroEL triggers the dissociation of GroES and the folded protein from GroEL.[152,153] Proteins that are heavily dependent on GroEL may require several rounds of interaction with GroEL to reach their native state.[151]

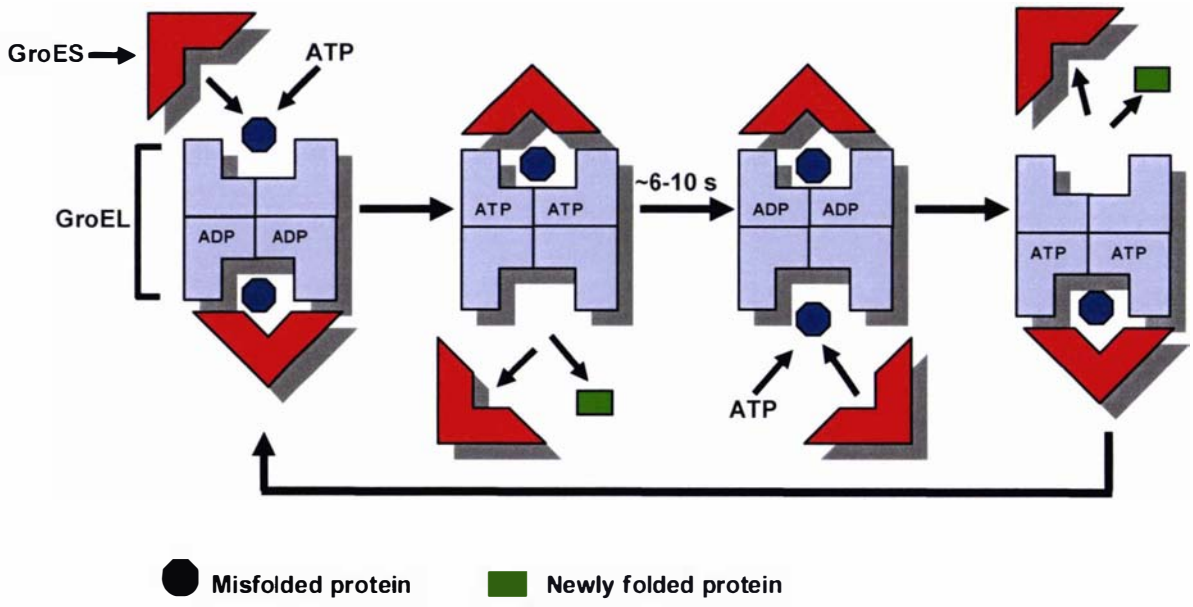


Figure 2.1 **Schematic diagram of GroEL-GroES cycle**

Incorrectly folded protein binds in the central cavity of GroEL (blue sphere), followed by the binding of GroES and ATP to GroEL. After 6-10 seconds of folding, when GroEL-bound ATP is hydrolyzed to ADP, the newly folded polypeptide (green rectangle) is released from the GroEL/GroES complex.

Numerous examples of GroEL/GroES-mediated *in vivo* solubilization of homologously[154] and heterologously[155,156] overexpressed proteins have been reported. Houry *et al.* found that substrates for GroEL/GroES consist preferentially of two or more domains containing αβ-folds, which contain α-helices and buried β-sheets with extensive hydrophobic surfaces. The theoretical size limit for GroEL/GroES is approximately 60 kDa.[153] The *E. coli* type 1 DAH7PSs belong to the aldolase superfamily,[71] a member of the TIM β/α-barrel fold. While an evolutionary relationship

between type I and type II enzymes had not been established at the time, some automated servers[157] predicted that *Hp*-DAH7PS is structurally related to the type I enzymes. A Superfamily[158] analysis of *Hp*-DAH7PS suggests that central regions of the sequence of this enzyme may belong to the TIM β/α-barrel fold aldolase and enolase C-terminal domain-like superfamilies.

The co-expression of *Hp*-DAH7PS and GroEL/GroES was performed by transforming pET-32a(+)-*Hp*-DAH7PS and a plasmid containing the ORFs of *E. coli* GroEL and GroES, pGroESL (kindly donated by Dr Mark Patchett), into *E. coli* BL21(DE3) cells. The first of the two plasmids was successfully introduced into the cells by chemical transformation, however the second plasmid could not be introduced into the competent cells using this protocol. To overcome this problem electro-competent BL21(DE3)-pGroESL cells were made and transformed with pET-*Hpy*DAH7PS by electroporation. It should be noted that *E. coli* BL21(DE3)/pGroESL/pET-*Hpy*DAH7PS cells took twice as long to grow compared with the growth of cells transformed with pET-*Hpy*DAH7PS alone.

The cells were grown at 37 °C until an OD_{600} of ~0.6 was reached when the cells were transferred to 25 °C and induced with 1 mM IPTG. The cells were harvested six hours after induction. The cells were lysed by sonication. After removal of cell debris by centrifugation, approximately 65 % of the *Hp*-DAH7PS was found in the supernatant (figure 2.2 lane 3), while the remaining 35 % was found in the pellet (figure 2.2, lane 4). The supernatant exhibited DAH7PS activity. Lysis conditions were varied in an attempt to optimize the proportion of soluble protein. The addition of salt (KCl), detergents (Thesit and *n*-octyl β-D-gluco-pyranoside), EDTA, PEP or metal ($MnSO_4$) to the lysis buffer did not enhance the solubility of the recombinant DAH7PS.

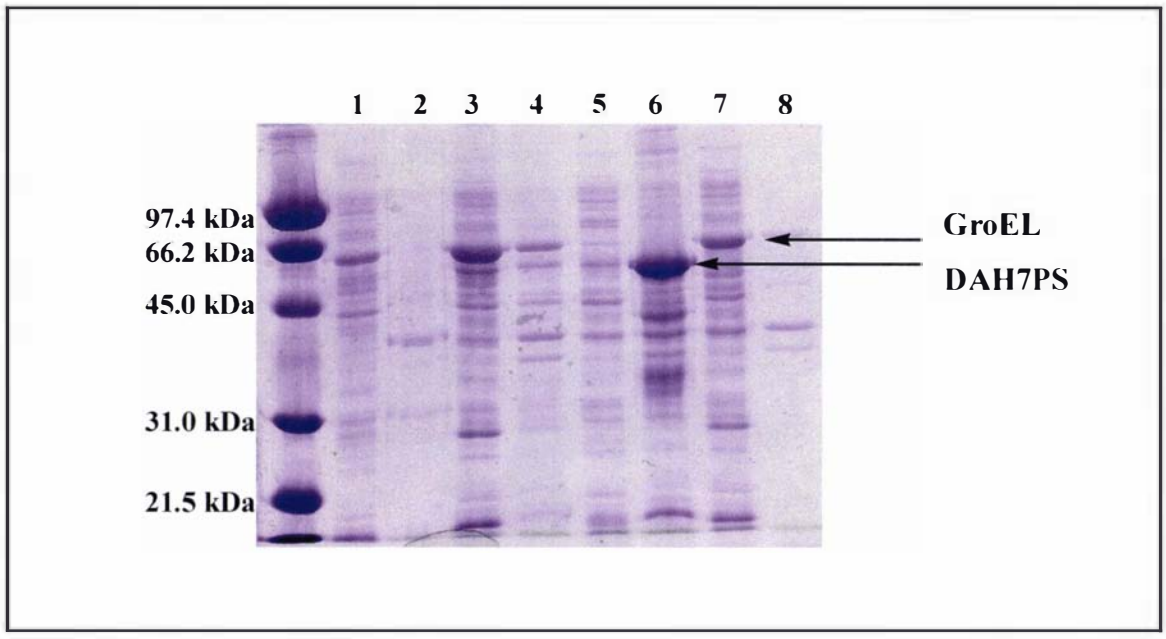


Figure 2.2 **SDS-PAGE analysis of *Hp*-DAH7PS in supernatants and pellets after cell lysis.** SDS-PAGE was performed on a 12 % polyacrylamide gel and protein bands visualized with Coomassie Brillant Blue R250. BL21(DE3)/pGroESL/pET-HpyDAH7PS, Preinduction: Lane 1 - supernatant; Lane 2 - pellet. BL21(DE3)/pGroESL/pET-HpyDAH7PS, Induced cells Lane 3 - supernatant; Lane 4 -pellet. BL21(DE3)/pET-HpyDAH7PS, Induced cells: Lane 5 - supernatant; Lane 6 pellet. BL21(DE3)/ pGroESL, Induced cells: Lane 7 - supernatant; Lane 8 -pellet.

2.4 Purification of *Hp*-DAH7PS

A two-step purification procedure based on the theoretical isoelectric point (pI) (7.5) of *Hp*-DAH7PS was developed. The results are summarized in table 2.1. A substantial amount of protein was lost throughout the purification as fractions pooled were selected for purity rather than for total enzyme recovery. It should also be noted that endogenous *E. coli* DAH7PS will account for some of the DAH7PS activity detected in the crude lysate fraction.

Whole cells were lysed using a French Press and the crude lysate was clarified by centrifugation. The supernatant was then diluted and applied to an anion exchange (Source 15Q®) column. To determine the optimum pH for loading *Hp*-DAH7PS onto the source 15Q® column a batch adsorption test was performed with an aliquot of

Source 15Q® resin at pH 6.5, 7.5 and 8.5. The highest pH was required for the binding of *Hp*-DAH7PS to the Source 15Q® resin (figure 2.3). This result was confirmed by activity assay, therefore crude lysate was loaded onto the column equilibrated in pH 8.5 buffer (BTP (10 mM), DTT (2 mM)). *Hp*-DAH7PS was eluted off the Source 15Q® column in approximately 150 mM NaCl.

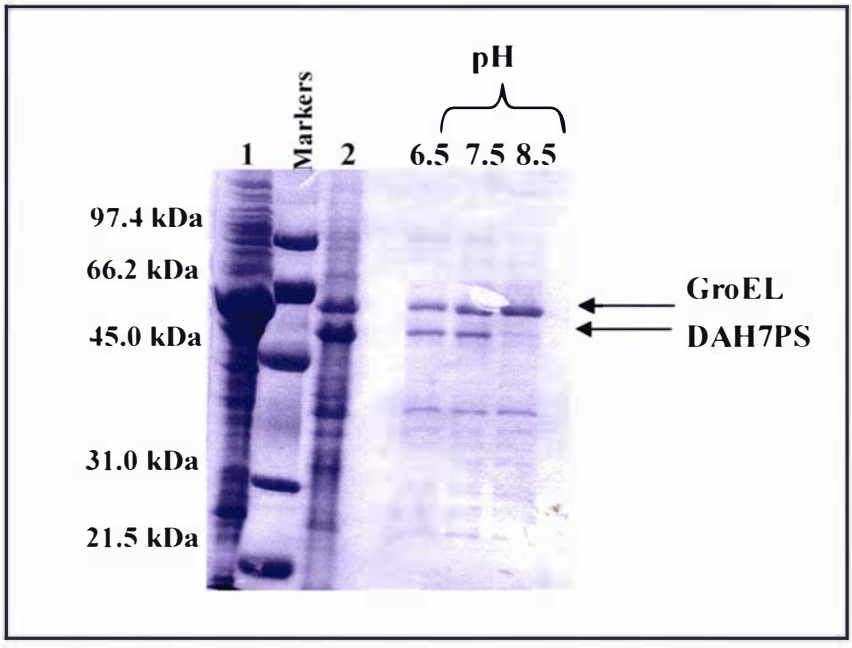


Figure 2.3 **SDS analysis of binding of *Hp*-DAH7PS onto Source 15Q® resin using different pHs.** Lane 1 and lane 2 are the supernatant and pellet fractions after cell lysis. The whole cells were resupsended in either pH 6.5, 7.5 or 8.5 buffer, lysed and added to Source 15Q® resin. The resin was then washed with the appropriate pH buffer and the unbound protein fractions analyzed by SDS-PAGE on a 12 % polyacrylamide gel and proteins visualized with Coomassie Brillant Blue R250.

In the first few purifications of *Hp*-DAH7PS the next step was a second passage down the Source 15Q® column equilibrated in pH 7.0 buffer. This acted as a negative purification in which *Hp*-DAH7PS passes through the column unbound, and contaminating *E. coli* proteins bind and elute in a salt gradient. Due to the instability of *Hp*-DAH7PS upon dilution of the protein (Section 2.5) future purifications of *Hp*-DAH7PS exchanged this purification step for a cation exchange (Mono S®) column. Fractions containing *Hp*-DAH7PS were pooled after the Source 15Q® column and buffer exchange in pH 6.5 buffer was performed using a 10 kDa MWCO concentrator

(Vivascience). A pH of 6.5 was chosen to load *Hp*-DAH7PS onto the cation exchange column as it is one unit below the predicted pI for the enzyme and activity assays indicated that *Hp*-DAH7PS is active at this pH, reassuring that the enzyme is stable at this pH. SDS-PAGE analysis of the pooled fractions from the cation exchange column containing *Hp*-DAH7PS showed that the GroEL protein had been substantially separated from *H. pylori* enzyme so that only a single major band of ~51 kDa was seen (figure 2.4, lane 3). In subsequent purifications a Source 15S® column was substituted for the Mono S® column with similar results. This purification step resulted in protein with >95 % purity. Purifications using both cation exchange columns resulted in *Hp*-DAH7PS being eluted off the column in approximately 130 mM NaCl.

In later purifications a size exclusion step was added which resulted in pure (>99 %) protein. A Superdex S200 column was used as the final step in the purification with the buffer consisting of BTP (10 mM), DTT (1 mM), PEP (200 μM) and $MnSO_4$ (100 μM) at pH 7.0). Initial concerns over enzyme stability meant that earlier purifications did not include this step. It should be noted that BTP was used as the buffer throughout the entire purification due to its wide pH buffering range and because it has a tertiary amine group that does not interact with E4P.

Purification Step	Total protein (mg)	Total enzyme activity (Units)[b]	Specific activity (Units mg^{-1})	% Yield	Relative purity
1. Crude lysate	333	69.0	0.21	100	1.0
2. Source 15Q®	12	11.5	0.96	17	4.5
3. Mono S®	2	7.4	3.70	11	17.6

Table 2.1 Two-step purification procedure of *Hp*-DAH7PS[a]

[a]Purification of a 333 mL culture of BL21(DE3)/pGroESL/pET-*Hpy*DAH7PS

[b]Determined by measuring the loss of PEP at 232 nm at 30 °C.

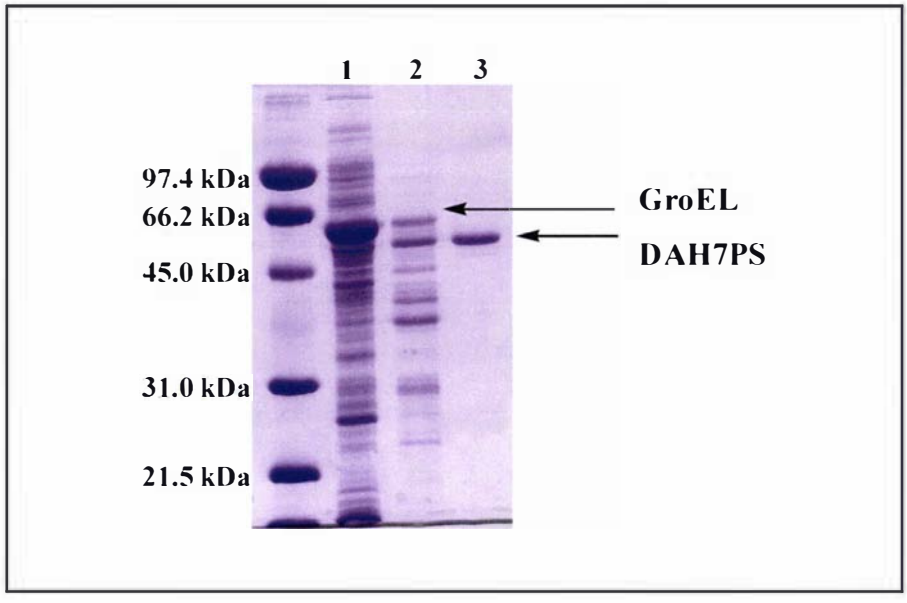


Figure 2.4 **SDS-PAGE analysis of purification of *Hp*-DAH7PS.**
Cells pelleted from a culture of *E. coli* BL21(DE3)/pGroESL/pET-*Hpy*-DAH7PS were re-suspended and the cell lysate clarified by centrifugation to yield a supernatant fraction (lane 1). Lane 2: Source 15Q® fractions, pooled and concentrated. Lane 3: Mono S® fractions, pooled and concentrated. SDS-PAGE analysis was performed on a 12 % polyacrylamide gel and protein bands visualized with Coomassie Brillant Blue R250.

2.5 Stability and DTT Reactivation of *Hp*-DAH7PS

The stability of *Hp*-DAH7PS (after passage through a Source 15Q® column, ~2 mg/mL) was tested at pH 8.5 under different storage conditions. About 10 % of enzyme activity was retained after two days stored at 4 °C, whereas enzyme stored at –80 °C after rapid freezing in liquid nitrogen retained the majority of its activity (~90 %). Protein stored at –80 °C appeared to retain activity for at least two months if stored at this temperature. Enzyme preparations with concentrations less than 2 mg/mL exhibited significant loss of activity over two days even when stored at –80 °C.

The instability of *Hp*-DAH7PS was partially overcome by the inclusion of DTT (1 mM) in all purification buffers. Whereas the absence of reducing agent necessitated a rapid two-step one-day purification, the addition of DTT allowed protein purification to be carried out in three-steps, over two days, with no significant loss of enzyme activity. Enzyme instability was also observed by Gosset *et al.* during the purification of type II

DAH7PS from *X. campestris*.[38] They reported increasing instability during purification or on dilution, which they overcame by carrying out a one-day purification of the enzyme. In contrast to DAH7PS from *H. pylori* the *X. campestris* type II enzyme was reportedly not stabilized by the presence of reducing agent.

DTT was found to protect *Hp*-DAH7PS from loss of activity at 4 °C, and DTT also restored activity to partially inactivated protein. The reactivation of *Hp*-DAH7PS by DTT was performed on purified enzyme, which had been stored at 4 °C for twenty-four hours without DTT present, retaining approximately 20 % of its original activity. Incubation of this partially inactivated enzyme with 1 mM DTT prior to initiation of the assay fully restored enzymic activity. The addition of 0.1 – 1 mM DTT to enzyme preparations prevented the loss of enzyme activity over twenty hours at 4 °C. *Hp*-DAH7PS appears to be indefinitely stable at -80 °C when stored in the presence of 1 mM DTT at concentrations greater than 2 mg/mL.

DTT activation has been reported for a number of type II DAH7PSs. Plant DAH7PSs have been shown to be hysteretically activated by DTT, and recently it has been reported that an *Arabidopsis* isoenzyme requires reduced thioredoxin for activity.[67] DTT enhanced the maintenance of activity of the microbial type II DAH7PSs from *S. aureofaciens* and *S. coelicolor*.[37,121] In contrast, the activity of the unstable partially purified DAH7PS from *X. campestris* was not enhanced by DTT.[38] While the activity of type I DAH7PSs appears to be dependent on conserved cysteine residues, no particular requirement for reducing reagents has been reported.[89]

2.6 Molecular Mass Determination of *Hp*-DAH7PS

To determine the molecular mass of *Hp*-DAH7PS size exclusion chromatography was used with protein standards of known molecular mass. The quaternary structure of *Hp*-DAH7PS in solution could then be determined by extrapolation from a standard curve of the log molecular mass against elution time for the protein standards (figure 2.5). *Hp*-DAH7PS was estimated to be 100 kDa which suggests the enzyme is dimeric in solution. This is consistent with other type I and type II DAH7PSs that have been reported to be either dimeric or tetrameric.

Protein	Molecular mass (kDa)	Elution time (min)
β-amylase	200	31.25
Alcohol dehydrogenase	150	34.8
Hp-DAH7PS	100 (by extrapolation)	36.0
Albumin	66	37.8
Cytochrome C	12.4	49.2

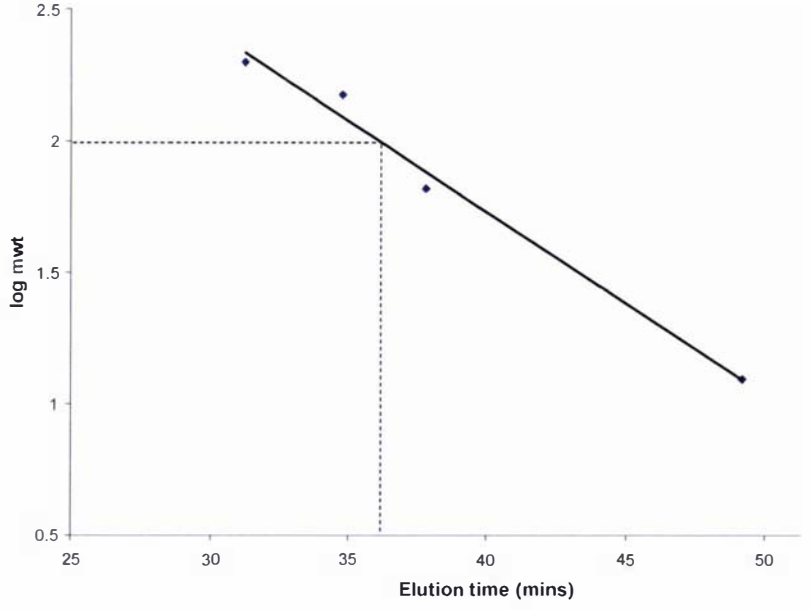


Figure 2.5 Standard curve of log molecular mass versus elution time for *Hp*-DAH7PS
The molecular mass of *Hp*-DAH7PS was determined by plotting the logarithms of the known masses of protein standards against their elution times. Refer to Chapter Six for experimental details.

2.7 Crystallization Trials with *Hp*-DAH7PS

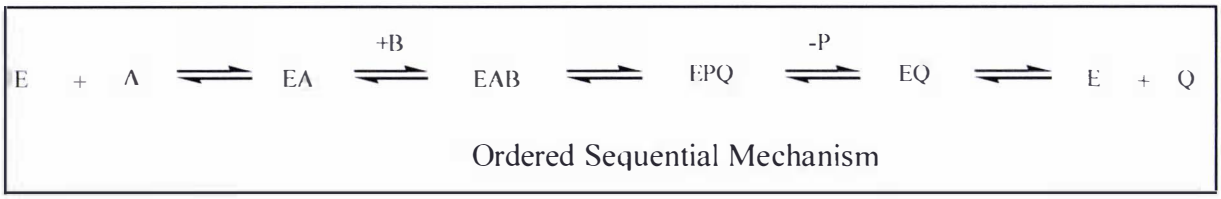
Over five hundred different conditions were screened using the Centre for Molecular Biodiscovery crystallization facility and all conditions were unsuccessful in producing crystals of *Hp*-DAH7PS. Crystallization trials were performed using freshly purified protein and protein stored at -80 °C, as well as different protein concentrations, and all were found to be unsuccessful. The focus of crystallization efforts then switched to another type II DAH7PS, *Mt*-DAH7PS (described in Chapter Three). Further investigation of the conditions required to crystallize the *H. pylori* protein are essential, as structural information may help in explaining the differences in feedback regulation between *Mt*-DAH7PS and *Hp*-DAH7PS (refer to Chapter Four), which cannot be explained at the primary sequence level.

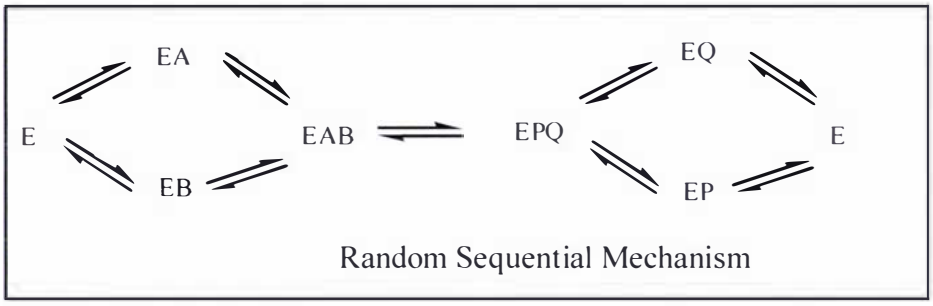
2.8 Kinetic Mechanism of *Hp*-DAH7PS

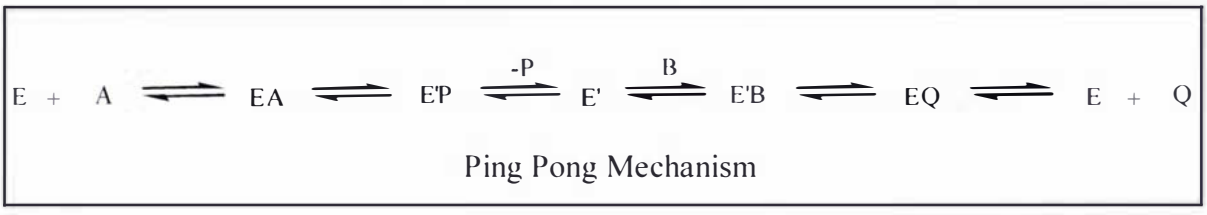
The reaction catalyzed by DAH7PS is a two-substrate two-product (bi-bi) reaction. There are three possible kinetic mechanisms for a two-substrate two-product reaction as seen in figure 2.6. The reaction mechanism may be sequential, where both substrates bind to the enzyme to form a ternary complex before the first product is formed, or ping-pong, where the first product leaves before the second substrate binds and the second substrate cannot bind to the enzyme but binds to the modified enzyme complex.[159] A sequential mechanism can be either random, where both substrates and both products bind to the same form of enzyme, or ordered, where the order of binding and release from the enzyme is compulsory as both substrates and products bind to different enzyme forms. To complicate things further, a ping-pong mechanism could appear to be sequential if there is a delay in the first product being released, giving the appearance that the second substrate has bound first.

Studies on type I DAH7PSs have indicated that this family of enzymes catalyze an ordered sequential reaction, where PEP binds first and DAH7P is the last product to leave the enzyme (refer to Chapter One, Section 1.9). To investigate the reaction catalyzed by the type II *Hp*-DAH7PS double-reciprocal (Lineweaver-Burk) plots of initial-velocity data were used to determine whether the reaction mechanism is sequential or ping-pong. To distinguish between random and ordered mechanisms product inhibition studies were performed using DAH7P.

Figure 2.6 **Possible kinetic mechanisms for two-substrate two-product reactions**







Sequential or Ping Pong Mechanism

Initial velocity values were determined as a function of the concentration of one substrate at various fixed concentrations of the other substrate. All measurements were done in duplicate to reduce error. Double-reciprocal plots generated from these measurements showed intersecting lines consistent with a sequential kinetic mechanism[160] (figure 2.7). This is also consistent with the mechanism determined for DAH7PSs from *E. coli* (*Ec*-DAH7PS(Tyr)),[75] *Salmonella typhimurium*[132] and *S. cerevisiae*.[65]

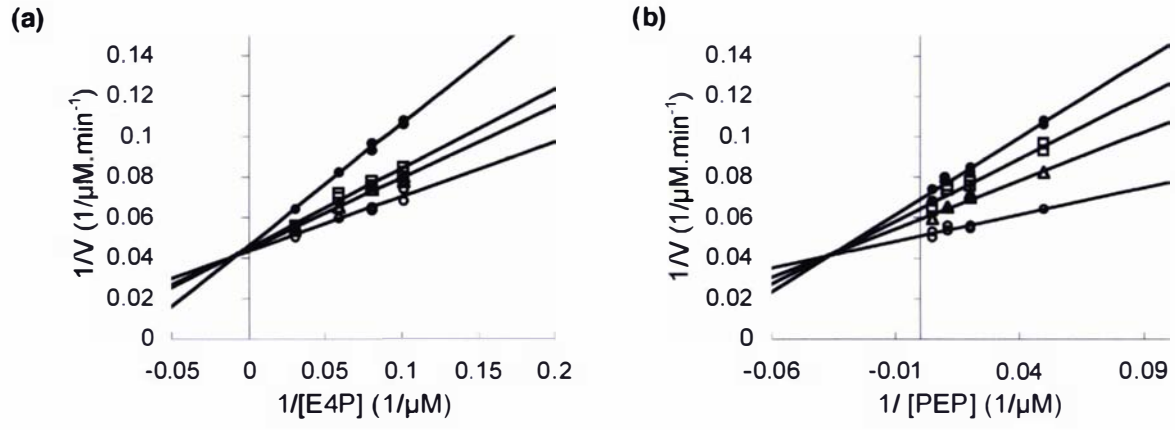


Figure 2.7 **Determination of the kinetic mechanism of *Hp*-DAH7PS**

(a) Double-reciprocal plots of initial velocities of DAH7PS as a function of E4P concentration at different PEP concentrations (32.5 μM (O), 16.8 μM (△), 12.3 μM (□), and 9.9 μM (●)). (b) Double-reciprocal plots of initial velocities of DAH7PS as a function of PEP concentration at different E4P concentrations (200 μM (O), 100 μM (△), 50 μM (□), and 20 μM (●)).

Random or Ordered Sequential Mechanism

The use of inhibitors, which compete with one of the substrates for a site on the enzyme, can give useful information as to the mechanism of catalysis. For Lineweaver-Burk plots drawn for experiments performed using a fixed concentration of one substrate and varying concentrations of the second substrate at different inhibitor concentrations, the overall pattern will be competitive (inhibitor binds to same enzyme-form as the variable substrate) if all the lines converge on the Y-axis. This is because the addition of a competitive inhibitor does not affect V_{max} but it alters the apparent K_M of the variable

substrate. On the other hand, if a series of parallel lines is obtained at different inhibitor concentrations the overall inhibition pattern is presumed to be non-competitive.[159] Non-competitive inhibitors alter V_{max} while leaving the apparent binding affinity of the enzyme for the substrate, K_M, unaltered.

Product inhibition studies were performed using DAH7P. Lineweaver-Burk plots were created as a function of the concentration of one substrate, PEP (figure 2.8 (a)) and E4P (figure 2.8 (b)), at various fixed DAH7P concentrations. Duplicate measurements were performed and in the case of discrepancies a triplicate measurement was taken. The convergence of the lines on the Y axis in figure 2.8 (a) indicates that DAH7P is a competitive inhibitor with respect to PEP. The parallel lines in figure 2.8 (b) indicate that DAH7P is a uncompetitive inhibitor with respect to E4P. Results therefore reveal that DAH7P binds to the same enzyme form as PEP and an ordered sequential mechanism is operating for *Hp*-DAH7PS where PEP binds first to the enzyme and DAH7P is the last product to be released (figure 2.9).

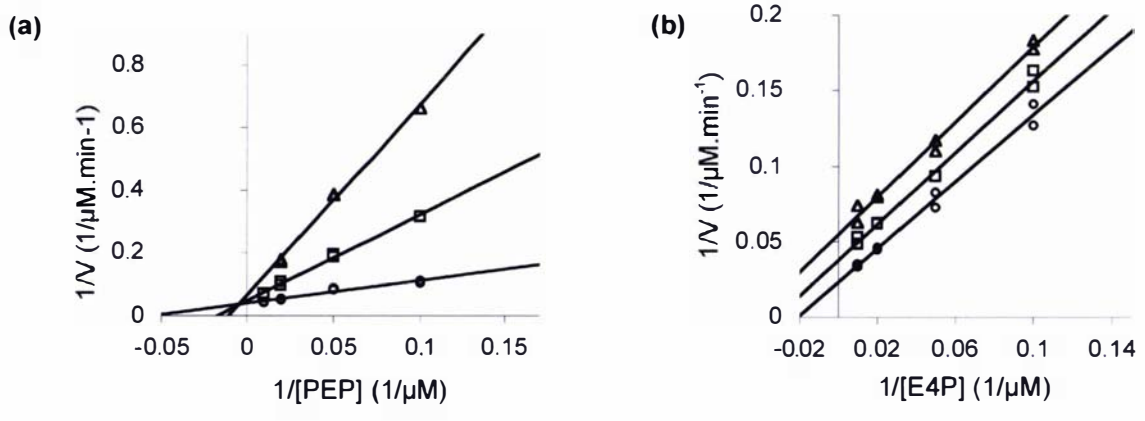


Figure 2.8 **Product inhibition of *Hp*-DAH7PS with DAH7P**

Double-reciprocal plots of initial velocities are given as a function of (**a**) PEP when the DAH7P concentrations were none (O) 1 mM (□) or 2 mM (△); (**b**) E4P when the DAH7P concentrations were none (O), 1 mM (□) or 2 mM (△).

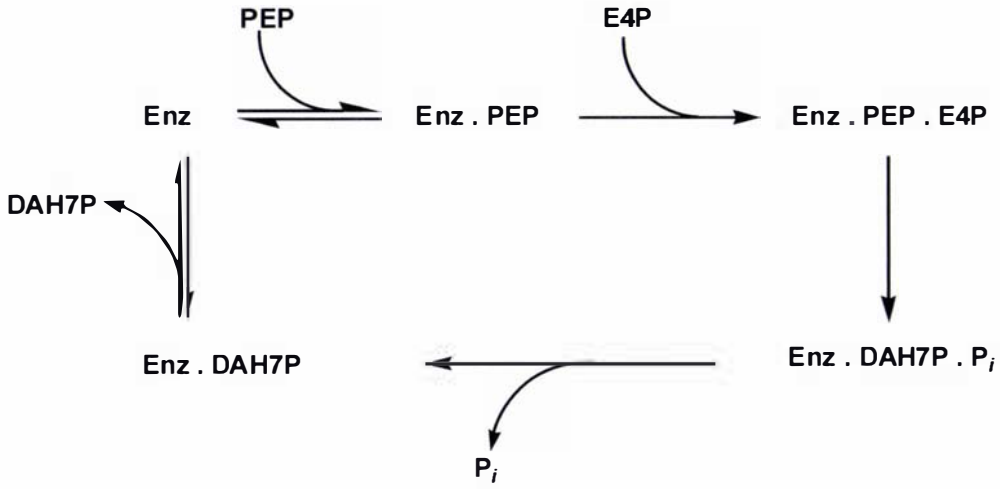


Figure 2.9 **Proposed order of substrate binding and product release**

The Lineweaver-Burk plot has been criticized on several grounds with the main criticism being the large weighting it gives measurements made at low substrate concentrations, which generally carry the most error. Compared with other linear transformations of the Michaelis-Menten equation (Eadie-Hofstee (V_0 vs $V_0/[S_0]$) and Hanes ($[S_0]/V_0$ vs $[S_0]$)) breaks in linearity are less obvious and this is important when investigating reaction mechanisms.[159] The Lineweaver-Burk plots have been used here to display steady-state kinetic results for *Hp*-DAH7PS for consistency with previously published DAH7PS kinetic studies.[132] Analysis of the *Hp*-DAH7PS data using other linear transformations yields identical conclusions.

2.9 Facial Selectivity of the *Hp*-DAH7PS-Catalyzed Reaction

While it is evident from the *arabino* configuration of the product that the *re* face of the aldehyde is attacked by PEP in the reaction, the use of a PEP analogue that distinguishes between the two geminal protons at C3 is required to determine the face of the enol phosphate that is involved (figure 2.10). (*Z*)-3-FluoroPEP (figure 2.11) acts as an alternative substrate to PEP for *Hp*-DAH7PS (V_{max} approximately 20 % of that with PEP). This compound was used to determine the stereochemical course of the enzymic reaction, by reacting (*Z*)-3-fluoroPEP with E4P and determining the position of fluorine in the product. The proton coupled ^{19}F NMR spectrum of the product 3-fluoroDAH7P showed a single resonance at -206.6 ppm that was split with coupling constants of 49.1

and 30.1 Hz (figure 2.12a). This spectrum is identical to that obtained when the type I *Ec*-DAH7PS(Phe) is used to catalyze the reaction.[161] The coupling constant value indicates that the fluorine occupies an axial position and that (3*S*)-3-fluoroDAH7P is generated in the enzymic reaction (figure 2.12b). It is therefore reasonable to conclude that the *Hp*-DAH7PS catalyzed reaction proceeds *via* interaction of the *si* face of PEP (equivalent to the *re* face of (*Z*)-3-fluoroPEP) and the *re* face of E4P.

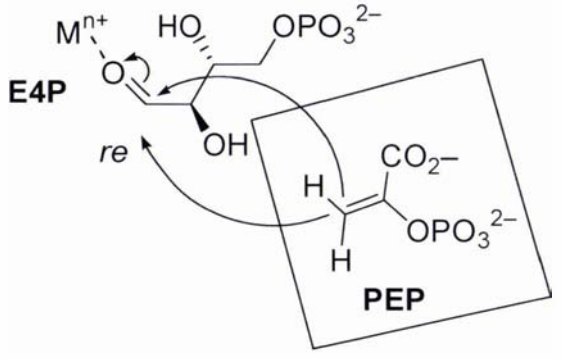


Figure 2.10 The two possible attacks of PEP on the *re* face of E4P

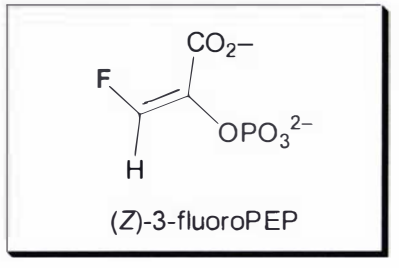


Figure 2.11

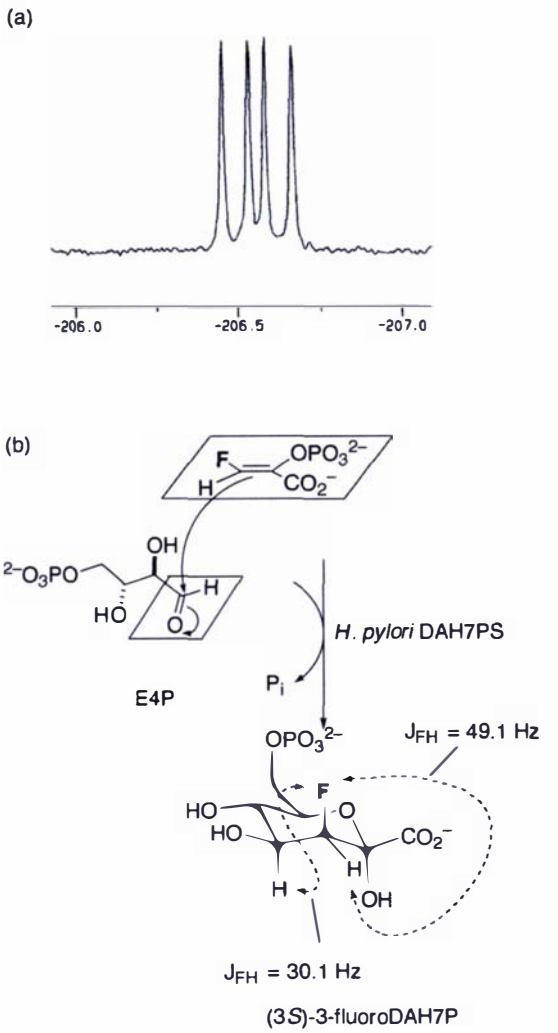


Figure 2.12 **Stereochemical course of the *Hp*-DAH7PS-catalyzed reaction.** (a) ^{19}F NMR spectrum (376 MHz) of product obtained by reaction of (*Z*)-3-fluoroPEP with E4P catalyzed by *H. pylori* DAH7PS. (b) Generation of (3*S*)-3-fluoroDAH7P. The axial position of the fluorine is confirmed by the large diaxial vicinial ^{19}F-^{1}H coupling constant of 30.1 Hz.

2.10 Initial Kinetic Parameters for *Hp*-DAH7PS

The steady-state kinetic constants for freshly purified *Hp*-DAH7PS were measured following the disappearance of PEP. Initial velocity values were determined as a function of the concentration of one substrate at various fixed concentrations of the other substrate. The apparent K_M values for E4P and PEP were 6 ± 1 μM and 3 ± 1 μM respectively, and the k_{cat} value was calculated as 3.0 ± 0.3 s^{-1} (figure 2.13).

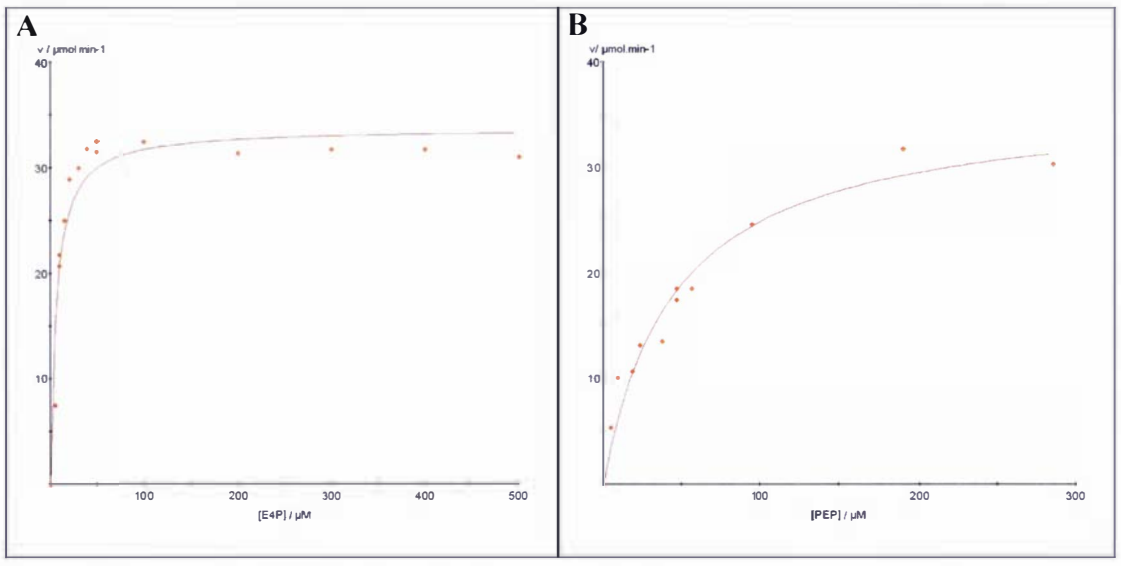


Figure 2.13 **Michaelis-Menten plot for determination of K_M values for E4P A, and PEP B.** Initial velocity values were determined as a function of E4P A, and PEP B, concentration at fixed concentrations of PEP A, and E4P B. The reaction consisted of E4P (7 µM to 670 µM), $MnSO_4$ (94 µM), and PEP (5 µM to 380 µM), in BTP buffer (50 mM, pH 7.5). The reaction was initiated by the addition of purified *Hp*-DAH7PS (2 µL, 10 mg/mL and carried out at 30 °C. K_M and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using Enzfitter (Biosoft).

The K_M values for *Hp*-DAH7PS are broadly in line with the kinetic constants reported for other type II DAH7PSs (table 2.3). It should be noted that the characterization of the *H. pylori* and *X. campestris* DAH7PSs has been performed on recombinant protein, while the other enzymes listed in table 2.3 were purified from native sources. As it is now known that some of these organisms possess more than one type II paralog, it is possible that the characteristics have been determined on a mixture of native isoenzymes.

2.11 Metal Reactivation of *Hp*-DAH7PS

As with all other DAH7PSs examined to date *Hp*-DAH7PS activity is dependent on the presence of a divalent metal ion (refer to Chapter One) (table 2.2). For *Hp*-DAH7PS metal ion reactivation studies buffer and substrate solutions were pretreated with Chelex to minimize the presence of metal ions from other sources. Enzyme activity was not detectable following treatment with EDTA (0.5 mM) for thirty minutes at 4 °C. This differs from the observations of Gosset *et al.*[38] Gorisch *et al.*,[121] and Yoo *et al*[118] for DAH7PSs from *X. campestris*, *S. aureofaciens* and *S. caespitosus* respectively, where EDTA treatment caused only partial inactivation. Enzyme activity could be restored by the inclusion of a variety of divalent metals in the activity assay. As can be seen in table 2.3 activity was restored upon addition of Co^{2+}, Mn^{2+}, Ca^{2+} Mg^{2+} or Cu^{2+}, however Zn^{2+}, Cd^{2+} and Ni^{2+} did not appreciably restore activity. Although Co^{2+}-activated enzyme gave the fastest reaction rates Mn^{2+} is used in all assays of *Hp*-DAH7PS. This is to allow comparisons with the type I DAH7PSs, which have generally been characterized in the presence of Mn^{2+}. The order of reactivation of *Hp*-DAH7PS is comparable with that reported for other microbial type II enzymes (table 2.2) that also show greatest activity with Co^{2+}.[38,118,162]

Divalent metal	% activity
Co^{2+}	100
Mn^{2+}	76
Ca^{2+}	58
Mg^{2+}	33
Cu^{2+}	12
Zn^{2+}	5
Cd^{2+}	2
Ni^{2+}	1
EDTA	0

Table 2.2 **Effect of divalent metals on *Hp*-DAH7PS activity[a]**

[a]Standard assay conditions detailed in Chapter Six were used.

Organism	K_M(PEP) (μM)	K_M(E4P) (μM)	Activation by divalent metal (decreasing order of effect)	Reference
H. pylori	3	6	Co^{2+}, Mn^{2+}, Ca^{2+}, Mg^{2+}, Cu^{2+}, Zn^{2+}	This study
X. campestris	130	230	Co^{2+}, Zn^{2+}, Mn^{2+}, Ni^{2+}, Mg^{2+}, Ca^{2+}, Cu^{2+}, Fe^{2+}	38
S. coelicolor	92	195	Not reported	37
S. caespitosus†	430	220	Co^{2+}, Zn^{2+} or Mn^{2+}restored activity to EDTA treated enzyme. Ni^{2+}, Fe^{2+} and Ca^{2+} inhibited activity.	162 118
S. rimosus†	6.7	2.6	Not reported	117
S. aurefaciens†	300	160	Partial inactivation with EDTA	121
N. crassa	12	2.7	EDTA-sensitive	119

Table 2.3 Comparison of properties of microbial type II DAH7PSs

†No primary sequence information is available for these enzymes, however the DAH7PSs from these species are highly likely to be type II enzymes as they all have subunit molecular masses of approximately 50 kDa.

2.12 Substrate Specificity

A number of monosaccharide phosphates were tested as alternative substrates to E4P (figure 2.14-2.18 and table 2.4). The three-carbon DL-glyceraldehyde 3-phosphate (G3P) and the six-carbon D-glucose 6-phosphate (G6P) were not substrates for *Hp*-DAH7PS. The E4P analogue 2-deoxyE4P, with the C2 hydroxyl group removed, was an excellent substrate. A variety of five-carbon sugar phosphates in which the C2 hydroxyl was either absent (2-deoxy-D-ribose 5-phosphate (R5P)), or in either possible configuration ((D-arabinose 5-phosphate (A5P) and R5P) were also able to act as substrates. A large-scale reaction of PEP with R5P catalyzed by *Hp*-DAH7PS was performed. The ^{1}H

NMR spectrum of the product isolated after purification was identical to the reported data for 3-deoxy-D-*altro*-octulosonate 8-phosphate,[77] confirming that the loss of PEP is catalyzed by *Hp*-DAH7PS.

The ability of *Hp*-DAH7PS to accept, although relatively poorly, five-carbon sugars, in which the C2 hydroxyl group is either absent or present in either possible configuration, and not accept three or six-carbon sugars as alternatives to E4P, parallels observations made with the type I DAH7PSs from *E. coli*[77] and *P. furiosus*.[47] However, while 2dE4P was found to be an excellent substrate for *H. pylori* and *P. furiosus* enzymes,[47] it was not as efficiently utilized by *E. coli* DAH7PS.[163] Substrate specificity studies have not been performed on any other microbial type II enzymes to date.

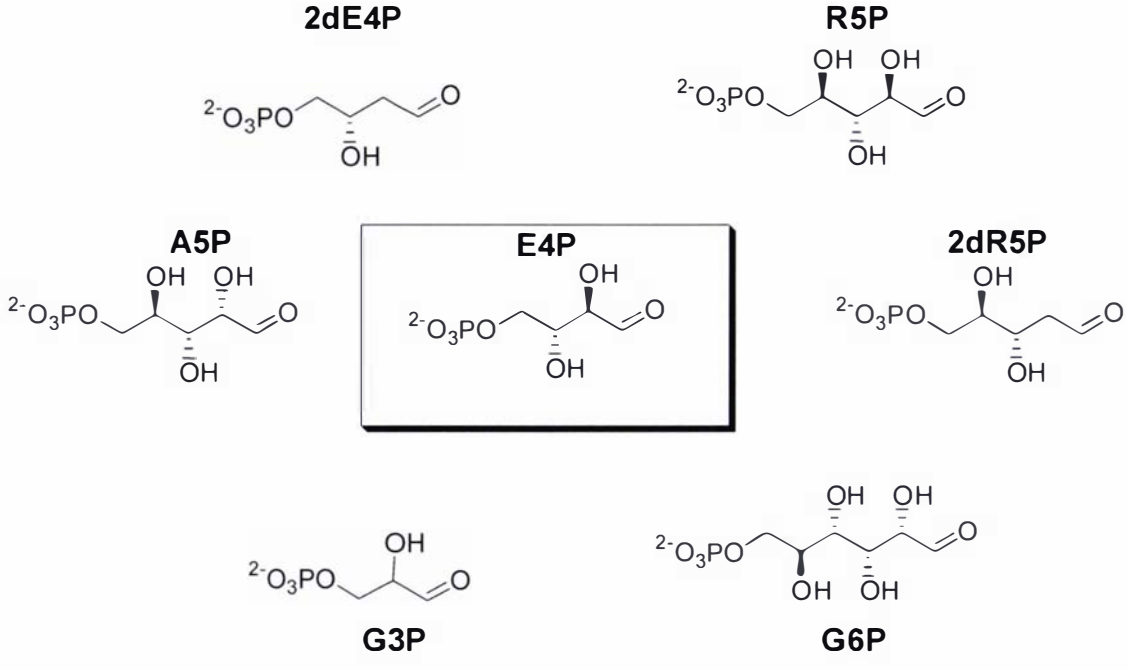


Figure 2.14 Phosphorylated monosaccharides tested as substrates for *Hp*-DAH7PS

Substrate	K_M (mono) (μM)	K_M (PEP) (μM)	k_{cat} (s^{-1})	k_{cat} / K_M (mono) ($\mu M^{-1}. s^{-1}$)
E4P	6 ± 1	3.1 ± 0.5	3.0 ± 0.3	0.5
2dE4P	18 ± 2	5 ± 0.3	6.0 ± 0.3	0.3
R5P	1050 ± 95	2.4 ± 0.2	1.2 ± 0.01	0.001
2dR5P	3040 ± 230	5.2 ± 0.5	1.0 ± 0.03	0.0003
A5P	4800 ± 360	8 ± 1	0.3 ± 0.01	0.00006

Table 2.4 Monosaccharide phosphates tested as alternative substrates to E4P

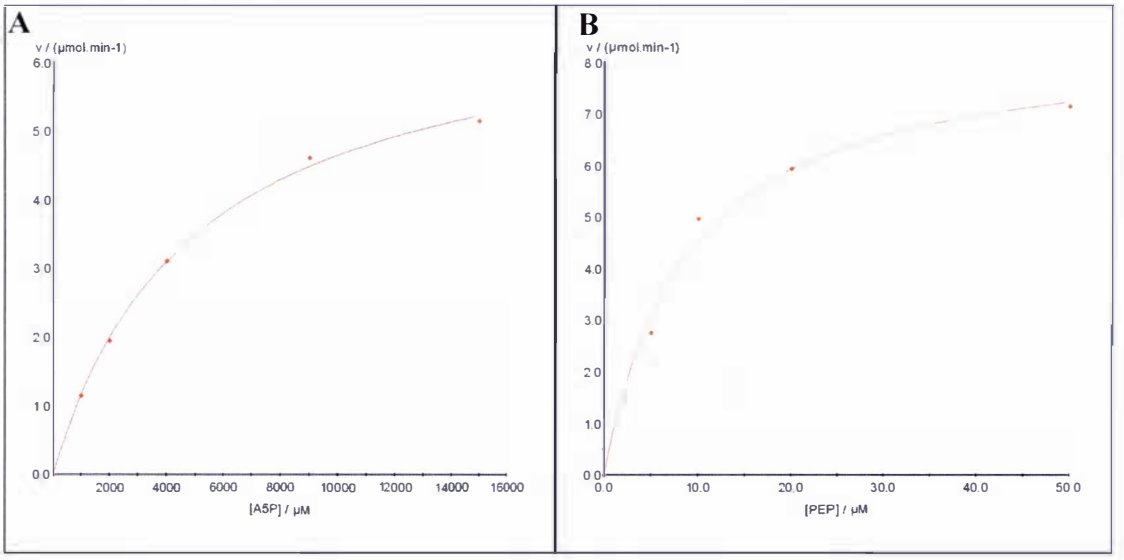


Figure 2.15 Michaelis-Menten plots for determination of K_M values for A5P A, and PEP B. Initial velocity values were determined as a function of A5P A, and PEP B, concentration at fixed concentrations of PEP A, and A5P B. For experimental details refer to Chapter Six.

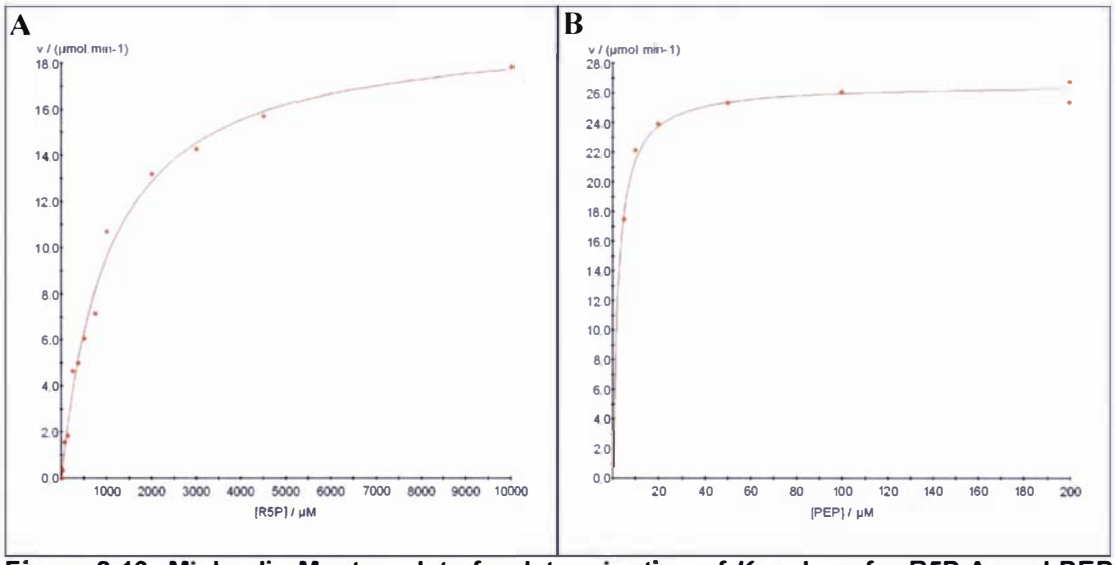


Figure 2.16 Michaelis-Menten plots for determination of K_M values for R5P A, and PEP B. Initial velocity values were determined as a function of R5P A, and PEP B, concentration at fixed concentrations of PEP A, and R5P B. For experimental details refer to Chapter Six.

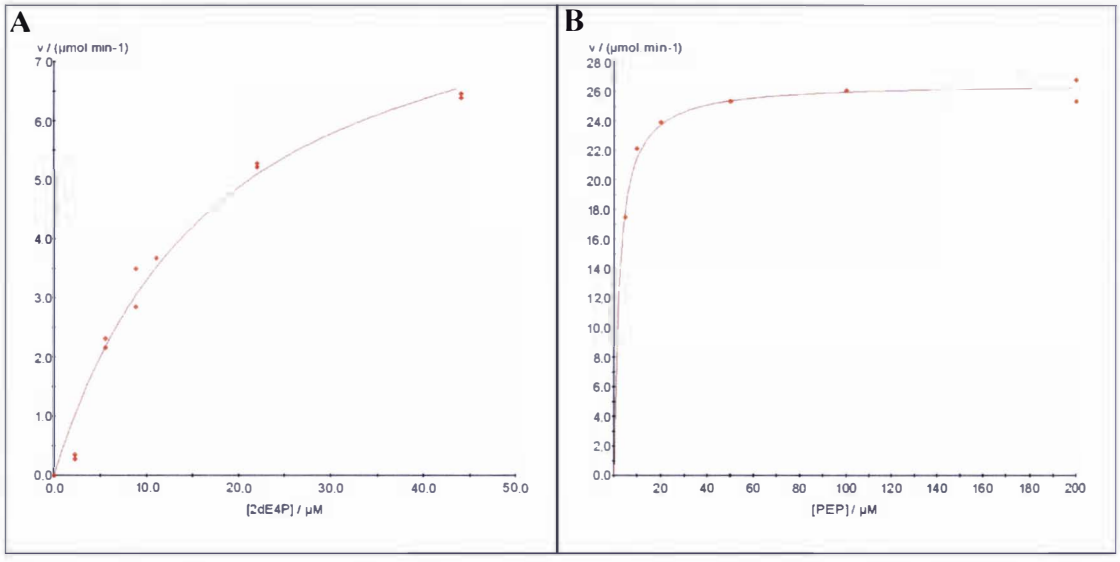


Figure 2.17 Michaelis-Menten plots for determination of K_M values for 2dE4P A, and PEP B. Initial velocity values were determined as a function of 2dE4P A, and PEP B, concentration at fixed concentrations of PEP A, and 2dE4P B. For experimental details refer to Chapter Six.

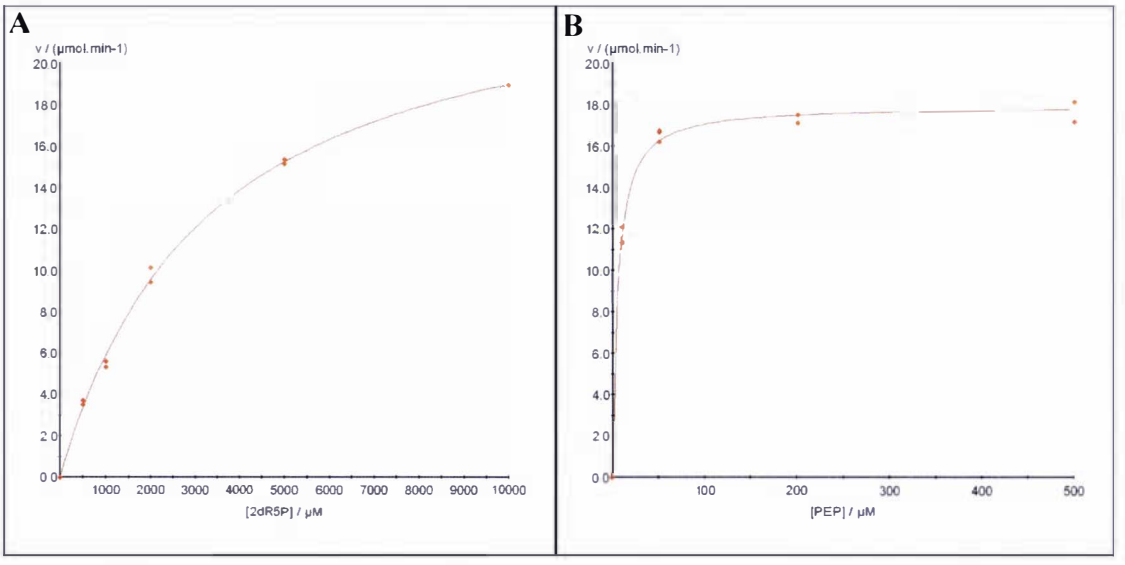


Figure 2.18 Michaelis-Menten plots for determination of K_M values for 2dR5P A, and PEP B. Initial velocity values were determined as a function of 2dR5P A, and PEP B, concentration at fixed concentrations of PEP A, and 2dR5P B. For experimental details refer to Chapter Six.

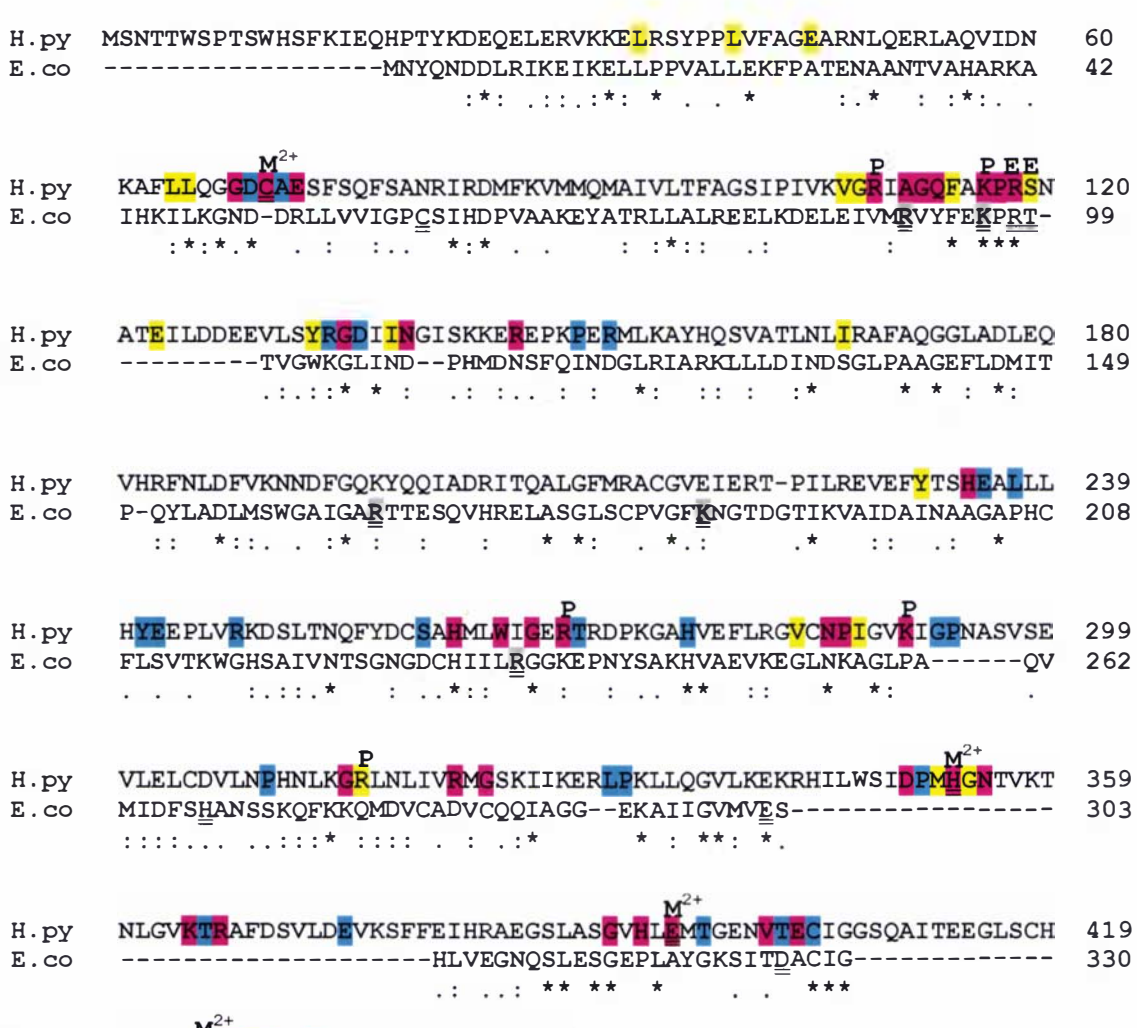
2.13 Analysis of Type II DAH7PS Sequences and Species Distribution

Bacterial genomes encode predominantly type Iα and type II DAH7PSs, with the type Iβ enzymes being largely restricted to extremophiles. Both type Iα and type II DAH7PSs are also present in eukaryotes. In some instances these are found together, *e.g. N. crassa* and *Aspergillus nidulans*, while in other eukaryotic genomes, *e.g. Toxoplasma gondii*, algae (*Cyanidioschyzon merolae*) and plants, only type II enzymes are encoded. Genes encoding type Iα or type II enzymes can be found either as singletons, exclusive paralogs, or in various combinations (table 2.5). The presence of multiple open reading frames encoding type II enzymes is typically associated with the production of secondary metabolites. Several species, including *H. pylori*, possess a single DAH7PS that is classified as a type II enzyme. This genomic evidence strongly supports a role for type II DAH7PSs in aromatic amino acid biosynthesis in these organisms.

Type Iα only	Type Iα and Type II	Type II only
E. coli {3}	*X. campestris*	*M. tuberculosis*
S. cerevisiae {2}	*S. aurantiaca* {2 II}	*H. pylori*
N. meningitidis	*Corynebacterium glutamicum*	*S. coelicolor* {2}
	P. aeruginosa {2 Iα, 3 II}	*C. diphtheriae*
	Pseudomonas putida {2 Iα}	*Streptomyces avermitilis*
	Mycobacterium smegmatis	*T. gondii*
	N. crassa	*Arabidopsis thaliana* {3}

Table 2.5 **Occurrence and multiplicity of type Iα and type II DAH7PS in selected organisms.** Organisms with multiple ORFs encoding type Iα or type II DAH7PSs have the number of paralogs indicated in parentheses {}.

A ClustalX[164] alignment of ninety type II DAH7PSs was generated by Dr Mark Patchett (figure 2.19, top line) (see Appendix One for species names and gene identifiers) which shows that these enzymes have several sequence motifs that are absolutely conserved across phylogenetically diverse species (*e.g.* Bacteria, fungi, *T. gondii*, plants). The identification of conserved residues within the type II family also provides insight into the relationship between the two enzyme families. The overall sequence identity between type I and type II DAH7PSs is low;[36] a ClustalW multiple alignment[164] (figure 2.19) of *Hp*-DAH7PS (top line) and *Ec*-DAH7PS(Phe) (bottom line) yields only 5 % sequence identity. However, a comparison of the type II alignment with sequence and structural information available for type Iα DAH7PSs shows that key residues implicated in metal, PEP and E4P binding are conserved between the two enzyme families (figure 2.19 and table 2.5). These residues are found in the same order and relative spacing in the primary sequence in both enzyme types, and are found in regions of relatively high sequence similarity, or conservation.


H.py MSNTTWSPTSWHSFKIEQHPTYKDEQELERVKKELRSYPPLVFAGEARNLQERLAQVIDN 60
E.co ----------MNYQNDDLRIKEIKELLPPVALLEKFPATENAANTVAHARKA 42
H.py KAFLLQGGDCAESFSQFSANRIRDMFKVMMQMAIVLTFAGSIPIVKVGRIAGQFAKPRSN 120
E.co IHKILKGND-DRLLVVIGPCSIHDPVAAKEYATRLLALREELKDELEIVMRVYFEKPRT- 99
H.py ATEILDDEEVLSYRGDIINGISKKEREPKPERMLKAYHQSVATLNLIRAFAQGGLADLEQ 180
E.co ---------TVGWKGLIND--PHMDNSFQINDGLRIARKLLLDINDSGLPAAGEFLDMIT 149
H.py VHRFNLDFVKNNDFGQKYQQIADRITQALGFMRACGVEIERT-PILREVEFYTSHEALLL 239
E.co P-QYLADLMSWGAIGARTTESQVHRELASGLSCPVGFKNGTDGTIKVAIDAINAAGAPHC 208
H.py HYEEPLVRKDSLTNQFYDCSAHMLWIGERTRDPKGAHVEFLRGVCNPIGVKIGPNASVSE 299
E.co FLSVTKWGHSAIVNTSGNGDCHIILRGGKEPNYSAKHVAEVKEGLNKAGLPA------QV 262
H.py VLELCDVLNPHNLKGRLNLIVRMGSKIIKERLPKLLQGVLKEKRHILWSIDPMHGNTVKT 359
E.co MIDFSHANSSKQFKKQMDVCADVCQQIAGG--EKAIIGVMVES---------------- 303
H.py NLGVKTRAFDSVLDEVKSFFEIHRAEGSLASGVHLEMTGENVTECIGGSQAITEEGLSCH 419
E.co --------------------HLVEGNQSLESGEPLAYGKSITDACIG---------- 330
H.py YYTQCDPRLNATQALELAFLIADMLKKQRT 449
E.co -WEDTDALLR---------QLANAVKARRG 350

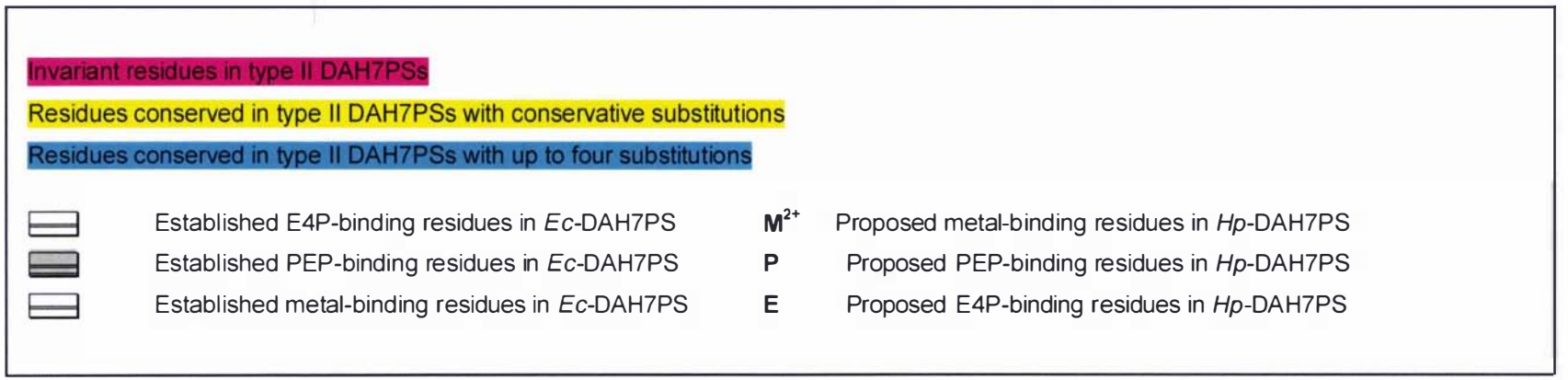


Figure 2.19 ClustalW alignment of *Hp*-DAH7PS (top line) and type Iα *Ec*-DAH7PS (bottom line). The pattern of invariant residues within the type II DAH7PSs is shown in relation to the *H. pylori* J99 DAH7PS sequence (GI:15611192, top line). Invariant residues were extracted from a multiple alignment (ClustalX) of ninety type II sequences (the organisms and, where available, the gene identifiers are in Appendix One).

Role of Residue		Crystallographically observed residues in *Ec*-DAH7PS(Phe)	Proposed residues in *Hp*-DAH7PS
Metal binding		Cys61 His268 Glu302 Asp326	Cys70 His353 Glu395 Asp425
PEP binding	**Carboxylate group**	Arg92 Lys97 Lys186	Arg109 Lys116 Lys290
	Phosphate group	Arg165 Lys186 Arg234	Arg268 Lys290 Arg321
E4P binding (Phosphate group)		Arg99 Thr100	Arg118 Ser119

Table 2.6 Table showing established metal and substrate binding residues of *Ec*-DAH7PS and equivalent proposed residues of *Hp*-DAH7PS

A subgroup of sequences with an approximately 45 amino acid deletion between residues 160 and 201 (*H. pylori* numbering) can be identified.[38] This deletion appears to be a characteristic of DAH7PSs involved in secondary metabolite biosynthesis where either chorismate or shikimate is a precursor; for example, the type II enzymes from phenazine-producing *P. aeruginosa* (*Pa*-DAH7PS Phz)[165] (figure 2.20 shows alignment of *Hp*-DAH7PS with *Pa*-DAH7PS Phz) and ansatrienin-producing *S. collinus*.[141] However, if the secondary metabolite precursor is derived exclusively from the

aminoshikimate pathway via AHBA (refer to Chapter One for more detail), then the associated aminoDAH7PS is substantially full length. The region corresponding to *H. pylori* residues 160 to 201 shows relatively poor amino acid sequence conservation in full-length type II DAH7PSs. It has been suggested that this region dictates allosteric specificity towards chorismate and tryptophan.[38] Interestingly, *Hp*-DAH7PS is full-length and not feedback regulated by any of the aromatic amino acids or chorismate (detailed in Chapter Four). The insensitivity towards Trp, Phe, Tyr and chorismate does not infer that this enzyme is unregulated it may simply mean that the metabolite(s) responsible for regulation have yet to be identified.

Figure 2.20 A ClustalW (1.83) sequence alignment of *Hp*-DAH7PS (top line) with *Pa*-DAH7PS (bottom line). Deletion (Δ160-201, *H. pylori* numbering) found in *Pa*-DAH7PS Phz.

```
H.py  MSNTTWSPTSWHSFKIEQHPTYKDEQELERVKKELRSYPPLVFAGEARNLQERLAQVIDN 60
P.ae  MDDLLQRVR---RCEALQQPEWGDPSRLRDVQAYLRGSPALIRAGDILALRATLARVARG 57
      *.:           :  *:* : * ..*. *:  **. *.*: **:   *:  **:*  .

H.py  KAFLLQGGDCAESFSQFSANRIRDMFKVMMQMAIVLTFAGSIPIVKVGRIAGQFAKPRSN 120
P.ae  EALVVQCGDCAEDMDDHHAENVARKAAVLELLAGALRLAGRRPVIRVGRIAGQYAKPRSK 117
      :*:::* *****.:.:. *:.:     *:  :* .* :**  *:::*******:*****:

H.py  ATEILDDEEVLSYRGDIING--ISKKEREPKPERMLKAYHQSVATLNLIRAFAQGGLADL 178
P.ae  PHEQVGEQTLPVYRGDMVNGREAHAEQRRADPQRILKGYAA------------------- 158
      . * :.:: :  ****::**     ::*...*:*:**.*

H.py  EQVHRFNLDFVKNNDFGQKYQQIADRITQALGFMRACGVEIERTPILREVEFYTSHEALL 238
P.ae  ----------------------ARNIMRHLGWDAASGQEANASPVW------TSHEMLL 189
                           * .* : **:  *.* * : :*:       **** **

H.py  LHYEEPLVRKDSLTNQFYDCSAHMLWIGERTRDPKGAHVEFLRGVCNPIGVKIGPNASVS 298
P.ae  LDYELSMLREDEQR-RVYLGSTHWPWIGERTRQVDGAHVALLAEVLNPVACKVGPEIGRD 248
      *.** .::*:*.   :.*  *:*  *******: .**** :*  * **:. *:**: . .

H.py  EVLELCDVLNPHNLKGRLNLIVRMGSKIIKERLPKLLQGVLKEKRHILWSIDPMHGNTVK 358
P.ae  QLLALCERLDPRREPGRLTLIARMGAQKVGERLPPLVEAVRAAGHPVIWLSDPMHGNTIV 308
      ::* **: *:*:.  ***.**.***:: : **** *::.*    : ::*  *******:

H.py  TNLGVKTRAFDSVLDEVKSFFEIHRAEGSLASGVHLEMTGENVTECIGGSQAITEEGLSC 418
P.ae  APCGNKTRLVRSIAEEVAAFRLAVSGSGGVAAGLHLETTPDDVTECVADSSGLHQV--SR 366
      :  * *** . *: :** :*     ..*.:*:*:*** * ::****:..*..: :   *

H.py  HYYTQCDPRLNATQALELAFLIADMLKKQRT-------- 449
P.ae  HYTSLCDPRLNPWQALSAVMAWSGAEAIPSATFPLETVA 405
```

* Conserved residues	: Conservative substitutions	. Semi-conservative substitutions

2.14 Conclusions

These studies indicate that the *H. pylori* type II DAH7PS catalyzes a metal-dependent, ordered sequential reaction with defined stereochemistry as seen for type I DAH7PSs. The results are consistent with a reaction mechanism where E4P binds to the PEP-DAH7PS complex and phosphate is released from the enzyme before DAH7P. These results suggest that catalysis by type I and type II enzymes occurs on a similar active site scaffold and that the two DAH7PS families may indeed be distantly related. Classification of the evolutionary relationship between the two DAH7PS families requires structural information for type II enzymes. The crystal structure of the type II DAH7PS from *M. tuberculosis* has been determined as part of the studies described in this thesis, and is discussed in Chapter Three.

Note:

The work in this chapter has been published in a modified form in Biochemical Journal. A copy of the paper is included in the back of this thesis.

CHAPTER THREE

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THE TYPE II DAH7PS FROM *M. TUBERCULOSIS*

3.1 Background

Mycobacterium tuberculosis is the causative agent of tuberculosis (TB),[166] one of the most serious health threats facing mankind. Worldwide deaths total 2-3 million annually,[167] more than for any other infectious disease, and it is estimated that approximately one third of the world's population is infected. Although effective anti-TB drugs exist, the long treatment times required, the problems of latent or persistent TB,[168] and the proliferation of multidrug-resistant strains of *M. tuberculosis* have all created an urgent need for the development of new antimycobacterial agents.

Recent gene disruption studies have shown that operation of the shikimate pathway is essential for the viability of *M. tuberculosis*,[169] validating the choice of enzymes from this pathway as targets for the development of novel antimycobacterial agents.[169] The presence of a type II DAH7PS in *M. tuberculosis* makes it a prime target for study. Not only is DAH7PS a key drug target in its own right, given its crucial role in the biosynthesis of many key aromatic molecules, but the exploitation of differences between the two types of DAH7PS may also enable the development of narrow-spectrum antibiotics.

3.2 Introduction

As was observed for the *H. pylori* enzyme, *Mt*-DAH7PS is insoluble when expressed in *E.coli* BL21(DE3) cells. Prior to the start of this project, Dr Shaun Lott and Dr Linley Schofield had attempted to solubilize the *Mt*-DAH7PS enzyme and had failed to find conditions which gave any soluble protein or enzyme activity. Following the successful solubilization of *Hp*-DAH7PS we decided to further investigate *Mt*-DAH7PS as part of this project. The *M. tuberculosis* protein was solubilized using a similar procedure, involving the *E. coli* GroEL/GroES chaperonins, to that described for *Hp*-DAH7PS in Chapter Two.

The functional characterization of *Hp*-DAH7PS (Chapter One) has revealed that type I and type II enzymes share a common kinetic mechanism, however structural information on type II enzymes is necessary to establish the evolutionary relationship between both families. Attempts to crystallize *Hp*-DAH7PS were unsuccessful and the reproducible crystallization of *Mt*-DAH7PS (as described in this chapter) has allowed us to gain valuable structural information about a type II enzyme by determining the crystal structure of *Mt*-DAH7PS.

This chapter describes the expression, purification, crystallization and crystal structure determination of DAH7PS from *M. tuberculosis*. This is the first structure of a type II DAH7PS and its comparison with the type I enzymes is discussed. The structural information of a type II enzyme has allowed us to establish the evolutionary relationship between the two families of DAH7PS.

3.3 Expression and Solubilization of *Mt*-DAH7PS

The cloning of the *Mt*-DAH7PS gene into a pProEx-HTa vector using *Nco*I and *Sac*I restriction sites was performed prior to the start of this PhD, by Dr Shaun Lott, University of Auckland. pProEx-HTa is a standard expression vector incorporating an N-terminal tobacco etch virus (TEV) protease-cleavable His_6-tag, which can be exploited to purify the protein of interest using immobilized metal affinity chromatography (IMAC). The expression of *Mt*-DAH7PS in *E. coli* BL21(DE3) cells produced protein of the appropriate molecular weight as seen by SDS-PAGE (figure 3.1). However, all the protein was insoluble, or in the pellet fraction after lysis and no DAH7PS activity was detected.

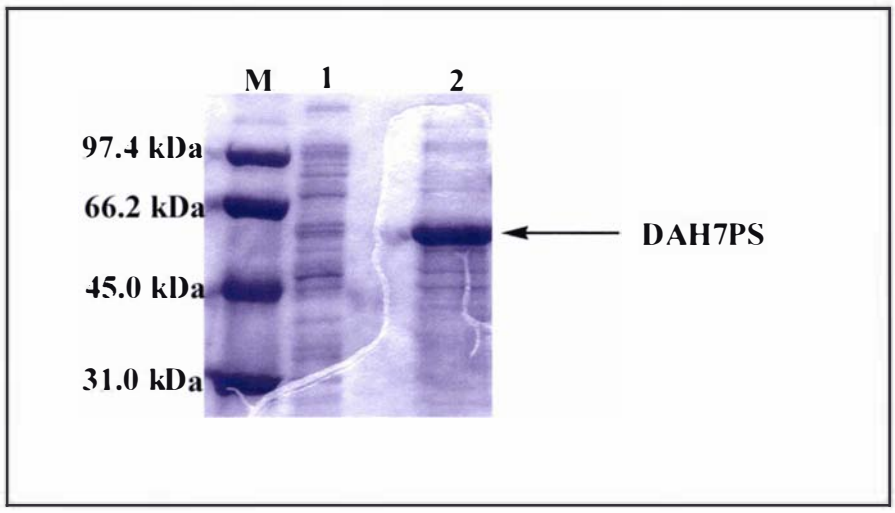


Figure 3.1 SDS-PAGE analysis of *Mt*-DAH7PS in supernatant and pellet fractions after cell lysis. SDS-PAGE analysis was performed on a 12 % polyacrylamide gel and visualized with Coomassie Brillant Blue R250. The samples were re-suspended in lysis buffer and sonicated. The ruptured cells were then centrifuged and the supernatant (lane 1) and pellet (lane 2) were analyzed.

Similar to what was observed with *Hp*-DAH7PS, the alteration of growth temperature and lysis conditions had no obvious effect on protein solubility (personal communication). Therefore, co-expression of *Mt*-DAH7PS with *E. coli* GroEL/GroES was performed, as described for *Hp*-DAH7PS in Chapter Two. pProEx-HTa-*Mt*-DAH7PS, was electroporated into BL21(DE3) cells that had been previously transformed with pGroESL.[170] Transformants were then grown in Luria broth supplemented with ampicillin (100 μg/mL) and chloramphenicol (25 μg/mL). The growth temperature, IPTG concentration and time of harvesting were identical to that described for *Hp*-DAH7PS (refer to Chapter Two, Section 2.3). Expression trials with 5 mL cultures were performed at two, four, six and eight hours after induction with IPTG (figure 3.2). There appeared to be no significant change in expression levels over time, although six hours after induction looked like it gave slightly higher levels of recombinant protein. It was noticed that when cultures were up-scaled (5 mL to 4 L) the expression levels of *Mt*-DAH7PS improved, as can be seen in lane 1 of figure 3.2.

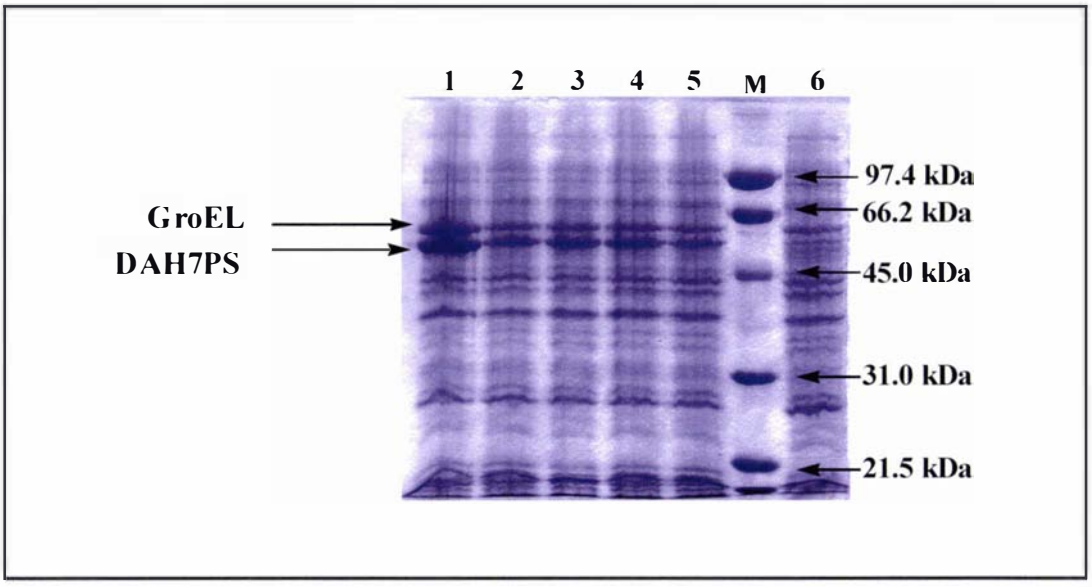


Figure 3.2 **SDS-PAGE analysis of *Mt*-DAH7PS and GroEL production at various time points.** SDS-PAGE analysis was performed on a 12 % polyacrylamide gel and visualized with Coomassie Brillant Blue R250. Whole cells (volume = 500 µL / $0D_{600}$) were re-suspended in lysis buffer and boiled prior to application. Lane 6 is whole cells which have not been induced. Lane 5 is whole cells which were induced 2 hours after induction, lane 4: 4 hours, lane 3: 6 hours, lane 2: 8hours after induction. Lane 1 is whole cells harvested from the large (4 L) growth of *Mt*-DAH7PS 6 hours after induction.

The co-expression of *Mt*-DAH7PS with the *E. coli* chaperonins resulted in soluble protein, as seen by protein of the appropriate molecular weight in the supernatant fraction (figure 3.3). This supernatant was also found to exhibit DAH7PS activity. SDS-PAGE analysis of the supernatant and pellet fractions after cell lysis showed that only 10 – 20 % of total protein expressed was present in the supernatant fraction (figure 3.3, lane 2). However, this was a considerable improvement to previous conditions that gave no soluble protein, so we continued on with purification of this enzyme.

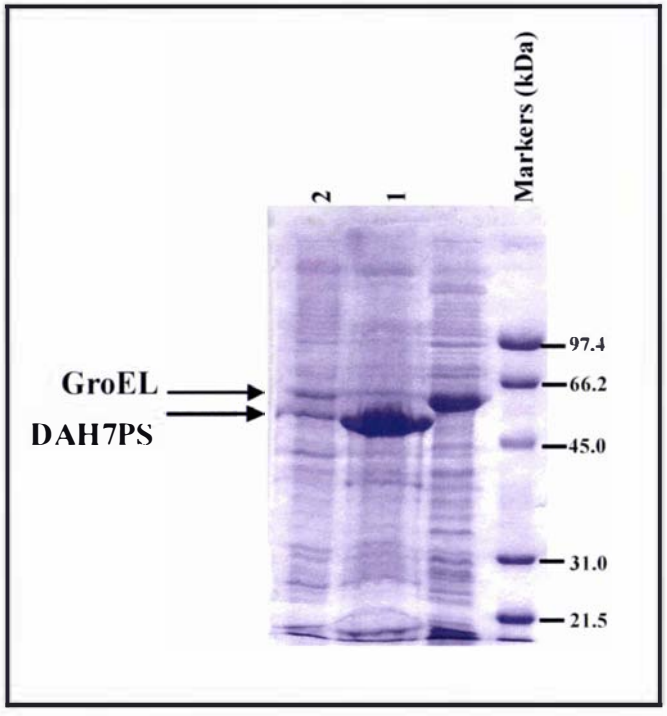


Figure 3.3 SDS-PAGE analysis of *Mt*-DAH7PS and GroEL in supernatant and pellet fraction after cell lysis. SDS-PAGE analysis was performed on a 12 % polyacrylamide gel and visualized with Coomassie Brillant Blue R250. Whole cells were lysed and clarified using centrifugation. The supernatant fraction was loaded directly into the well and the pellet fraction was resupended in lysis buffer and boiled prior to loading into the well. Lane 1 is the pellet fraction, and lane 2 the supernatant fraction, after cell lysis.

3.4 Purification of *Mt*-DAH7PS

A simple two-step purification procedure exploiting the His_6-tag at the N-terminus of *Mt*-DAH7PS, was developed. The results are summarized in table 3.1. The cells were re-suspended in lysis buffer (BTP (20 mM), NaCl (150 mM) pH 7.5 buffer) and lysed by sonication for small volumes (5 mL or less) and cell disruption for larger volumes. The cell debris was removed using centrifugation and the supernatant was filtered (0.8 μM) and passed down a chelating sepharose column charged with Ni^{2+} ions. *Mt*-DAH7PS was eluted from the column by using an imidazole gradient (figure 3.4). The protein was observed to elute at approximately 200 mM imidazole. This purification step yielded very pure (>95 %) protein as estimated by SDS-PAGE (figure 3.5). Purification via passage down a Ni^{2+} column also successfully separated *Mt*-DAH7PS away from the *E. coli* chaperonins, with GroEL eluting from the column before the imidazole gradient was applied (figure 3.5, lane 3).

Purification step	Total Protein (mg)	Total enzyme Activity (Units)[b]	Specific activity (Units/mg)	% Yield	Relative purity
1. Crude lysate	378	71.8	0.19	100	1.0
2. Ni^{2+} affinity	16	33.8	2.11	47.1	11.1
3. Superdex S200	3.4	9.0	2.65	12.5	13.9

Table 3.1 **Table showing two-step purification procedure of *Mt*-DAH7PS[a]**

[a]Purification of a 1 L culture of BL21(DE3)/pGroESL/pET-*Mt*-DAH7PS

[b]Determined by measuring the loss of PEP at 232 nm at 30 °C

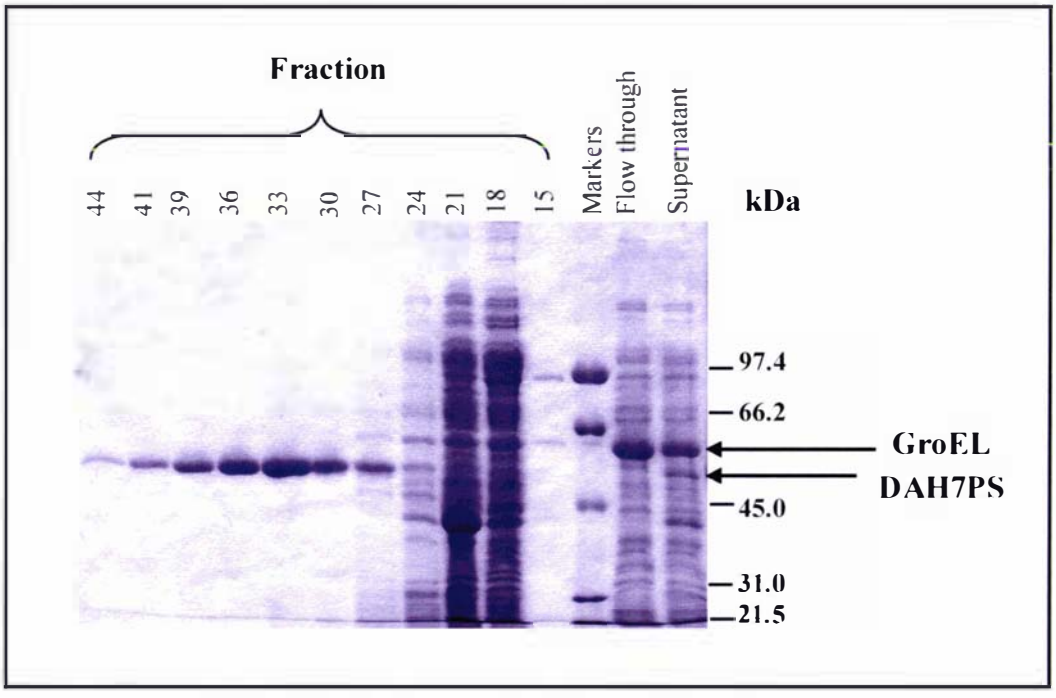


Figure 3.4 **SDS-PAGE analysis of purification by Ni^{2+} affinity chromatography.**

SDS-PAGE analysis was performed on a 12 % polyacrylamide gel and visualized with Coomassie Brillant Blue R250. 10 µL of each fraction were loaded per well.

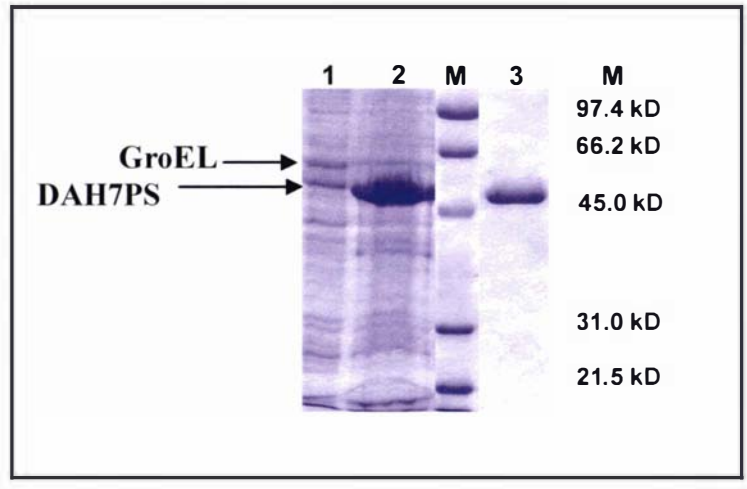


Figure 3.5 **SDS-PAGE analysis of purification of *Mt*-DAH7PS.**
Cells pelleted from a 1 L culture of IPTG-induced *E. coli* BL21(DE3) /pGroESL /pET-*Mt*-DAH7PS were re-suspended and lysed and the cell lysate clarified by centrifugation to yield a supernatant (lane 1) and pellet fractions (lane 2). Lane 3: Ni^{2+} affinity fractions pooled and concentrated. SDS-PAGE analysis was performed on a 12 % polyacrylamide gel and visualized with Coomassie Brillant Blue R250.

Fractions containing pure protein of the correct molecular weight were then pooled and the imidazole was removed by successive concentrating and washing using a 10k MWCO concentrator. This technique was used over the more tradition method of dialysis as buffer exchange this way was considerably faster allowing the purification to be performed in two days. The final step in the purification was size exclusion chromatography (SEC), using a Superdex S200 column. This was added to purify the protein further in preparation for crystallization trials. However, the chromatogram from the Superdex S200 column showed high molecular weight protein aggregates which exhibited DAH7PS activity. It has been reported for *Ec*-DAH7PS(Phe) that the inclusion of PEP stabilizes the enzyme during purification.[74] Therefore, PEP (100 μM) and $MnSO_4$ (100 μM) (which we believed may help in stabilizing the *M. tuberculosis* enzyme) were added to all purification buffers (BTP (20 mM), TCEP (0.2 mM) and NaCl (150 mM)). Thesit (polyethyleneglycoldodecyl ether) (0.005 %), a non-ionic detergent, was also included in the lysis buffer to help prevent protein aggregation. The inclusion of these additives appeared to stabilize the *M. tuberculosis* enzyme, as determined by size exclusion chromatography (described in Section 3.8). There is a significant loss of protein in going from purification step 2 to step 3 (table 3.1), this is

because the protein tended to precipitate from solution while performing buffer exchange. This was overcome slightly in future purifications by carefully monitoring protein concentration using Bradford determination. Concentrations any higher than 5 mg/mL appeared to cause protein precipitation. This procedure yielded approximately 4 mg of homogeneous *Mt*-DAH7PS from a 1 L culture.

3.5 Cleavage of the N-terminal His_6-tag

To investigate whether the His_6-tag at the N-terminus of *Mt*-DAH7PS was interfering with enzyme catalysis and crystallization of the protein, the affinity tag was cleaved off protein that had been purified by Ni^{2+} affinity chromatography. Cleavage of the His_6-tag was performed by incubation with recombinant tobacco etch virus (TEV) protease overnight, and SEC successfully removed the TEV protease and the cleaved His_6-tag. The completeness of the cleavage by TEV protease was followed using 7.5 % SDS-PAGE, which was able to resolve cleaved and uncleaved protein. The His_6-protein had similar specific activity to that of protein with the affinity tag cleaved off. Crystallization trials were performed on cleaved and uncleaved protein preparations using the robotics system at the University of Auckland (described in further detail in Section 3.12). Several conditions produced crystals of the cleaved protein, however protein with the His_6-tag still attached did not crystallize under any conditions. Therefore, successive purifications included cleaving the His_6-tag off with TEV protease before SEC.

3.6 Requirement for a Reducing Agent

Similar to the observations made with *Hp*-DAH7PS, the purification of *Mt*-DAH7PS required the inclusion of reducing agent in all purification buffers to maintain full enzyme activity (discussed in further detail in Chapter Five). The reducing agent used for *Hp*-DAH7PS was DTT, however this is not suitable for purification by Ni^{2+} affinity chromatography as it creates thiol reducing complexes which make the resin unstable. Instead, tris (2-carboxyethyl) phosphine (TCEP)[171] was used in the purification of *Mt*-DAH7PS due to its known compatibility with chelating sepharose resin.

Further studies were performed to show that the addition of PEP or EDTA alone do not stop inactivation of *Mt*-DAH7PS, which has been reported to stabilize *Ec*-DAH7PS.[87] *Mt*-DAH7PS protein purified in the presence of TCEP, was simultaneously passed down two identical NAP-5 columns, one equilibrated with buffer containing no TCEP and the second with reducing agent. The buffer consisted of BTP (20 mM, pH 7.0), NaCl (150 mM), PEP (200 μM) and $MnSO_4$ (200 μM). Columns were run with and without TCEP under several further conditions; (1) with EDTA (1 mM) substituting $MnSO_4$ and (2) with no PEP. The specific activity (μmol/min/mg protein) of enzyme eluted from the column with no TCEP was approximately one-third of that of protein passed down a column equilibrated in buffer containing TCEP. To determine whether the loss in specific activity could be restored upon addition of TCEP, both enzyme preparations (with and without TCEP) were then put down a second column both in the presence of TCEP. This resulted in an increase in specific activity to the level of that seen for protein kept in TCEP. The presence of EDTA and the absence of PEP appeared to have minimal effect on enzyme activity. This suggests that the maintenance of full enzyme activity is absolutely dependent on the presence of reducing agent.

3.7 Determination of Quaternary Structure in Solution

To determine the quaternary structure of *Mt*-DAH7PS SEC was used with protein standards of known molecular mass (refer to Chapter Six for a description of protein standards and their molecular mass). The quaternary structure was estimated off a standard curve of log molecular mass of protein standards against elution time from the Superdex S200 column. *Mt*-DAH7PS was estimated to be approximately 200 kDa, which is consistent with it being tetrameric in solution (figure 3.6). This is in contrast to what is observed for the type II enzyme from *H. pylori*, which is predicted to be a dimer (refer to Chapter Two, Section 2.6 for more details).

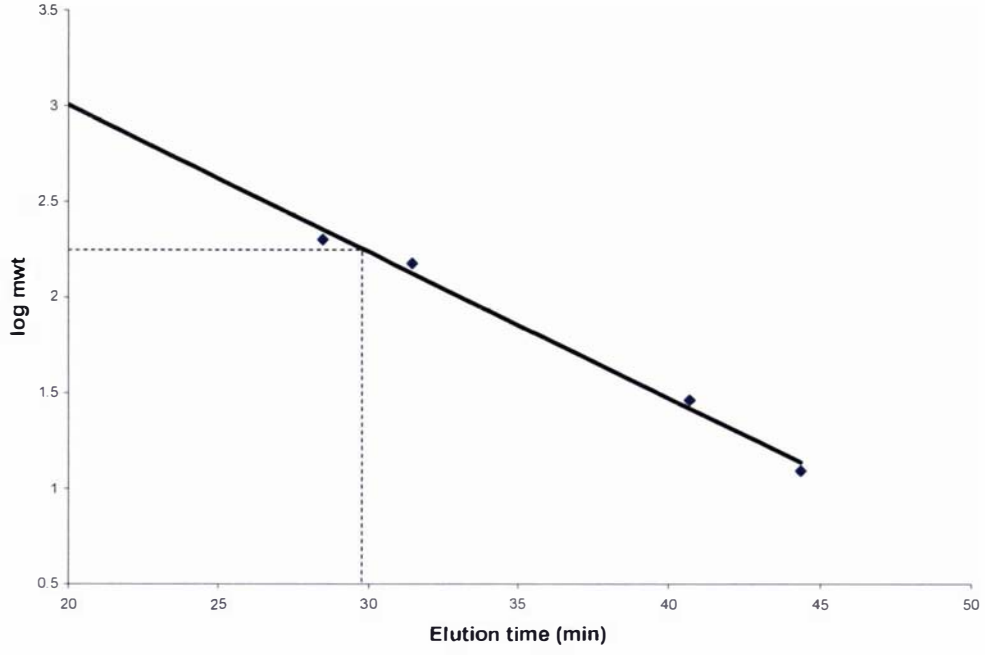


Figure 3.6 Standard curve of log molecular mass versus elution time. The molecular mass of *Mt*-DAH7PS was determined by extrapolation from a standard curve of the log molecular mass of protein standards against their elution times. The buffer and flow rate used for the protein standards was identical to that used for *Mt*-DAH7PS. Refer to Chapter Six for experimental details.

3.8 Stabilization of *Mt*-DAH7PS by PEP and $MnSO_4$

To investigate the influence PEP and $MnSO_4$ have on the quaternary structure of *Mt*-DAH7PS protein was purified in the presence of EDTA (and no PEP). Size exclusion columns were performed with this protein and with protein incubated overnight with $MnSO_4$ (~70 μM) and PEP (~70 μM). The chromatograms (figure 3.7) clearly demonstrate that the binding of PEP and Mn^{2+} in the active site of *Mt*-DAH7PS stabilizes the protein by shifting the equilibrium from higher molecular weight protein aggregates to tetramer. The protein that was eluted off the column as 'aggregates' is active but this is expected if there is an equilibrium in place. To investigate whether the change in quaternary structure is associated with loss of enzyme activity protein that was purified in the presence of PEP and $MnSO_4$ was incubated overnight in EDTA (10 mM). Incubation with the metal-chelator resulted in protein with specific activity of approximately half that observed for protein left in $MnSO_4$. Interestingly, the loss of activity could be completely restored by incubation overnight with $MnSO_4$ (~10 mM).

This suggests that the higher molecular weight protein aggregates that we are observing by SEC are either inactive or retain minimal enzyme activity.

PEP and $MnSO_4$ appear to have no significant effect on the quaternary structure of *Hp*-DAH7PS, as there appears to be no difference in how the samples run on the Superdex S200 column in the presence or absence of PEP and $MnSO_4$.

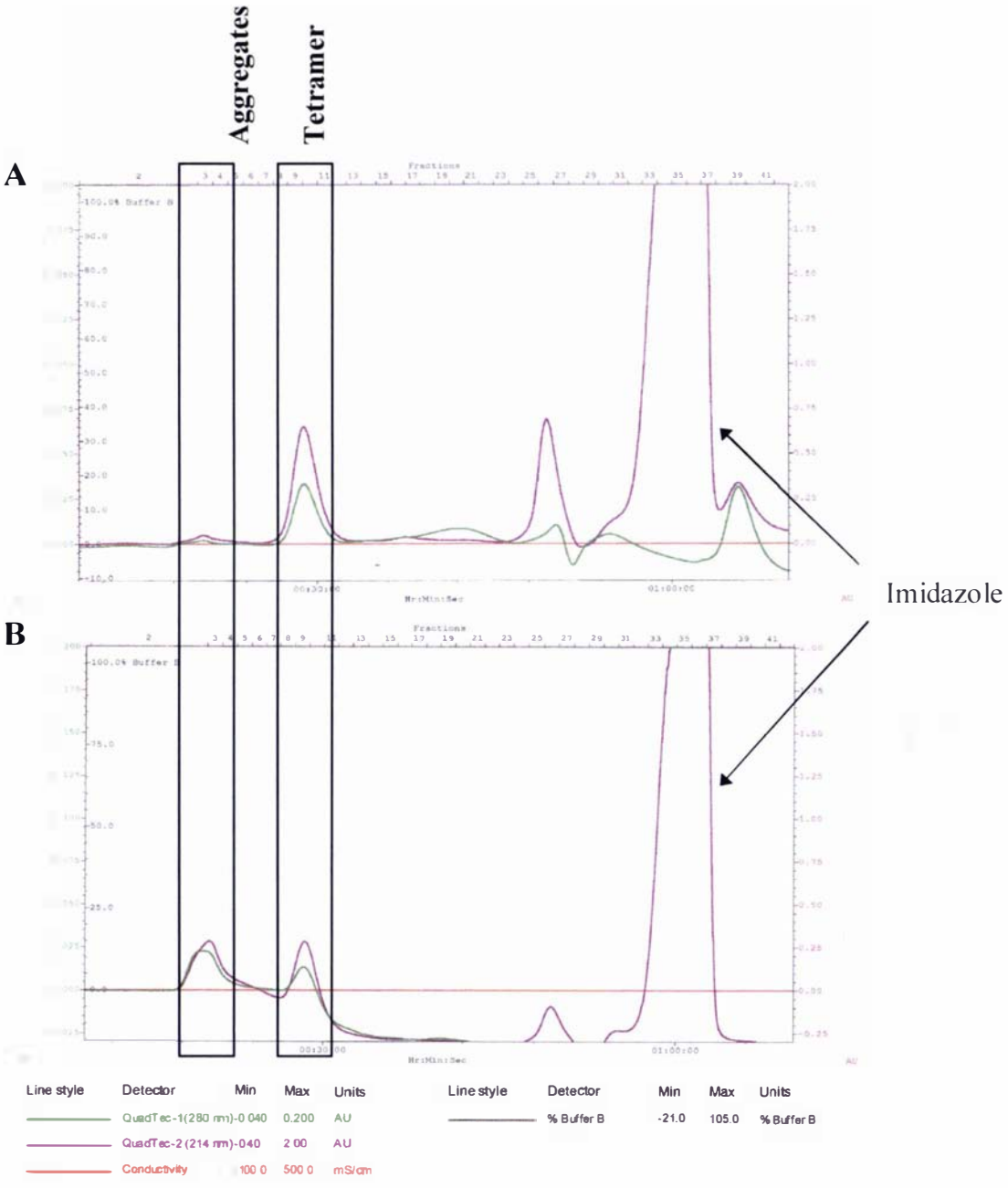


Figure 3.7 Chromatogram trace of Superdex S200 run (A) in the presence of PEP and $MnSO_4$, and (B) in EDTA

3.9 Initial Kinetic Parameters for *Mt*-DAH7PS

The apparent K_M values for E4P and PEP were 17.0 ± 1.5 μM and 11 ± 0.6 μM, respectively and the k_{cat} value was calculated as 2.5 ± 0.03 s^{-1} (figure 3.8). These values are similar to those obtained for the *H. pylori* enzyme (refer to Chapter Two, Section 2.10).

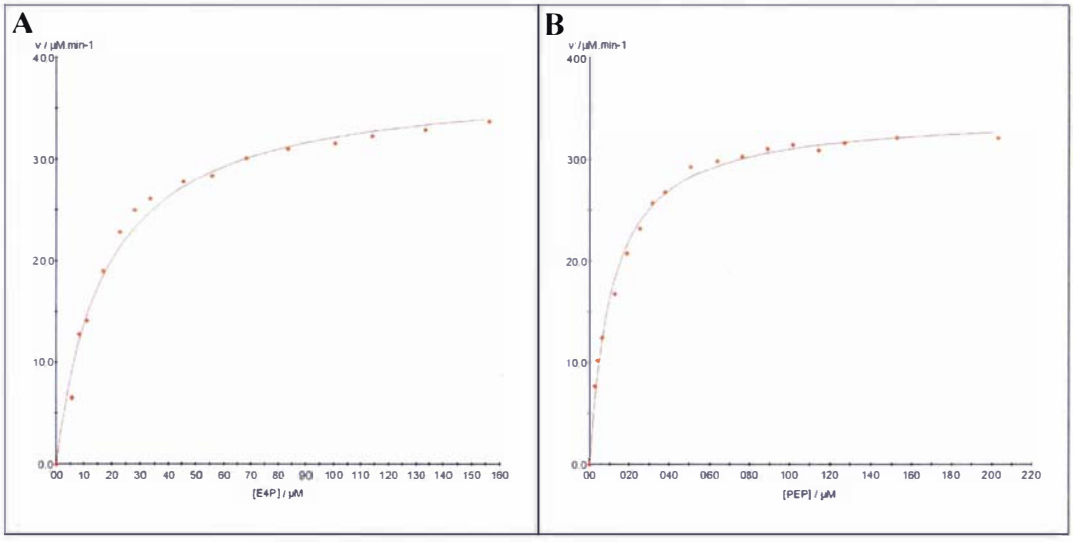


Figure 3.8 Michaelis-Menten plots for determination of K_M values for E4P (A) and PEP (B). The reaction mixtures for the determination of the K_M of E4P consisted of PEP (300 μM), $MnSO_4$ (100 μM) and E4P (6 to 156 μM), in BTP (50 mM) pH 7.5 buffer. The determination of the K_M of PEP consisted of reaction mixtures of E4P (140 μM), $MnSO_4$ (100 μM) and PEP (6 μM to 200 μM), in BTP (50 mM) pH 7.5 buffer. The reaction was initiated by the addition of purified *Mt*-DAH7PS (2 μL, 5.8 mg/mL) and carried out at 30 °C. K_M and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using Enzfitter (Biosoft).

3.10 Metal Dependency

Like all other DAH7PSs characterized to date, *Mt*-DAH7PS was sensitive to the presence of EDTA. Incubation with the metal-chelator (EDTA, 250 mM pH 8.0, 4 °C) took approximately twenty-four hours to completely inactivate the enzyme. Buffer and substrate solutions were pre-treated with Chelex to minimize the presence of metal ions from other sources. Activity of the apoenzyme could be restored by the inclusion of a variety of divalent metals in the activity assay (figure 3.9). Co^{2+} and Cd^{2+}, besides Mn^{2+}, which is generally used to assay type I DAH7PSs, gave the fastest reaction rates. Whereas, Mg^{2+}, Fe^{2+} and Ni^{2+} had no detectable effect on enzyme activity. The metal

reactivation series for the *M. tuberculosis* enzyme is comparable to that determined for *Hp*-DAH7PS. However, there is a significant difference in the time taken for EDTA to completely remove enzyme activity. While incubation with EDTA took twenty-four hours to completely inactivate the *Mt*-DAH7PS it only required incubation for thirty minutes to remove enzyme activity for the *H. pylori* enzyme. Further studies with *Mt*-DAH7PS have shown that incubation with EDTA destabilizes the quaternary structure of the enzyme (described in Section 3.8), and that full enzyme activity is restored only after twenty-four hours incubation with $MnSO_4$. Therefore, because reactivation studies were performed by the addition of metal to the reaction mixture, the metal dependency of *Mt*-DAH7PS was determined for partially activated enzyme. EDTA does not appear to have the same effect with *Hp*-DAH7PS (refer to Section 3.8).

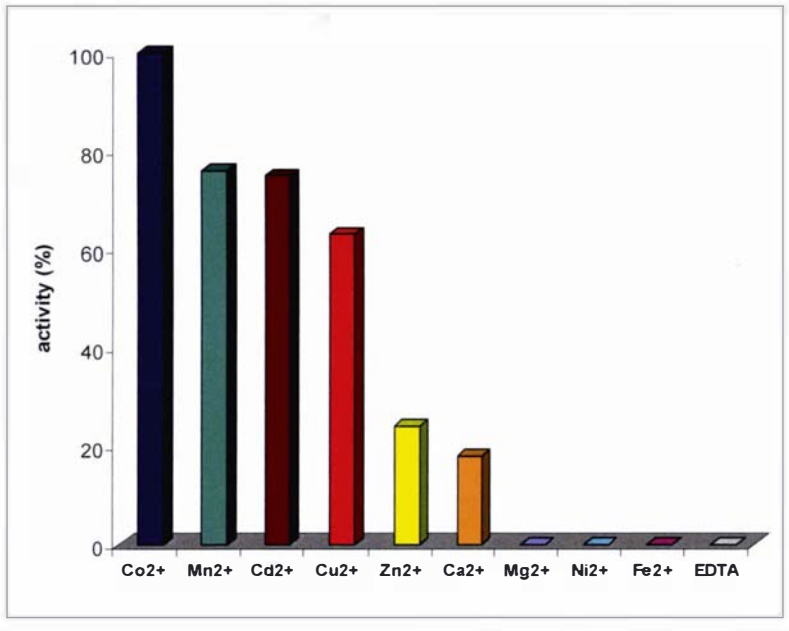


Figure 3.9 Effect of divalent metal ions on the activity of *Mt*-DAH7PS

3.11 pH Profile of *Mt*-DAH7PS

Mt-DAH7PS was active over a broad pH range with the highest activity at pH 6.4 (figure 3.10). BTP (100 mM) was used to buffer solutions of pH 6 to 9, while acetate (100 mM) was used to buffer solutions of pH 5 to 6. Activity assays were performed at pH 7.5 rather than pH 6.4 to allow comparison with other type II DAH7PSs that have been partially characterized.

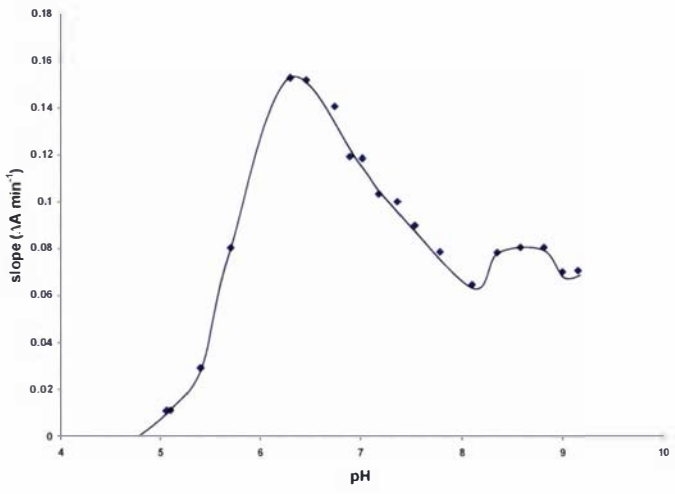


Figure 3.10 pH profile of *Mt*-DAH7PS

3.12 Crystallization of Native *Mt*-DAH7PS

Crystallization trials were performed using protein concentrations no greater than 5 mg/mL, as protein precipitation was observed at higher concentrations. Screening for crystallization conditions was achieved using sitting-drop vapour diffusion in 96-well Intelliplates (Hampton Research). The initial screens were set down using the Centre for Molecular Biodiscovery Crystallization facility, where they have set up screening for approximately five hundred different conditions. Small crystals appeared within two to three hours in several conditions, and larger needle crystals in one to two days in many different conditions. Further manual screening in twenty-four-well plates (1 μL + 1 μL drops) was performed using these lead conditions. The precipitant condition that gave the most reproducible crystals comprised 0.1 M Tris-HCl pH 8.5, 1.5 M ammonium sulfate, 12 % (v/v) glycerol (Hampton II # 42). Refinement of crystallization conditions was achieved by altering the precipitant concentration and pH. The best diffracting crystals were grown with 0.1 M Tris-HCl pH 8.0, 1.5 M ammonium sulfate, 12 % (v/v) glycerol. The crystals grew to approximately 0.3 × 0.05 × 0.05 mm (figure 3.11).

Figure 3.11 Crystals of native *Mt*-DAH7PS.
Crystal dimensions are 0.3 x 0.05 x 0.05 mm

3.13 Data Collection of Native *Mt*-DAH7PS

Native crystals were dipped momentarily in cryoprotectant solution consisting of 0.1 M Tris-HCl pH 8.0, 1.5 M ammonium sulfate supplemented with 25 % glycerol (v/v), and were flash-frozen under a stream of cold nitrogen at 110 K. X-ray diffraction data were collected using a Rigaku MicroMax007 generator with Osmic blue optics and an RAxis IV++ detector. A highly redundant native data set was collected to a maximum resolution of 2.5 Å and an overall R_{merge} of 9.9 % on intensities, using non-overlapping 0.3° oscillations collected for fifteen minutes per frame at a crystal-detector separation of 180 mm (table 3.2). Data were processed using CrystalClear 1.3.6 (Rigaku).

Table 3.2 Table showing data collection and processing statistics

Resolution range (outer shell) (Å)	39.37 – 2.50 (2.59 – 2.50)
Wavelength (Å)	1.542
Total reflections	416784
Unique reflections	54710
Completeness (%)	100 (100)
Redundancy	7.62 (7.54)
I/σ	13.6 (5.4)
R_{merge} (%)	9.9 (34.7)

The crystals were found to be trigonal, with systematic absences consistent with the space group $P3_121$ or $P3_221$. The unit-cell parameters, $a = b = 203.61$ Å, $c = 66.39$ Å, are consistent with the presence of two or three molecules in the crystal asymmetric unit, corresponding to values of the Matthews coefficient V_M of 3.9 and 2.6 $Å^3\ Da^{-1}$ and solvent contents of 68 % and 53 %, respectively. Molecular replacement was attempted by Prof Geoff Jameson using *E. coli* type I DAH7PS as a model and found to be unsuccessful.

3.14 Crystallization of Se-Met-Substituted *Mt*-DAH7PS

The next step was to try and solve the structure of *Mt*-DAH7PS by single-wavelength anomalous dispersion (SAD) methods using selenomethionine (Se-Met)-substituted protein, given that the *Mt*-DAH7PS monomer contains thirteen methionines. For the production of Se-Met protein, the plasmid bearing *Mt*-DAH7PS was transformed into cells of the methionine auxotroph DL41(DE3) that also contained pGroESL. The transformed cells were grown in minimal media with Se-Met as the only methionine source.[172] The Se-Met protein was then purified using an identical procedure to that described for the native protein (refer to Section 3.4). Se-Met-protein formed crystals using the same protein concentration and were of similar size and morphology. The best Se-Met-substituted crystals were grown in the following three conditions: 0.1 M Na-Hepes pH 7.5, 0.8 M Na,K tartrate (Hampton I #29); 0.1 M Na-Hepes pH 7.5, 1.5 M lithium sulfate (Hampton I # 16); and 0.1 M Na citrate pH 3.6, 1 M ammonium phosphate (Hampton I # 11). The crystal used for data collection (described below) was grown in 0.1 M Na-Hepes pH 7.5, 0.8 M Na,K tartrate.

3.15 Data Collection of Set-Met-Substituted *Mt*-DAH7PS

For data collection, the crystals were soaked in a cryoprotectant comprising mother liquor plus 20 % (v/v) glycerol and flash-frozen in liquid nitrogen. These crystals proved to be trigonal, with space group $P3_221$, unit cell dimensions a=b=204.1, c=66.23 Å, $\alpha=\beta=90^o$, $\gamma=120^o$, and two DAH7PS molecules in the asymmetric unit, with a solvent content of 68 %. Data were collected at a wavelength of 0.9796 Å (the peak of the Se absorption edge) on beamline 8.2 at the Advanced Light Source (Lawrence Berkeley National Laboratory, CA). The data were processed using DENZO and

SCALEPACK[173] by Minmin Yu at the Advanced Light Source (Berkeley, CA). Full details of the data collection statistics are in table 3.3.

Table 3.3 Table showing data collection and processing statistics	
Resolution range (outer shell) (Å)	50.0 – 2.30 (2.38 – 2.30)
Wavelength	0.97964 (Se peak)
Total measurements	787621
Unique reflections	70357
Completeness (%)	100.0 (100.0)
Redundancy	11.2 (10.4)
I/δ	23.9 (3.9)
R_{merge}	0.124 (0.652)

3.16 Structure Determination and Refinement of Set-Met-Substituted *Mt*-DAH7PS

The crystal structure of Set-Met-substituted *Mt*-DAH7PS was solved using SAD data collected at the peak of the Se absorption edge, and was refined at 2.3 Å resolution to a final *R* factor of 0.181 (R_{free} = 0.219). The asymmetric unit of the crystal contains two molecules. In the final model, molecule A comprises the complete polypeptide chain (residues 1-462) except for residues 422-436, which are disordered, and molecule B is similarly complete except for residues 376-378 and 418-431. Two additional residues (Gly-Ala) from the linker to the cleaved His_6-tag are also modeled for both molecules. The main-chain torsion angles molecules correspond well with allowed values, with 90.1 % of residues in the most favored region of the Ramachandran plot, as defined by PROCHECK,[174] and no outliers. The final model also contains two Mn^{2+} ions, two PEP molecules, two SO_4^{2-} ions, a molecule of the detergent Thesit and three hundred and forty-five water molecules.

The Se sites were found using SHELXD[175] which located all of the expected twenty-six sites from the anomalous differences in the Se peak data set. The space group was confirmed as $P3_221$ and the initial phases were improved and extended to 2.3 Å using

cycles of density modification in SHELXE[176] which returned a final pseudo-free correlation coefficient of 0.73. Further cycles of maximum likelihood density modification and two-fold Non-Crystallographic Symmetry (NCS) averaging, followed by automated model building in RESOLVE[177] improved the figure of merit from 0.79 to 0.84, and resulted in the building of over 60 % of the structure (566 out of 924 residues from the two DAH7PS molecules, This was all performed by Minmin Yu (Advanced Light Source,Berkeley, CA) and the remainder of the model was built by myself with the help of Heather Baker and Bryan Anderson.

The remainder of the model was built manually with TURBO FRODO (http://www.afmb.univ-mrs.fr/TURBO-) and COOT[178] during successive cycles of refinement with REFMAC 5[179] using non-crystallographic symmetry restraints for residues 20-462 of both molecules in the asymmetric unit. Ordered water molecules were added into discrete Fo-Fc electron density peaks where the height was greater than 3 σ above the mean and there were potential hydrogen bonding contacts with the surrounding structure. In each molecule, a Mn^{2+} ion, a PEP molecule and a SO_4^{2-} ion (all identified by their shapes and environments) could be added into extra density, in equivalent positions in the two molecules. Subsequent refinement indicated partial occupancies for each of these species, estimated at 0.75 for Mn^{2+} and PEP and 0.5 for SO_4^{2-}. Density for a molecule of the detergent Thesit was also found in the dimer interface between molecules A and B; its inclusion reduced the value of R_{free} by 0.003. Final refinement statistics are in table 3.4. The Se-Met crystal structure has been deposited into the Protein Data Bank (PDB) with the accession code 2B70.

Table 3.4 Table showing refinement Statistics	
Resolution range (Å)	47 – 2.30
Number of reflections (test set)	66794 (3563)
R factor (R_{free})	0.181 (0.219)
Number of non-hydrogen atoms	
Protein (2 molecules)	6931
Water	345
Mn^{2+}/PEP/SO_4^{2-}	2/40/10
Thesit	26
Average B factors (Å^2)	
Protein	43.5
Water	48.1
Mn^{2+}/PEP	41.2/44.4
Rms deviations from ideality	
Bond lengths (Å)	0.019
Bond angles (deg)	1.7
Residues in most favored region of Ramachandran plot (%)	90.1

3.17 Description of Crystal Structure of *Mt*-DAH7PS

3.17.1 Monomer Structure

Both *Mt*-DAH7PS molecules in the asymmetric unit of the crystal have the same fold, with 412 out of 462 residues matching with a root-mean-square (rms) difference in Cα atom positions of 0.42 Å; only in the N-terminal twenty-one residues is there any significant deviation. The *Mt*-DAH7PS monomer (figure 3.12) consists of a core $(\beta/\alpha)_8$ TIM barrel domain in which the eight parallel β-strands (residues 80-85, 120-127, 243-249, 277-281, 302-307, 331-337, 361-366 and 406-412) are each followed by α-helices. Two major additions decorate the barrel. At the N-terminus, an extended β-strand plays a key role in dimerization (see below) and is followed by three helices (α0a, α0b and α0c), the last of which (α0c, residues 60-78) closes off the N-terminal end of the barrel. The second addition consists of a pair of helices (α2a and α2b) that extend the α2-β3 connecting loop and pack against the outside of the barrel. The connecting loops at the

C-terminal end of the barrel, where they help to form the active site (described below), are markedly longer than those at the N-terminal end (figure 3.12B). The β2-α2 and β8-α8 loops are 37 and 32 residues long, respectively, the β7-α7 and β6-α6 loops are slightly shorter (16 and 10 residues, respectively), and the other loops are 4-8 residues in length. In contrast, at the N-terminal end of the barrel six of the eight connecting loops are short (2-7 residues), and only the α2-β3 connection, with its additional two helices, and the 19-residue α3-β4 connection, are longer.

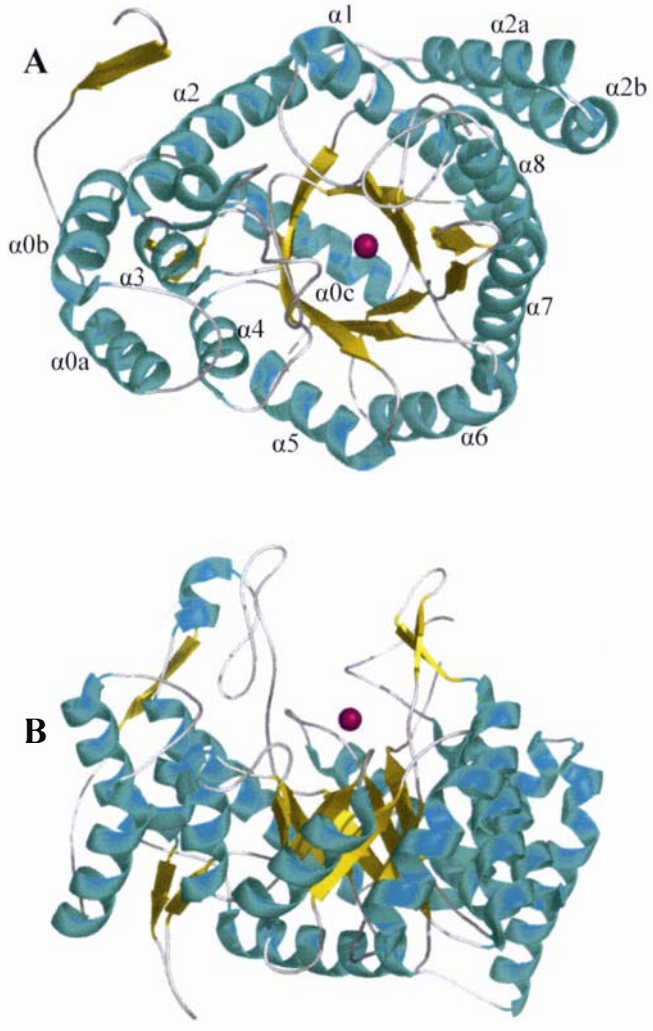


Figure 3.12 ***Mt*-DAH7PS monomer**. **A**, View of *Mt*-DAH7PS monomer, looking into the $(\beta/\alpha)_8$-barrel. α-Helices are shown in cyan and β-strands in gold. The bound Mn^{2+} ion (magenta sphere) marks the location of the active site. **B**, Side-on view of the *Mt*-DAH7PS monomer.

3.17.2 Quaternary Structure

The quaternary structure of *Mt*-DAH7PS is a homotetramer, which is generated by application of a crystallographic two-fold symmetry operation. The tetramer consists of two tight dimers of this enzyme (figure 3.14). The major contributors to the dimer interface are the extended N-terminal residues 1-13, parts of helix α0b and the α0b-α0c loop from the N-terminal addition, together with the long helix α2, residues 165-189 (a list of residues contributing to the dimer interface is given in Appendix Two). The N-termini form linked arms, with the N-terminal arm from molecule A fitting between the body of molecule B and the molecule B arm (and vice versa) and residues 3-9 from the two molecules forming an antiparallel β-ribbon (figure 3.13). This association involving the N-terminal arms contributes about 40 % of the total buried surface in the dimer. The other major contribution to dimerization comes from the two α2 helices, which cross in the interface. Overall, a total of 3782 $Å^2$ of solvent accessible surface is found to be buried by dimer formation (approximately 1900 $Å^2$ per monomer, or 10 % of the monomer surface), indicating a very stable dimer. The interface is 66 % hydrophobic, but also includes one salt-bridge ($Asp10_A...Arg171_B$) and hydrogen bonds involving the side chains of Gln11, Asn181 and Arg184, in addition to the hydrogen bonding of the N-terminal arms.

The interaction between the dimers to form the tetramer is relatively small and predominately hydrophobic, with no polar contacts between the two monomers. The formation of the weak tetramer buries 3996 $Å^2$ (1998 Å between subunit A and C (B and D) of solvent-accessible surface area (approximately 1000 $Å^2$ per monomer, or 5 % of the monomer surface). The interfacial area between the two subunits is largely through residues of the additional α2b helix and α2b-β3 loop, with minor contributions from the core α1 helix (for a list of residues contributing to the tetramer interface refer to Appendix Three).

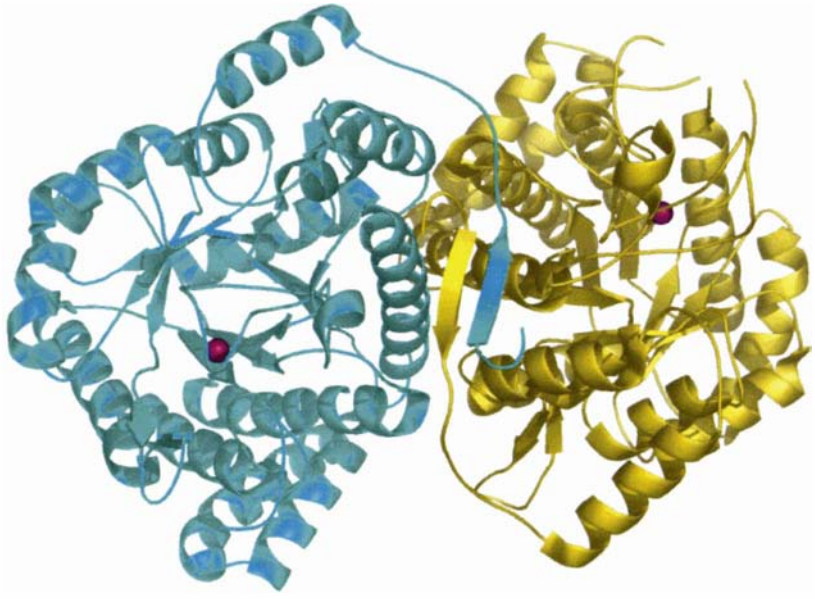


Figure 3.13 **The *Mt*-DAH7PS dimer**, with one monomer in cyan and the other in gold.

Figure 3.14 **The *Mt*-DAH7PS tetramer**, with the two monomers making up the tight dimer in different shades of the same color (either pink or purple) and the subunits making up the tetramer in contrasting colors (one monomer purple and the other pink).

3.17.3 Active Site Architecture

As is typical for $(\beta/\alpha)_8$-barrel enzymes, the active site of *Mt*-DAH7PS is located at the C-terminal end of the barrel, where it is clearly indicated by the presence of the bound metal ion cofactor, Mn^{2+}, and PEP (figure 3.12B), which are found in the same position in both subunits, at the bottom of a deep cavity. The active site residues are contributed by the C-terminal ends of the β-strands and the connecting loops between β-strands and α-helices. Residues contributing to the active site of *Mt*-DAH7PS are shown in figure 3.15.

The Metal Binding Site

The metal binding site is formed by four protein ligands, S^{γ} of Cys87 (from the β1-α1 loop), $N^{\varepsilon 2}$ of His369 (from the β7-α7 loop), $O^{\varepsilon 2}$ of Glu411 (from the end of strand β8) and $O^{\delta 2}$ of Asp441 (from the β8-α8 loop). Between them these ligands fill four positions of what most closely approximates a trigonal bipyramidal metal coordination site; Cys87 and His369 are the axial ligands, with Glu411 and Asp441 in the equatorial trigonal plane, leaving one equatorial site available for the binding of the E4P substrate or water. No water ligand is apparent in the present structure, although we have observed a water molecule in other crystal structures of *Mt*-DAH7PS discussd in this thesis. An intriguing feature of the metal binding site is the presence of Cys440 close to the metal ligand Cys87. A simple rotation about Cys440 C^{α}-C^{β} to another rotamer would allow formation of a disulfide bond with Cys87, precluding metal binding and explaining the need for a reducing agent to maintain full activity of the enzyme (refer to Chapter Five for further detail).

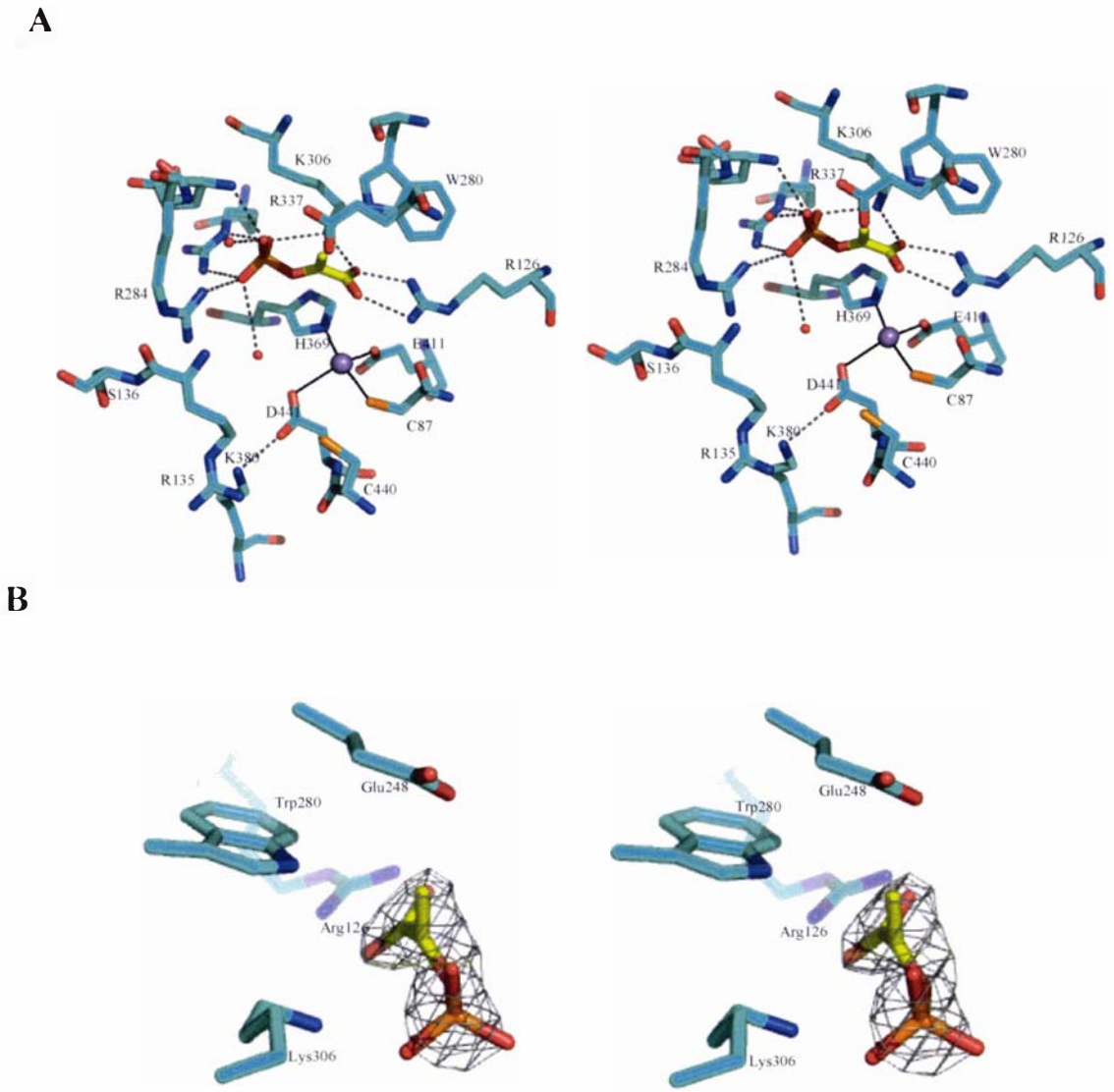


Figure 3.15 **Active site of *Mt*-DAH7PS**. **A**, stereo view showing the interactions with the Mn^{2+} ion (magenta sphere), the PEP substrate (yellow and orange stick model) and the sulfate ion (green) that marks the likely position of the phosphate group of the E4P substrate. Water molecules are shown as small red spheres. Metal-ligand bonds are shown with thin black lines, and hydrogen bonds with broken lines. Key residues that contribute to metal, PEP or sulfate binding are labeled. The invariant Glu248 would sit directly above the PEP C3 atom in this view, but has been removed for clarity. **B**, Stereo view of the electron density for PEP, from a 2Fo-Fc map, contoured at 1 σ. The side chains of Glu248 and Trp280 approach to ~3 Å from the methylene carbon (C3) of PEP.

The PEP Binding Site

The PEP binding site is defined by a network of hydrogen bonds between the protein and the PEP phosphate and carboxylate groups. The phosphate group of PEP is hydrogen bonded by the peptide NH of Glu283 and Arg284 $N^{\eta 1}$ (both located on the β4-α4 loop), Lys 306 N^{ζ} (from the end of strand β5), Arg337 N^{ε} and $N^{\eta 2}$ (from the β6-α6 loop) and two water molecules; each phosphate oxygen receives at least two hydrogen bonds. The PEP carboxylate group forms a doubly hydrogen bonded ion pair with Arg126 (from the end of strand β2), receiving hydrogen bonds from its $N^{\eta 1}$ and $N^{\eta 2}$ atoms, and is additionally hydrogen bonded to the ε-guanidine group of Lys306 and one water. Each carboxylate oxygen thus receives two hydrogen bonds of extremely favorable geometry, indicating exquisite molecular recognition. The PEP electron density (figure 3.15B) indicates a configuration in which the central C2 carbon is trigonal planar, but the carboxylate and enol groups are not coplanar, as a result of a twist of 30-40° about the C1-C2 bond. In this configuration, the side chains of Trp280 and Glu248 are on either side of the PEP methylene carbon (C3), with the Trp280 indole NH pointing at C3, ~3 Å away, and Glu248 $O^{\varepsilon 2}$ also ~3 Å away.

Potential E4P Binding Site

Also in the active site region, near its opening and ~10 Å further out from the PEP is a bound SO_4^{2-} ion. This ion, which marks the likely position of the E4P phosphate, is in the same position in both molecules, bound to Arg135 N^{ε}, Arg284$N^{\eta 2}$, Ser136 O^{γ} and the peptide NH of Ser136. Arg284 provides a bridge between the PEP and sulfate sites, being hydrogen bonded to both the PEP phosphate and the sulfate ion.

3.18 Comparison with Type I DAH7PS Enzymes

The structure determined for *Mt*-DAH7PS is a $(\beta/\alpha)_8$ TIM barrel, which is also observed for the type I DAH7PSs,[47,49,51,53,71,92,93] and KDO8PS.[103,115] We therefore superimposed the *Mt*-DAH7PS structure on to that of the type Iα *Ec*-DAH7PS,[92,93] to test the extent of structural similarity despite the lack of recognizable sequence identity. Superimposing first just the eight β-strands of the core $(\beta/\alpha)_8$ barrel, and then extending the structural alignment incrementally, with a maximum disagreement threshold of 3 Å, we could match 171 residues with an rmsd in Cα atom positions of 2.12 Å. The matching

structural elements comprised the eight β-strands, helices α2, α4, α5, part of α6, and α8, and extensive regions of the β-α connecting loops at the C-terminal end of the barrel, and included 37 sequence identities. This superposition, and a similar comparison with the type Iβ *Pf*-DAH7PS, gives the structure-based sequence alignment shown in figure 3.16. The superpositions and structure-based sequence alignment in figure 3.16 were kindly performed and constructed by Prof Ted Baker, Auckland University.

Importantly, the alignment of the type II *Mt*-DAH7PS on to the type I enzymes, purely on the basis of its fold, also aligns all the active site residues. The metal ligands (Cys87, His369, Glu411 and Asp441), the residues that interact with PEP and the SO_4^{2-} ion (Arg126, Arg135, Ser136, Arg284, Lys306 and Arg337), and other conserved residues around the active site (Pro134, Ile281 and Gly282), match the equivalent residues in the type I *Ec*-DAH7PS with an rmsd of only 1.03 Å for 111 atoms. The bound metal ion, PEP and sulfate are also within ~ 1 Å of their positions in the *E. coli* enzyme, consistent with a common reaction chemistry and mechanism. Equivalent residues in the active sites of *Ec*-DAH7PS, *Pf*-DAH7PS and *Mt*-DAH7PS are shown in table 3.5.


Mtu 1 ----------MNWTVDIPIDQLPSLPPLPTDLRTRLDAALAKPAAQQPTWPADQALAMRTVLESVPPVTVPSEI
Pfu 1 ----------MKYSKEYK-EK
Tma 1 MIVVLKPGSTEEDIRKVVKLAESYNLKCHISKGQERTVIGIIGDDRYVVADKFESLDCVESVVRVLKPYKLVSREFHPED
Eco 1 ----------MNYQNDDLRIKEIKELLPPVALLEKFPATENAANTVAHA
Sce 1 ----------MSESPMFAANGMPKVNQGAEEDVRILGYDPLASPALLQVQIPATPTSLETAKRG
α0c β1 M α1 β2 P P EE
Mtu 65 VRLQEQLAQVAKGEAFLLQGGDCAETFMDNTEPHIRGNVRALLQMAVVLTYGASMPVVKVARIAGQYAKPRSADIDALG-
Pfu 11 TVVKINDVKFG--EGFTIIAGPCSIE----SRDQIMKVAEFLAEVG----------IKVLRG--GAFKPRTSPY-----
Tma 81 TVIDLGDVKIGN-GYFTIIAGPCSVE----GREMLMETAHFLSELG----------VKVLRG--GAYKPRTSPY-----
Eco 40 RKAIHKILKG-NDDRLLVVIGPCSIH----DPVAAKEYATRLLALREELKD----ELEIVMRV--YFEKPRTTV------
Sce 55 RREAIDIITG-KDDRVLVIVGPCSIH----DLEAAQEYALRLKKLSDELKG----DLSIIMRA--YLEKPRTTV------
α2 α2a
Mtu 144 --LRSYRGDMINGFAPDAAAREHDPSRLVRAYANASAAMNLVRALTSSGLASLHLVHDWNREFVRTSPAGARYEALATE
Pfu 67 ----SFQGYG----------EKALRWMREAADEYG----------
Tma 138 ----SFQGLG----------EKGLEYLREAADKYG----------
Eco 103 ----GWKGLINDPHMDNSF ------QINDGLRIARKLLLDINDSG----------
Sce 118 ----GWKGLINDPDVNNTF-------NINKGLQSARQLFVNLTNIG----------
α2b β3 α3 β4 P α4
Mtu 221 IDRGLRFMSACGVADRNLQTAEIYASHEALVLDYERAMLRLSDGDDGEPQLFDLSAHTVWIGERTRQIDGAHIAFAQVI
Pfu 88 ----------LVTVTEVMDTRHVELVAKY----------SDILQIGARNSQ-NFELLKEVGKV
Tma 159 ----------MYVVTEALGEDDLPKVAEY----------ADIIQIGARNAQ-NFRLLSKAGSY
Eco 138 ----------LPAAGEFLDMITPQYLADLM----------SWGAIGARTTE-SQVHRELASGL
Sce 153 ----------LPIGSEMLDTISPQYLADLV----------SFGAIGARTTE-SQLHRELASGL
β5 P α5 β6 P α6
Mtu 300 ANPVGVKLGPNMTPELAVEYVERLDPHNK----------PGRLTLVSRMGNH----KVRDLLPPIVEKVQAT
Pfu 130 ENPVLLKRGMGNTIQELLYSAEYIMAQ----------GNENVILCERGIRTFETATRFTLDISAVPVVKEL
Tma 201 NKPVLLKRGFMNTIEEFLLSAEYIANS----------GNTKIILCERGIRTFEKATRNTLDISAVPIIRKE
Eco 180 SCPVGFKNGTDGTIKVAIDAINAAGAPHCFLSVTKWGHSAIVNTSGNGDCHIILRGGK--EPNYSAKHVAEVKEGLNKA
Sce 195 SFPVGFKNGTDGTLNVAVDACQAAAHSHHFMGVTKHGVAAITTTKGNEHCFVILRGGK-----KGTNYDAKSVAEAKAQ

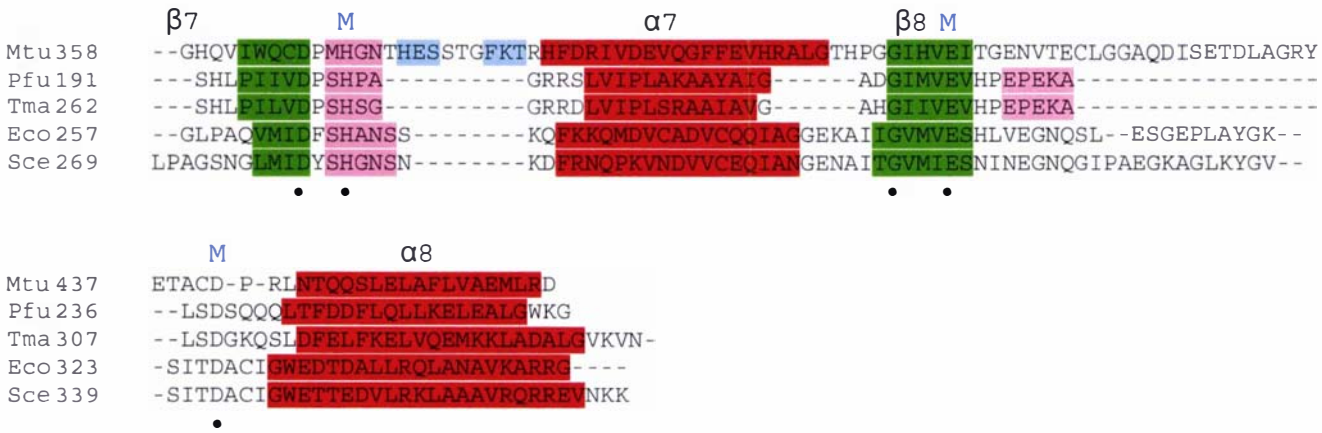


Figure 3.16 **Structure-based sequence alignment of DAH7PS enzymes**. The sequences are for the *M. tuberculosis* (Mtu), *P. furiosus* (Pfu), *T. maritima* (Tma), *E. coli* (Eco) and *S. cerevisiae* (Sce) DAH7PSs. Core β-strands are highlighted in green and core α-helices in red. Non-core α-helical regions are highlighted in pink and non-core β-strands in light blue. Residues in contact with metal or PEP are denoted with M or P respectively. The major secondary structure elements in *Mt*-DAH7PS are labeled. Residues that are conserved in all five structures are indicated with •.

Role of Residue	*Ec*-DAH7PS (type Iα)	*Pf*-DAH7PS (type Iβ)	*Mt*-DAH7PS (type II)
PEP - Phosphate binding	Ala164 Arg165 Lys186 Arg234	Ala114 Arg115 Lys136 Arg166	Glu283 Arg284 Lys306 Arg337
PEP - Carboxylate binding	Arg92 Lys97 Lys186	Arg55 Lys136 GlnIII	Arg126 Lys306
Metal binding	Cys61 His268 Glu302 Asp326	Cys31 His201 Glu227 Asp238	Cys87 His369 Glu411 Asp441
E4P - Phosphate binding	Arg99 Thr100	Arg62 Thr63	Arg135 Ser136

Table 3.5 **Table showing equivalent residues in the active sites of *Ec*-DAH7PS, *Pf*-DAH7PS and *Mt*-DAH7PS**

The only significant active site differences in *Mt*-DAH7PS appear to be the presence of the aromatic ring of Trp280 adjacent to the PEP, and the altered PEP conformation at C2-C3. Whether the contacts made by the PEP methylene carbon (C3) with Glu248 and Trp280 are significant for the reaction mechanism we cannot say at present. However, Glu248 is invariant in all DAH7PSs, and Trp280 is conserved in all type II enzymes. Trp280 replaces a Gln residue in *Pf*-DAH7PS and *Tm*-DAH7PS and Ala in *Ec*-DAH7PS and *Sc*-DAH7PS, though a Tyr residue from a different part of the structure

occupies part of this space in the latter two enzymes. One other active site difference is the presence of Lys380, which is contributed by an insertion in the *Mt*-DAH7PS sequence and has no counterpart in the Type I enzymes. Lys380 is hydrogen bonded to the metal ligand Asp441 and is 4 Å from the SO_4^{2-}, in a position where it could interact with E4P.

3.19 Discussion

DAH7PS is essential for the viability of most organisms because of its key role in the biosynthesis of aromatic compounds. Several subtypes of DAH7PS are often found in a given organism, subject to different regulation and with different metabolic roles. *M. tuberculosis*, in contrast, has only a single open reading frame (Rv2178c) that corresponds to a DAH7PS. This gives it a critical role and makes it a key drug target, since not only are the enzymes of the shikimate pathway, including DAH7PS, essential for the viability of *M. tuberculosis*,[3,180] but so are the enzymes of downstream pathways leading to the aromatic amino acids, folic acid and mycobactin.[180]

The amino acid sequence of *Mt*-DAH7PS classifies it as a type II DAH7PS. Type II DAH7PSs are found in both bacterial and eukaryotic phylogenetic domains, either in combination with type Iα enzymes or alone. Several of the microbial type II DAH7PSs have a predicted function in secondary metabolism, on the basis of genomic context,[141,143] but in organisms such as *M. tuberculosis*, where the sole DAH7PS belongs to the type II family, a pivotal role in primary metabolism is implied. The type II DAH7PSs are, however, poorly characterized compared with the type I DAH7PSs. In particular, the lack of tertiary structural information for the type II enzymes, which have no obvious sequence homology with the type I DAH7PSs, has made it unclear whether they have any evolutionary relationship.[36]

As in the case of all other DAH7PSs, *Mt*-DAH7PS activity is dependent on the presence of a divalent metal ion, and sensitive to treatment with metal ion chelators.[65,81,94,106] As was observed for the type II DAH7PSs from *H. pylori*, *X. campestris* and *S. caespitosus*,[38,118,181] highest enzyme activities were observed when the enzyme was treated with Co^{2+}. The kinetic parameters for *Mt*-DAH7PS are also similar to those observed for the type II DAH7PSs from *H. pylori* and *X. campestris*.[38,181]

The most striking finding from the present work is the clear evolutionary relationship between *Mt*-DAH7PS, the first type II DAH7PS to be structurally characterized, and the type Iα and type Iβ enzymes. Crystal structures have been reported for type Iα DAH7PS enzymes from *E. coli* and *S. cerevisiae*,[49,53] and for type Iβ DAH7PS enzymes from *T. maritima* and *P. furiosus*.[107] All of these type I enzymes are homotetramers, with $(\beta/\alpha)_8$ TIM-barrel subunits containing four apparently independent active sites. *Mt*-DAH7PS has the same monomer-fold, and although TIM-barrel enzymes are very common it is significant that the only close matches in the PDB are other DAH7PS enzymes and the related KDO8PSs. The key observation, however, is that when *Mt*-DAH7PS is superimposed on to the type I DAH7PSs on the basis of its core barrel structure, all of the active site residues match, both spatially and chemically. This congruence of both fold and catalytic apparatus makes it unequivocal that the type I and type II enzymes are related by divergent evolution from a common ancestor.

While the *Mt*-DAH7PS structure shows that the type I and type II enzymes share a common fold and active site architecture, there are significant differences both in the additional structural elements that adorn the core $(\beta/\alpha)_8$ barrel and in quaternary structure. As shown in figure 3.17, the type Iβ *Pf*-DAH7PS is the most stripped-down DAH7PS, its monomer comprising just the core $(\beta/\alpha)_8$ barrel. All of the structures have a structural element (α-helix in *Mt*-DAH7PS, *Ec*-DAH7PS and *Sc*-DAH7PS, and a β-ribbon in *Tm*-DAH7PS and *Pf*-DAH7PS) that closes the N-terminal end of the barrel. Most importantly, however, all except the *P. furiosus* enzyme possess extra structural motifs (domains or subdomains) that are involved in allosteric regulation (described in further detail in Chapter Four).

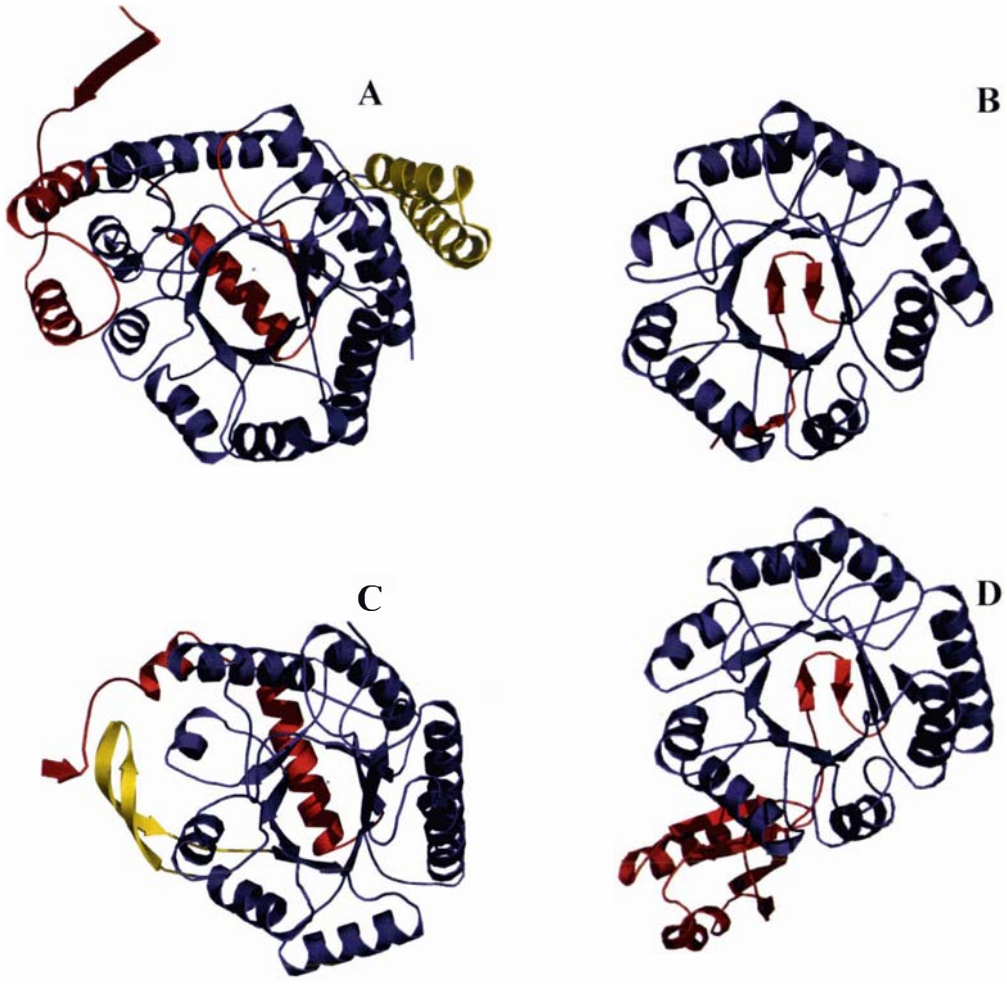


Figure 3.17 Comparison of subunit structure of DAH7PS enzymes. The core $(\beta/\alpha)_8$-barrel common to all DAH7PSs is shown in blue. N-terminal extensions or domains are shown in red and other protrusions from the barrel are in yellow. The structures shown are for A, *Mt*-DAH7PS, B, *Pf*-DAH7PS, C, *Ec*-DAH7PS, and D, *Tm*-DAH7PS.

A further difference between the various DAH7PS enzymes is in their quaternary structures (figure 3.18). The type Iα and type Iβ DAH7PSs are homotetrameric, but only one monomer-monomer interface is conserved between these two subfamilies. This common interface primarily involves helices α3, α4 and α5, and loops β2-α2, β3-α3, β4-α4 and β5-α5, and impacts on the positioning of the α2-β2 loop that provides part of the E4P binding site in the active site of the protein. The same interface is also observed in the closely related KDO8PS enzymes.[103,115] Despite this common dimer, the type Iα and Iβ DAH7PSs form tetramers using completely different structural

elements, meaning that while the dimers can be superimposed, the tetramers cannot. In contrast, the monomer-monomer interface of *Mt*-DAH7PS shows no commonality with any subunit interface found in any previously reported DAH7PS (or KDO8PS) structure.

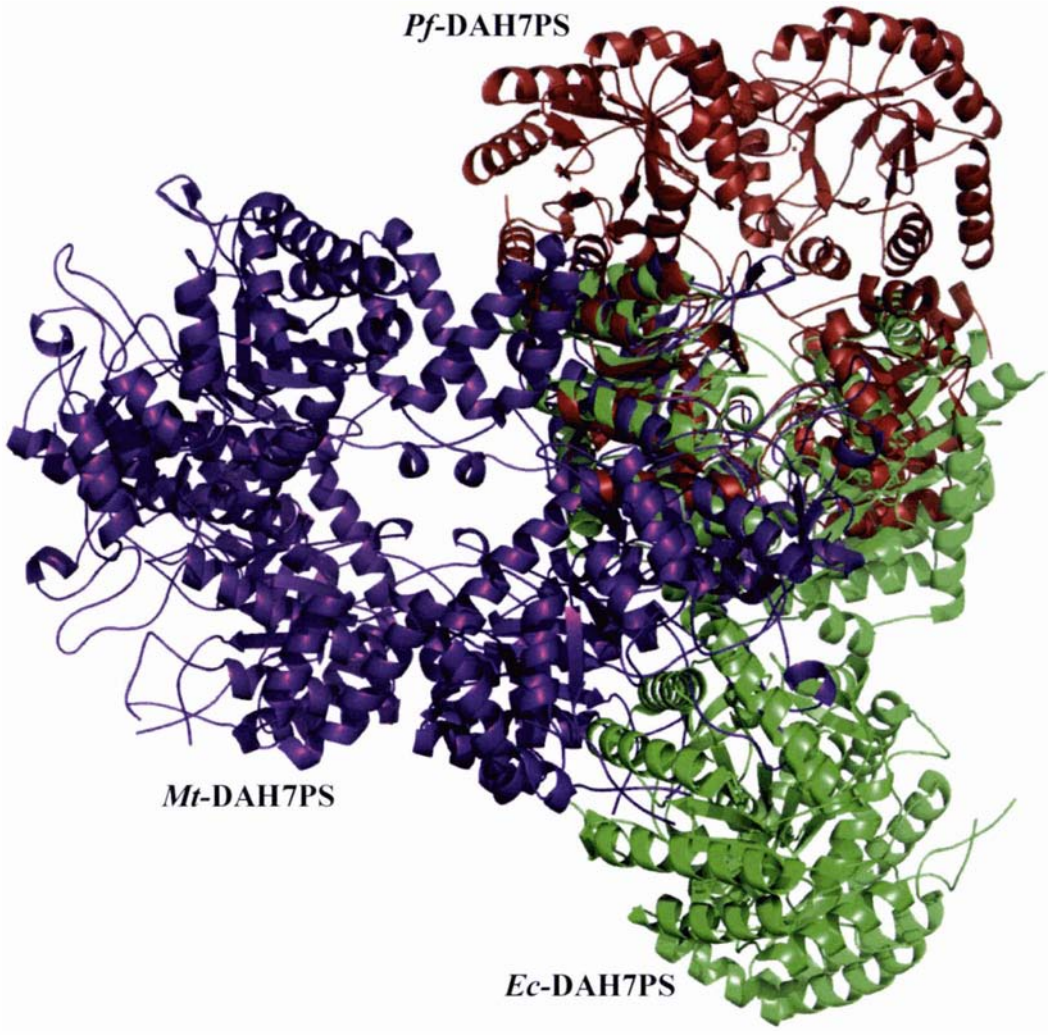


Figure 3.18 Comparison of quaternary structures of DAH7PS enzymes. Superposition of the type II *Mt*-DAH7PS tetramer (purple) onto the type Iα *Ec*-DAH7PS tetramer (green) and the type Iβ *Pf*-DAH7PS tetramer (fire brick).

3.20 Conclusions

These studies on the type II DAH7PS from *M. tuberculosis* reveal a remarkable similarity in fold and function between type I and II enzymes. Despite their minimal sequence identity, and the significant differences in quaternary structure and extensions to the core $(\beta/\alpha)_8$ barrel, the key residues that interact with PEP and the divalent metal ion are completely conserved and positioned almost identically in the two DAH7PS types. The common active site architecture and chemistry, housed within a shared protein fold, suggest very strongly that the type I and II enzymes have arisen from a common DAH7PS ancestor. The structural similarity in the active sites suggests that inhibitors designed for other DAH7PS enzymes could make useful lead compounds for the development of new anti-TB drugs. The presence of one or two key differences, notably the presence of Trp280 adjacent to the PEP site, further suggests that more specific inhibitors that are selective for type II DAH7PSs could be developed on the basis of the present structure.

Note:

The work in this chapter is in collaboration with Prof Ted Baker from the University of Auckland, and has been published in a modified form in the Journal of Molecular Biology. A copy of this paper is in the back of this thesis.

CHAPTER FOUR

FEEDBACK-REGULATION OF TYPE II DAH7PS FROM *M. TUBERCULOSIS* AND *H. PYLORI*.

4.1 Introduction

Feedback-inhibition is a mechanism by which biosynthetic pathways regulate the flux from starting materials to pathway end-products. That is, if there is too much end-product present it inhibits an earlier reaction in the pathway. The first irreversible reaction of a biosynthetic pathway, called the committed step, is usually an important regulatory target. The final product (Z) often inhibits the enzyme that catalyzes the committed step (A to B) (figure 4.1).

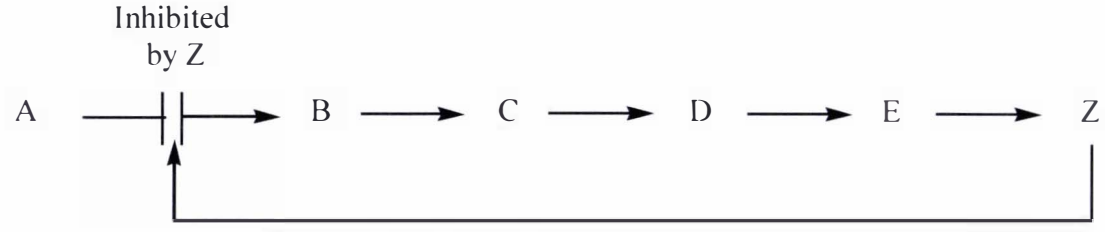


Figure 4.1 Feedback-inhibition of the first reaction of a biosynthetic pathway (A to B) by the end-product of the pathway (Z).

The biosynthesis of the aromatic amino acids is via the shikimate pathway, which branches at chorismate. Tyr and Phe are synthesized via prephenate and the biosynthesis of Trp is via anthranilate (figure 4.2). In fungi and bacteria, aromatic biosynthesis is primarily regulated by feedback-inhibition of DAH7PS,[46] which catalyzes the first committed step of the shikimate pathway. Given that there are three aromatic amino acids (end-products), the feedback mechanism for the regulation of DAH7PS is more complicated than what is described in figure 4.1 (refer to Chapter One, Section 1.3.1).

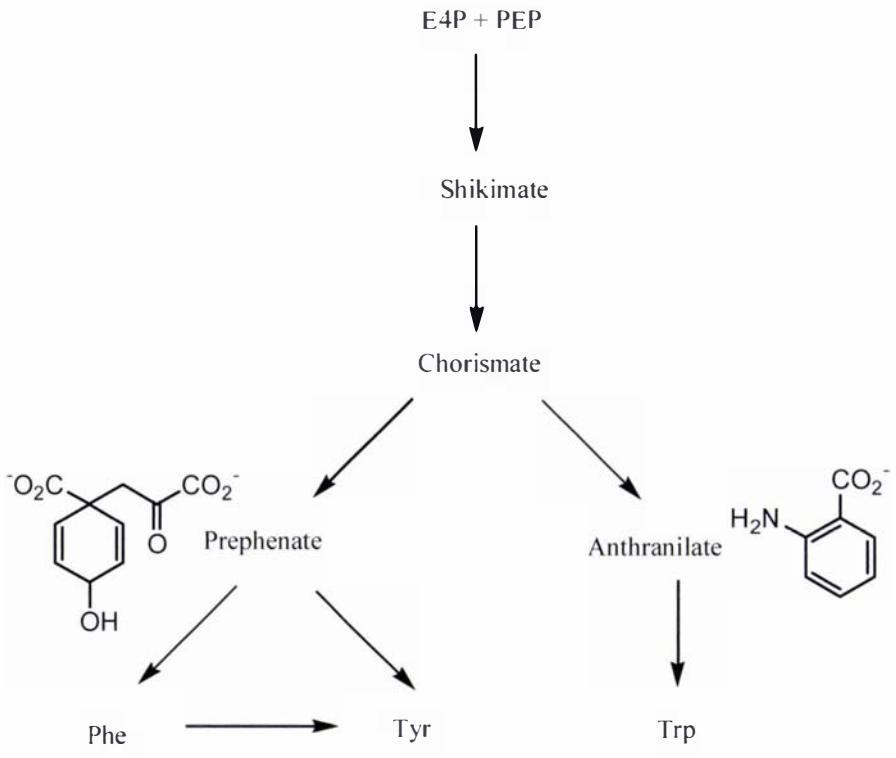


Figure 4.2 **Pathway for the biosynthesis of aromatic amino acids in microorganisms and plants**

In *E. coli* three DAH7PSs are responsible for catalyzing the first reaction of the shikimate pathway, each subject to feedback-inhibition by one of the aromatic amino acids. *S. cerevisiae* possesses two isoenzymes regulated by Phe and Tyr, respectively.[34,49] To completely prevent the conversion of PEP and E4P to DAH7P, all respective aromatic amino acids must be present at high levels. For example, complete inhibition of the shikimate pathway in *E. coli* requires the presence of all three aromatic amino acids, each directly inhibiting one of the DAH7PS isoenzymes (a schematic of this type of feedback-inhibition is in figure 4.3). In contrast to *E. coli* and *S. cerevisiae*, the genomes of *M. tuberculosis* and *H. pylori* contain a single gene encoding a type II DAH7PS. Therefore, as part of this project we decided to investigate the feedback regulation of both these type II enzymes by the three aromatic amino acids.

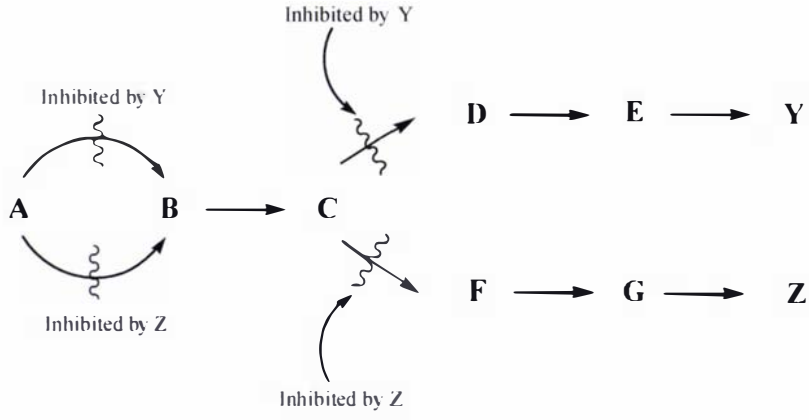


Figure 4.3 Feedback-inhibition mechanism involving multiple enzymes that catalyze the conversion of A to B. In the case of two different enzymes catalyzing the conversion of A to B, one is directly inhibited by Y and the other by Z.

The characterization of Iα enzymes from *E. coli* and *S. cerevisiae* and of Iβ enzymes from *P. furiosus* and *T. maritima* reveals a correlation between feedback regulation and the addition of structural elements decorating the basic $(\beta/\alpha)_8$ barrel. All structurally characterized DAH7PSs possess extra structural motifs that are involved in allosteric inhibition, except for the *P. furiosus* enzyme which by itself is not regulated allosterically as it lacks these extra structural motifs. The type Iβ *Tm*-DAH7PS, which is very similar in the core $(\beta/\alpha)_8$ barrel structure to *Pf*-DAH7PS, has a ferredoxin-like domain attached to the N-terminus of the barrel.[51] This domain has been implicated in Phe and Tyr inhibition.[95] The crystal structures of type Iα *Ec*-DAH7PS(Phe) and *Sc*-DAH7PS(Tyr), in complex with Phe and Tyr, respectively, have been solved. Both amino acids bind in very similar locations involving extensions of the N-terminus and an extended α5-β6 loop.[49,52-54]

The crystal structure of *Mt*-DAH7PS described in Chapter Three reveals a similar protein-fold to that observed for type I DAH7PSs. The *Mt*-DAH7PS monomer (462 residues) consists of a core $(\beta/\alpha)_8$ TIM barrel domain with two major additions decorating the barrel, an extended β-strand followed by three helices (α0a, α0b and α0c) at the N-terminus, and a pair of helices (α2a and α2b) that extend the α2-β3 connecting loop. It is tempting to speculate that these two additions to the barrel may provide two distinct inhibitory binding sites for aromatic amino acids. To investigate this further we

have solved the crystal structure of *Mt*-DAH7PS in the presence of Trp and Phe, and we have performed kinetic studies in the presence of the three aromatic amino acids and chorismate.

This chapter details kinetic studies performed with not only *Mt*-DAH7PS but also *Hp*-DAH7PS in the presence of the three aromatic amino acids and chorismate. The crystal structures of *Mt*-DAH7PS in complex with PEP, Mn^{2+} and either Trp and Phe alone or in combination are described. The diversity of feedback-regulation displayed across all characterized DAH7PS enzymes is also discussed.

4.2 Feedback-inhibition Studies with *Hp*-DAH7PS

The sensitivity of *Hp*-DAH7PS towards Trp, Tyr, Phe and chorismate (250 μM), either alone or in combination, was investigated using standard assay conditions, as described in Chapter Six. The three aromatic amino acids and chorismate had no detectable effect on the activity of *Hp*-DAH7PS suggesting that this enzyme by itself is unregulated by the four products of the shikimate pathway. This does not necessarily mean that this type II enzyme is unregulated; it could be that we are yet to identify the metabolite/s or protein partner(s) responsible for regulating *Hp*-DAH7PS's activity. Further studies involving the determination of the crystal structure of *Hp*-DAH7PS may allow us to identify potential metabolites that regulate the activity of the *H. pylori* enzyme. The testing of other metabolites of the aromatic amino acid biosynthetic pathway as potential feedback regulators of *Hp*-DAH7PS is required, for example prephenate and anthranilate, to rule out feedback-inhibition of this enzyme by intermediates of the pathway.

4.3 Feedback-inhibition Studies with *Mt*-DAH7PS

The three aromatic amino acids and chorismate were tested as potential feedback-regulators of *Mt*-DAH7PS. Similar to *Hp*-DAH7PS, this was achieved using standard assay conditions, described in Chapter Six. The three aromatic amino acids (200 μM) and chorismate (200 μM) were tested individually and in combination (table 4.1). Phe and to a lesser extent Tyr, appear to activate *Mt*-DAH7PS, whereas Trp and chorismate

show no significant effect on enzyme activity. An interesting observation is the inhibition displayed when Trp was added in combination with either Phe or Tyr (to a lesser extent). The inhibition was considerably greater than the additive effect of the aromatic amino acids alone, suggesting synergistic inhibition with this combination of aromatic amino acids. The combinations of Tyr and Phe, Trp and chorismate, and chorismate and Phe do not show any signs of synergistic inhibition. The synergistic inhibition displayed with Trp and Phe or Tyr provides functional evidence to support the hypothesis that the two additions to the core $(\beta/\alpha)_8$ barrel of *Mt*-DAH7PS (refer to Chapter Three, Section 3.17.1) are involved in the binding of Trp and Phe or Tyr. In order to investigate this further a crystal structure in the presence of Trp and Phe was determined (refer to Section 4.6.3). It should be noted that the level of inhibition of *Mt*-DAH7PS by Trp and Phe is very sensitive to the concentration of E4P. The inhibition of *Mt*-DAH7PS is significantly increased at lower concentrations of E4P (refer to Section 4.10.1 for further detail).

Inhibitor	**% activity**
None	100
Trp	95
Tyr	109
Phe	152
Chorismate	100
Trp + Phe	54
Trp + Phe*	69
Trp + Tyr	78
Trp + Chorismate	93
Tyr + Phe	153
Chorismate + Phe	143

Table 4.1 **Table showing effect of aromatic amino acids and chorismate on *Mt*-DAH7PS activity.** Standard assays were performed (refer to Chapter Six) which consisted of E4P (142 µM), PEP (200 µM), $MnSO_4$ (100 µM) and either aromatic amino acids (200 µM) or chorismate (200 µM). The combination of Trp and Phe annotated with an asterisk (*) is in the presence of 100 µM of each amino acid. All reactions were initiated by the addition of *Mt*-DAH7PS, which had been purified in the presence of Thesit (2 µL, 4.6 mg/mL). Assays were

performed in triplicate, and results were averaged with an estimated error no greater than 10 %.

The activation of *Mt*-DAH7PS by Phe (and Tyr to a lesser extent) is very interesting and not expected. At first we considered that it was possible Thesit, a non-ionic detergent that is present in the Se-Met structure and displaced by two Phe molecules in the Trp and Phe structure (refer to Section 4.9.1), may inhibit the enzyme. Therefore, the addition of either Phe or Tyr removes Thesit and eliminates the inhibition of *Mt*-DAH7PS by the detergent molecule. To investigate this *Mt*-DAH7PS was purified in the absence of Thesit and standard assays were performed with the addition of Thesit (0.005 % and 0.025 %). Results showed that the addition of Thesit to the reaction mixture has no detectable effect on the activity of *Mt*-DAH7PS. The kinetic parameters were determined for the Thesit-free *Mt*-DAH7PS protein (figure 4.4). The apparent K_M values for E4P and PEP were 12.0 ± 1.0 μM and 21 ± 1.3 μM, respectively and the k_{cat} value was calculated as 4.5 ± 0.1 s^{-1}. These values were found to be very similar to those reported for protein purified in the presence of detergent (refer to Chapter Three, Section 3.9), suggesting that Thesit does not inhibit *Mt*-DAH7PS.

The activation of DAH7PS by the aromatic amino acids has been reported for the plant type II enzymes (refer to Chapter One, Section 1.4). The feedback regulation of *Mt*-DAH7PS by the aromatic amino acids is more complex than that reported for any other DAH7PS. The difference in the levels of activation of *Mt*-DAH7PS by Phe and Tyr could be due to the latter amino acid not being essential; that is, Tyr can be synthesized via the hydroxylation of Phe.

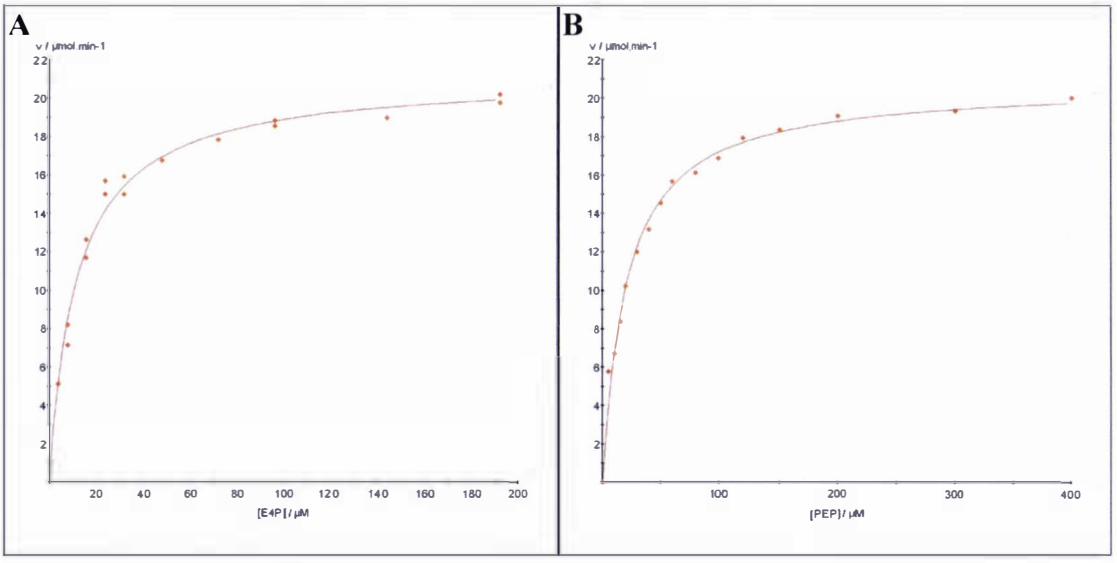


Figure 4.4 **Michaelis-Menten plots for determination of K_M values for E4P A, and PEP B.** The reaction mixtures for the determination of the K_M of E4P consisted of PEP (225 µM), $MnSO_4$ (100 µM) and E4P (8 to 192 µM), in BTP (50 mM) pH 7.5 buffer. The determination of the K_M of PEP consisted of reaction mixtures of E4P (200 µM), $MnSO_4$ (100 µM) and PEP (5 µM to 400 µM), in BTP (50 mM) pH 7.5 buffer. The reaction was initiated by the addition of purified *Mt*-DAH7PS (2.5 µL, 2 mg/mL) and carried out at 30 °C. K_M and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using Enzfitter (Biosoft).

To investigate the unique mode of feedback regulation of *Mt*-DAH7PS by all three aromatic amino acids further studies were performed. These included a series of assays with varying ratios of Trp and Phe (figure 4.5) or Trp and Tyr (figure 4.6) so that the total aromatic amino acid concentration remained constant at 200 µM. Results showed that the level of inhibition is sensitive to the ratio of concentrations of Trp and Phe. At either 200 µM of Phe or Trp no inhibition is observed but as soon as the two are added in combination (even 20 µM:180µM) inhibition of *Mt*-DAH7PS is displayed. It is also noticed that 180 µM Phe + 20 µM Trp shows a higher level of inhibition than 20 µM Phe + 180 µM Trp. The synergistic inhibition displayed in the presence of Trp and Tyr is not as significant as that observed with Trp and Phe. It is only when the ratio of [Trp]:[Tyr] is close to one, that is 100 µM Trp + 100 µM Tyr that *Mt*-DAH7PS activity is inhibited. The enzyme concentration in the assays is approximately 10 nM which is significantly lower than the concentration of aromatic amino acids added (20-200 µM). Further studies involving the determination of the binding constants of Trp and Phe are

required to help explain the results displayed here as it would appear that the binding of Trp to the enzyme is stronger than that of Phe. This could be achieved by labeling Trp and Phe and using NMR spectroscopy to look at the effect on the signal in the presence of *Mt*-DAH7PS. A series of experiments of different concentrations of aromatic amino acids may allow a binding contant to be determined. A second option would be to use fluorescence spectroscopy and match changes in fluorescence with the binding of Trp to *Mt*-DAH7PS.

Our results indicate that the presence of both Trp and Phe (or Tyr to a lesser extent) is necessary for the inhibition of *Mt*-DAH7PS. In the absence of either Trp or Phe the enzyme is available to supply more chorismate as a precursor of the deficient amino acid to restore or maintain the aromatic ratio at the optimum set point for *M. tuberculosis* (under particular environmental conditions and physiological requirements). As well as feedback-inhibition of DAH7PS, there is evidence for additional feedback-regulatory strategies for each of the amino acid biosynthetic pathways after the branch point at the common metabolite chorismate. Anthranilate synthase (Trp branch) has been shown to be feedback-regulated by Trp[182] and prephenate dehydratase (Phe and Tyr branch) is feedback-regulated by Phe and Tyr.[41] In the case of *Mt*-DAH7PS it is interesting that only Phe or Tyr activate the enzyme and not Trp.

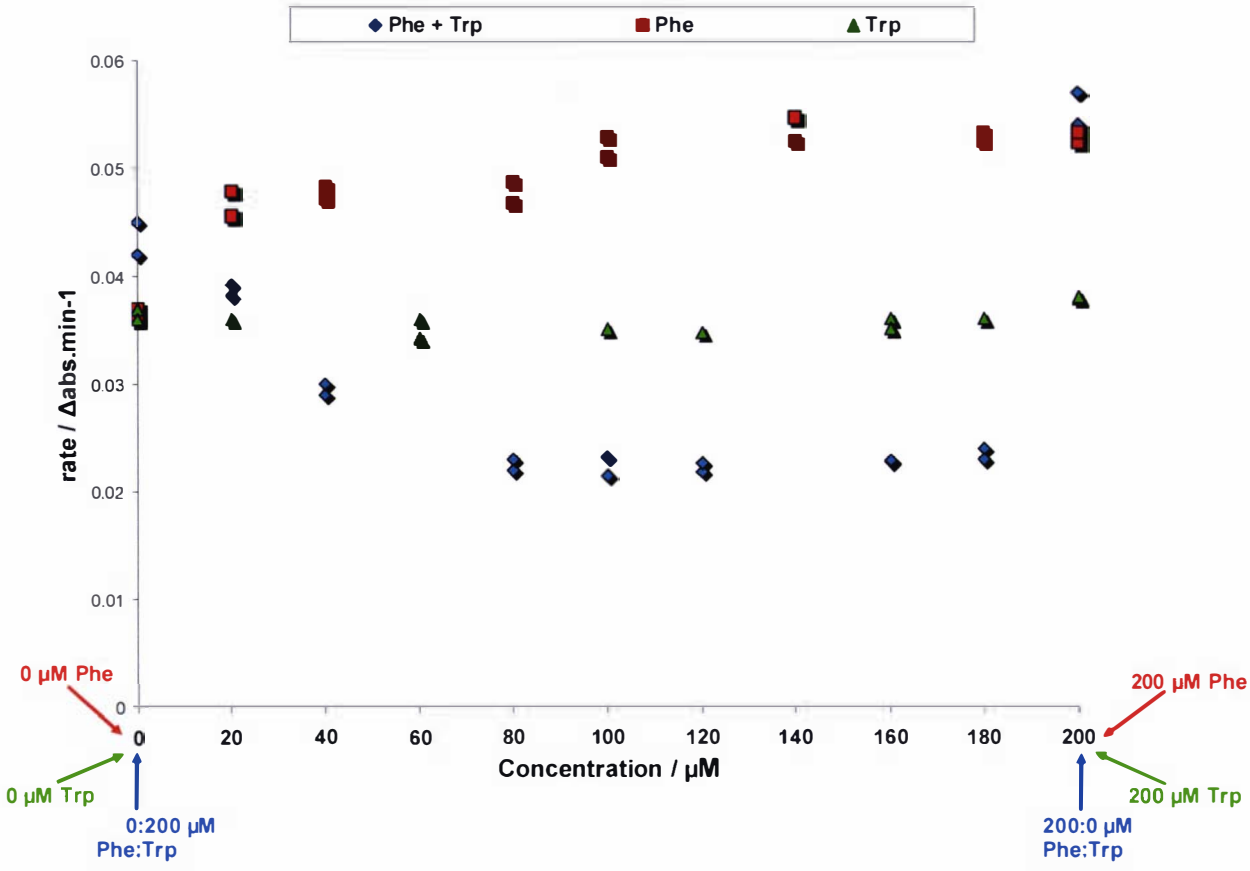


Figure 4.5 **Effect of different ratios of Trp and Phe on the activity of *Mt*-DAH7PS.** Standard assays were performed (refer to Chapter Six for details) consisting of E4P (140 μM), PEP (250 μM), $MnSO_4$ (100 μM) and varying concentrations of Trp and Phe (◆): 0 μM Trp + 200 μM Phe, 20 μM Trp + 180 μM Phe, 40 μM Trp + 160 μM Phe, 80 μM Trp + 120 μM Phe, 100 μM Trp + 100 μM Phe, 120 μM Trp + 80 μM Phe, 160 μM Trp + 40 μM Phe, 180 μM Trp + 20 μM Phe, and 200 μM Trp + 0 μM Phe. The reaction was initiated by the addition of *Mt*-DAH7PS (1 μL, 4.6 mg/mL). The dashed line shows enzyme activity in the absence of aromatic amino acids. As reference lines 0-200 μM Phe (■) and 0-200 μM Trp (▲) are also included on this plot.

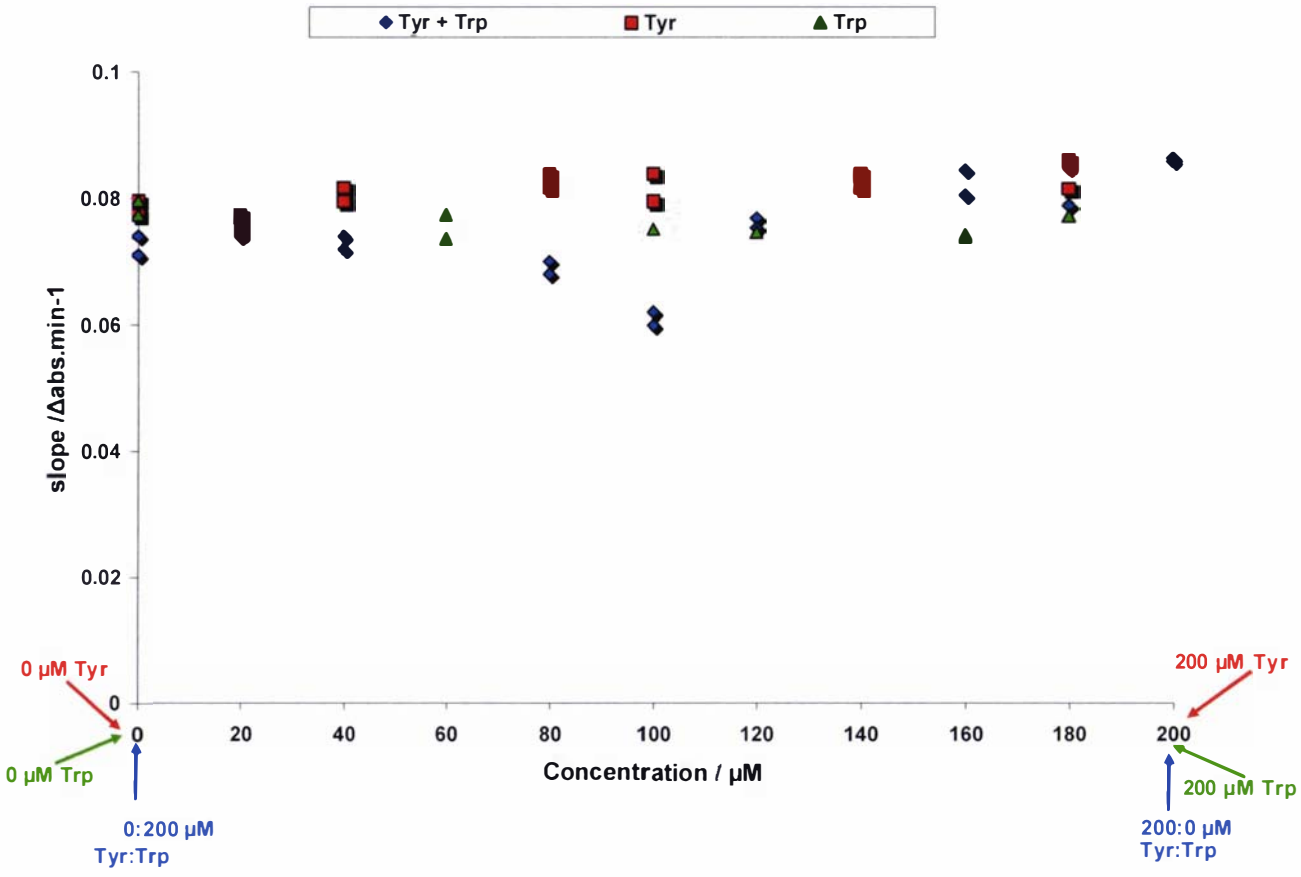


Figure 4.6 **Effect of different ratios of Trp and Tyr on the activity of *Mt*-DAH7PS.** Standard assays were performed (refer to Chapter Six for details) consisting of E4P (142 μM), PEP (200 μM), $MnSO_4$ (100 μM) and varying concentrations of Trp and Tyr (◆): 0 μM Trp + 200 μM Tyr, 20 μM Trp + 180 μM Tyr, 40 μM Trp + 160 μM Tyr, 80 μM Trp + 120 μM Tyr, 100 μM Trp + 100 μM Tyr, 120 μM Trp + 80 μM Tyr, 160 μM Trp + 40 μM Tyr, 180 μM Trp + 20 μM Tyr, and 200 μM Trp + 0 μM Tyr. The reaction was initiated by the addition of *Mt*-DAH7PS (2 μL, 4.6 mg/mL). The dashed line shows enzyme activity in the absence of aromatic amino acids. As reference lines 0-200 μM Tyr (■) and 0-200 μM Trp (▲) are also included on this plot.

The regulation of aromatic amino acid biosynthesis in *M. tuberculosis* is completely different to that observed for *E. coli* and *S. cerevisiae*, which possess multiple isoenzymes of DAH7PS that are regulated by one of the aromatic amino acids (refer to introduction of this chapter). In the case of *M. tuberculosis*, which possesses a single

DAH7PS, the presence of both end-products of the pathway, Trp (e.g Y) and Phe or Tyr (e.g Z) are necessary to significantly inhibit the type II DAH7PS (figure 4.7).

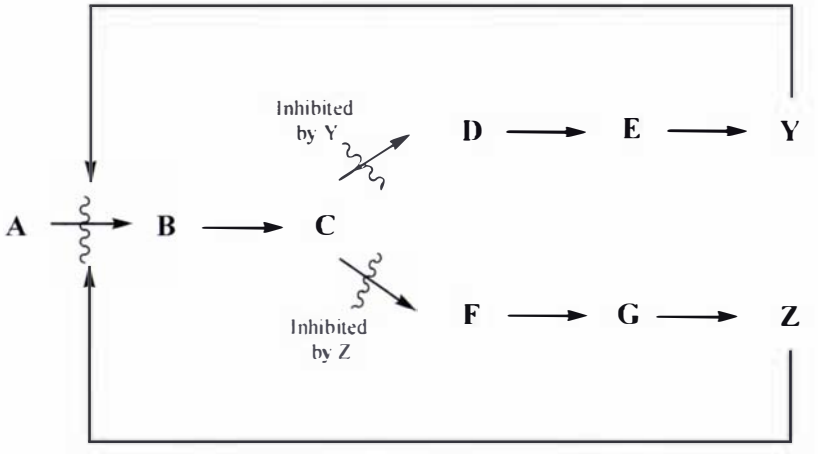


Figure 4.7 **Concerted feedback-inhibition mechanism.** Inhibition of the first step of the biosynthetic pathway requires both Y and Z to be present. The conversion of A to B is not inhibited by a single product.

4.4 Comparison with Other Type II Enzymes

Hp-DAH7PS activity was unaffected by the addition of aromatic amino acids or chorismate, and at the other extreme the *M. tuberculosis* enzyme was found to be sensitive to all three of the aromatic amino acids. Other microbial type II enzymes have been reported to be at least partially inhibited by Trp (table 4.2). The *X. campestris* enzyme is also inhibited by chorismate. Plant type II DAH7PSs appear to be insensitive to feedback-inhibition by aromatic amino acids, and in some cases aromatic amino acids have even been reported to enhance enzyme activity. For example, Trp was found to activate DAH7PS from carrot and potato.[34,68] DAH7PS from *S. antibioticus* is reported to be significantly inhibited by Trp and, interestingly, chorismate has been reported to exhibit a synergistic effect at all concentrations of Trp assayed.[147]

Organism	Feedback-inhibition	Reference
H. pylori	no	Section 4.2
M. tuberculosis	Trp/Phe or Tyr	Section 4.3
X. campestris	Trp/ chorismate	38
S. coelicolor	Trp	37
S. caespitosus†	Trp	118,162
S. rimosus†	Trp	117
S. aurefaciens†	Trp	121
N. crassa	Trp	148
S. antibioticus	Trp and synergistic inhibition with Trp and chorismate	147

Table 4.2 **Comparison of properties of microbial type II DAH7PSs**

†No primary sequence information is available for these enzymes; however, the DAH7PSs from these species are highly likely to be type II enzymes as they all have subunit molecular masses around 50 kDa

4.5 Crystallization and Data Collection of *Mt*-DAH7PS with Trp and Phe

To investigate the synergistic inhibition displayed with Trp and Phe or Tyr (refer to Section 4.3) crystal structures in the presence of Trp and Phe alone and in combination were sought after. Originally co-crystallization of *Mt*-DAH7PS and the aromatic amino acids using the original crystallization conditions for native *Mt*-DAH7PS were tried; however, all attempts were unsuccessful in producing crystals. The next step was to try soaking native *Mt*-DAH7PS crystals in mother liquor supplemented with either combination of amino acids. Native *Mt*-DAH7PS protein was crystallized using conditions described in Chapter Three. Several datasets were collected but only the highest resolution datasets for crystals soaked in Trp and Phe alone and in combination will be discussed in this thesis.

4.5.1 *Mt*-DAH7PS and Trp

A native crystal was soaked in Tris.HCl (pH 7.0, 0.1 M), ammonium sulfate (1.5 M), glycerol (15 % (v/v)) and Trp (2mM). Crystals were left overnight at 10 °C then flash-frozen in liquid nitrogen prior to transport to the Stanford synchrotron. Cryoprotectant consisted of mother liquor with an increased concentration of glycerol (25 % (v/v)). This crystal proved to be trigonal, with space group $P3_221$, unit cell dimensions $a=b=204.71$ Å, $c=66.26$ Å, $\alpha=\beta=90°$, $\gamma=120°$, and two molecules in the asymmetric unit, with a solvent content of 68 % (v/v). Data were collected at a wavelength of 0.97929 Å at the Stanford Synchrotron Radiation Laboratory. The data were processed using DENZO and SCALEPACK by scientists at the synchrotron. Full details of the data collection statistics are in table 4.3.

Table 4.3 Data collection statistics for *Mt*-DAH7PS-Trp

Values in parentheses are for the outermost shell	
Space group	$P3_221$
Unit cell dimensions (Å)	
$a=b$	204.71
c	66.26
Resolution range (outer shell) (Å)	50.0-2.5 (2.59-2.5)
Wavelength (Å)	0.97929
No. measured reflections	107433
No. unique reflections	52344
Completeness (%)	100 (100)
Redundancy	2.8 (2.8)
$\langle I/\sigma \rangle$	10.4 (2.1)
R_{merge} (%)	13.4 (63.3)

4.5.2 *Mt*-DAH7PS and Phe

A native crystal was soaked in Tris.HCl (pH 7.0, 0.1 M), ammonium sulfate (1.5 M), glycerol (15 % (v/v)) and Phe (2 mM). Crystals were left soaking for seven days at 10 °C. For data collection the crystals were soaked in a cryoprotectant comprised of mother liquor plus an increased concentration of glycerol (25 % (v/v)). The crystal was then flash-frozen in a cold-stream of gaseous nitrogen at 110 K. This crystal proved to be trigonal, with space group $P3_221$, unit cell dimensions $a=b=204.65$ Å, $c=66.61$ Å, $\alpha=\beta=90°$, $\gamma=120°$, and two molecules in the asymmetric unit, with a solvent content of 68 % (v/v). Data were collected at a wavelength of 1.542 Å (rotating copper anode)

using a Rigaku MicroMax007 generator with Osmic Blue optics and an RAxis IV++ detector. Full details of the data collection statistics are in table 4.4. A moderately redundant dataset was collected to a maximum resolution of 3.0 Å and an overall R_{merge} of 16.4 %, using non-overlapping 0.25° oscillations collected for fifteen minutes per frame at a crystal-detector distance of 180 mm. Data were processed using d*TREK as part of CrystalClear version 1.3.6 (Rigaku).

Table 4.4 Data collection statistics for *Mt*-DAH7PS-Phe

Values in parentheses are for the outermost shell	
Space group	$P3_221$
Unit cell dimensions (Å)	
$a=b$	204.65
c	66.61
Resolution range (outer shell) (Å)	39.6-3.0 (3.11-3.0)
Wavelength (Å)	1.542
No. measured reflections	82552
No. unique reflections	31363
Completeness (%)	97.2 (97.2)
Redundancy	2.63 (2.68)
$\langle I/\sigma \rangle$	5.6 (3.0)
R_{merge}	16.4 (34.4)

4.5.3 *Mt*-DAH7PS and Trp and Phe

A native crystal was soaked in a cryoprotectant solution containing Tris.HCl (0.1 mM, pH 7.5), ammonium sulfate (1.4 mM), Trp (2.2 mM), Phe (2.2 mM) and glycerol (25 % (v/v)), for approximately 30 minutes. The crystal was then flash-frozen in a stream of gaseous nitrogen at 110 K. This crystal proved to be trigonal, with space group $P3_221$, unit cell dimensions $a=b=204.54$ Å, $c=66.46$ Å, $\alpha=\beta=90°$, $\gamma=120°$, and two molecules in the asymmetric unit, with a solvent content of 68 % (v/v). Data was collected at a wavelength of 1.542 Å (rotating copper anode) using a Rigaku MicroMax007 generator with Osmic Blue optics and an RAxis IV++ detector. Full details of the data collection statistics are in table 4.5. A moderately redundant dataset was collected to a maximum resolution of 2.0 Å and an overall R_{merge} of 5.7 %, using non-overlapping 0.25° oscillations collected for fifteen minutes per frame at a crystal-detector distance of 180 mm. Data were processed using d*TREK as part of CrystalClear 1.3.6 (Rigaku).

Table 4.5 Data collection statistics for *Mt*-DAH7PS-Trp-Phe

Values in parentheses are for the outermost shell	
Space group	$P3_221$
Unit cell dimensions (Å)	
$a=b$	204.61
c	66.41
Resolution range (outer shell) (Å)	39.5-2.0 (2.07-2.0)
Wavelength (Å)	1.542
No. measured reflections	204630
No. unique reflections	98952
Completeness (%)	92.2 (71.9)
Redundancy	2.07 (1.73)
$\langle I/\sigma \rangle$	8.3 (1.9)
R_{merge} (%)	5.7 (33.4)

4.6 Crystal Structure Determination and Refinement

The soaking of the native crystals did not alter the crystal packing in any way, in that the space group and cell dimension parameters are essentially identical to the Se-Met *Mt*-DAH7PS dataset. As all structures were isomorphous we were able to use phase information from the Se-Met structure to determine the crystal structures of *Mt*-DAH7PS in complex with either aromatic amino acid and in combination. Measured intensities were converted to amplitudes using TRUNCATE as part of Scalepack2mtz or dtrek2mtz (CCP4[183]). The structure was then solved by performing two rounds of rigid-body refinement (CCP4) calculating phase information from the Se-Met coordinate (PDB) file. Optimization of the model consisted of repetitive cycles of rebuilding using COOT and refinement with REFMAC 5 (CCP4). Water molecules were added automatically in COOT and verified using 2Fo-Fc and Fo-Fc maps and by their potential to hydrogen bond to at least one protein atom or water molecule.

4.6.1 *Mt*-DAH7PS and Trp

The initial refinement statistics had values for R of ~24 %, R_{free} of 28 % and FOM of ~0.8. In the final model for *Mt*-DAH7PS in complex with Trp, molecule A comprises the complete polypeptide chain (residues 1-462) except for residues 266-268, 374-377, 420-421 and 426-431, and molecule B is similarly complete with the exception of

residues 372-379, 419-431 and 435-437. The two residues (Gly-Ala) from the linker to the cleaved His_6-tag are also modeled for subunit B and the alanine is modeled for subunit A. The main-chain torsion angles correspond well with the allowed values, 88.8 % in most favored region of Ramachandran plot, as defined by PROCHECK. In the current structure $Arg25_A$, $Ala240_A$, $Ala424_A$, $Asp2_B$ and $Leu12_B$ are in the disallowed region of the Ramachandran plot. In the current structure two manganese ions, three glycerol molecules, four sulfate ions and two Trp molecules (all identified by their shapes and environments) were modeled into extra density. Subsequent refinement indicated that the occupancy of Mn^{2+} was ~30 % and that of the SO_4^{2-} ions, glycerol and Trp molecules was ~100 %. Two phosphate ions were modeled into both subunits (100 % occupancy) where PEP is found in the Se-Met structure. The 2Fo-Fc map does not show electron density for the remainder of the PEP molecule in the current Trp structure. Density for a molecule of detergent Thesit was also found bound at the interface between subunit A and B with occupancy of 100 %. It was modeled in as a dodecyl ether moiety with one polyethyleneglycol unit attached. Final refinement statistics are given in table 4.6. The crystal structure reveals a disulfide linkage between the thiol groups of the metal-binding cysteine, Cys87, and Cys440 and this is modeled in at ~70 % occupancy. The superposition of the two subunits in the asymmetric unit of the Trp and Se-Met structures yields a root-mean-square deviation (rmsd) of 0.72 Å for the main-chain atoms.

4.6.2 *Mt*-DAH7PS and Phe

The initial refinement statistics had values for R of ~23 %, R_{free} of ~29 % and FOM of ~0.82. In the current model for *Mt*-DAH7PS in complex with Phe, molecule A comprises the complete polypeptide chain (residues 1-462) except for residues 419-431, and molecule B is similarly complete with the exception of residues 376-378 and 422-431. The two residues (Gly-Ala) from the linker to the cleaved His_6-tag are modeled for subunit B. The main-chain torsion angles correspond well with the allowed values, 84.2 % in most favored region of the Ramachandran plot, as defined by PROCHECK. In the current structure $Ser189_A$, $Ser375_A$, $Ser14_B$ and $Ala234_B$ are in the disallowed region of the Ramachandran plot. In the current structure two sulfate ions, two phosphate ions and four Phe molecules were modeled into both subunits of the asymmetric unit. The two phosphate and sulfate ions, and the four Phe molecules were modeled into both

subunits at 100 % occupancy. The phosphate ions are modeled in at the same place that the phosphate group of PEP is found in the Se-Met structure. The 2Fo-Fc map does not show electron density for the remainder of the PEP molecule in the current Phe structure and no electron density was observed for any manganese ions. The final refinement statistics for the Phe structure are given in table 4.6. The superposition of the two subunits in the asymmetric unit of the Phe and Se-Met structures yields an rmsd of 0.74 Å for the main-chain atoms.

4.6.3 *Mt*-DAH7PS and Trp and Phe

The following sections in this chapter refer to the Trp and Phe binding sites of this structure. This is a higher resolution structure and we are interested in the structural changes that occur to *Mt*-DAH7PS upon binding of both aromatic amino acids. The complete solution had an initial R and R_{free} of ~35 %, and FOM of 0.67. In the current model for *Mt*-DAH7PS in complex with Trp and Phe, molecule A comprises the complete polypeptide chain (residues 1-462) except for residues 11-15, and molecule B is similarly complete with the exception of residues 10-14. In each molecule, a manganese ion, a PEP molecule and a Trp and Phe molecule (all identified by their shapes and environments) were added into extra density, in equivalent positions in the two molecules. Subsequent refinement indicated partial occupancies for each of these species, estimated 75 % for Mn^{2+} and 50 % for PEP. A sulfate ion was modeled into subunit B with an occupancy of 50 %. There are also seven glycerol molecules modeled in to the Trp and Phe structure. The main-chain torsion angles correspond well with the allowed values, 91.4 % in most favored region (and no residues in the disallowed region) of the Ramachandran plot, as defined by PROCHECK. Final refinement statistics are given in table 4.6.

Refinement statistics for *Mt*-DAH7PS-Trp	
Resolution range (Å)	50.0-2.5 (2.58-2.5)
Number of reflections (test set)	52344 (2815)
R factor	0.195
R_{free}	0.240
Number of non-hydrogen atoms	
Protein (two molecules)	6818
Water	351
Mn^{2+}	2
PO_4^{3-}	10
SO_4^{2-}	20
Glycerol	18
Thesit	16
Trp	30
Mean *B* value (Å^2)	28.8
rms deviations from ideality	
Bond lengths (Å)	0.02
Bond angles (deg.)	2.0
Residues in the most favored region of Ramachandran plot (%)	88.8

Refinement statistics for *Mt*-DAH7PS-Phe	
Resolution range (Å)	39.6 (3.11-3.0)
Number of reflections (test set)	29549 (1585)
R factor	0.235
R_{free}	0.288
Number of non-hydrogen atoms	
Protein (two molecules)	6900
Water	41
PO_4^{3-}	10
SO_4^{2-}	10
Phe	48
Mean *B* value (Å^2)	25.4
rms deviations from ideality	
Bond lengths (Å)	0.016
Bond angles (deg.)	1.7
Residues in the most favored region of Ramachandran plot (%)	84.2

Refinement statistics for *Mt*-DAH7PS-Trp-Phe	
Resolution range (Å)	39.5-2.0 (2.07-2.0)
Number of reflections (test set)	99865 (5245)
R factor	0.213
R_{free}	0.247
Number of non-hydrogen atoms	
Protein (two molecules)	7006
Water	577
Mn^{2+}	2
PEP	20
SO_4^{2-}	5
Glycerol	42
Trp	30
Phe	24
Mean *B* value (Å^2)	29.0
rms deviations from ideality	
Bond lengths (Å)	0.014
Bond angles (deg.)	1.5
Residues in the most favored region of Ramachandran plot (%)	91.4

Table 4.6 Refinement statistics for *Mt*-DAH7PS in complex with Trp, Phe, and Trp and Phe

The binding of Trp and Phe leaves the *Mt*-DAH7PS subunit largely unaffected with the largest structural changes observed involving residues surrounding the Phe and Trp binding sites. The two molecules in the asymmetric unit of the Trp and Phe and Se-Met structures superimpose with a rmsd of 0.97 Å for all main-chain atoms. When residues involved in Phe and Trp binding are excluded in the superposition (residues 1-18 and 188-241), the main-chains are very similar, with an rmsd of 0.45 Å, the excluded segments have an rmsd of 2.17 Å (refer to Section 4.7 and 4.8 for further information on residues involved in Trp and Phe binding).

4.7 The Trp Binding Site

The Trp binding site in both the Trp only and the Trp and Phe structure is essentially identical. One molecule of Trp is bound in a cavity between the extra-barrel pair of helices (α2a and α2b) and the core barrel (α1) in each monomer, and is located ~25 Å away from Mn^{2+} at the active site (figure 4.8). The Trp binding site is predominately formed by residues of the α2b and α1 helices, and the α2b-β3 loop, which are also involved in tetramer association. The carboxylate group of Trp is coordinated by N^{ζ} of Lys123 (2.9 Å and 3.1 Å) and two water molecules (2.6 Å and 2.7 Å). The amino group of Trp interacts with the side chain of Asn237 (O^{δ}, 2.7 Å), the carbonyl oxygen of Thr240 (2.9 Å) and a water molecule (3.0 Å). The hydrophobic ring is surrounded by hydrophobic residues, Val111 C^{γ} (4.3 Å), Leu107 C^{η} (4.3 Å), Leu194 C^{γ} (4.3 Å) and Val197 $C^{\gamma 2}$ (4.5 Å). Trp is positioned in the hydrophobic pocket so that its N^{ε} hydrogen bonds to the main-chain carbonyl oxygen of Ala192 (2.7 Å) (figure 4.9). Upon inhibitor binding the hydrophobic sidechains of Leu194, Val111 and Leu107 move away slightly (Leu194 CG 0.9 Å, Val111 Cβ 0.6 Å and Leu107 0.7 Å). Movements are in relation to the positions these residues occupy in the Se-Met structure. The largest movement is seen by residue Gln239, with its Cα atom moving more than 6 Å resulting in the enlargement of the binding cavity (figure 4.10), allowing Trp to bind and interact with Asn237, Thr240 and Lys123, which have moved up to 6 Å closer (Thr240 CO 3.5 Å, Asn237 CO 6.0 Å, Lys123 N^{ζ} 1.2 Å) as a result of inhibitor binding. The distances described here are for subunit B; similar contacts are observed in subunit A.

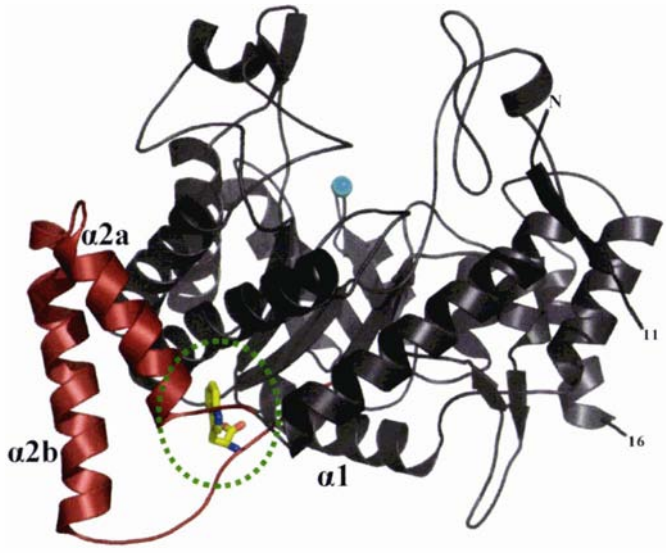


Figure 4.8 **Trp binding site in the Trp and Phe structure.** The site is located between the extra-barrel pair of helices (α2a and α2b) (firebrick) and the core barrel (α1) (grey). Subunit B is shown in this figure. Trp is shown in yellow inside the green circle and Mn^{2+} is shown as a sphere (cyan).

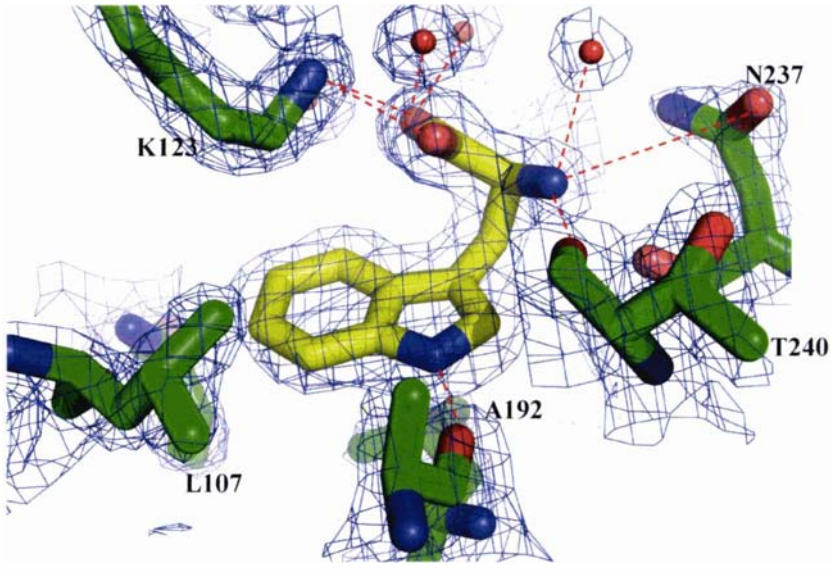


Figure 4.9 **Residues involved in the binding of Trp in the Trp and Phe structure.** 2Fo-Fc map in blue at contour level of 1.5 σ. Interactions between Trp and surrounding water molecules and protein residues in subunit B of *Mt*-DAH7PS are represented as red dashed lines.

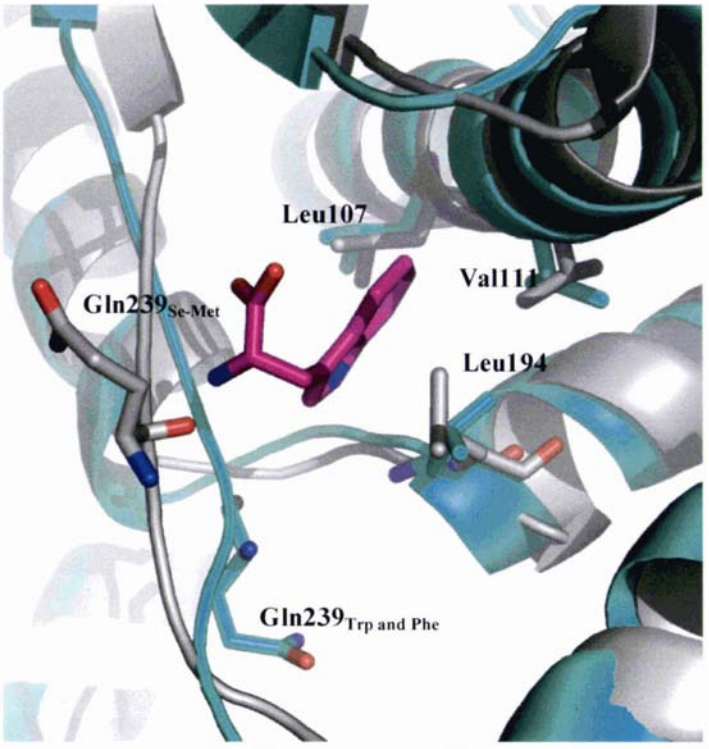


Figure 4.10 **Superposition of the Trp binding site in subunit B of the Trp and Phe structure (cyan) and the equivalent residues in subunit B of the Se-Met structure (grey).**

The residues involved in making up the Trp binding pocket are all within one subunit of the protein; however, several of these residues also contribute to the tetramer interface of *Mt*-DAH7PS (Gln239, Leu194, Val111, Asn237). The formation of the tetramer, using crystallographic symmetry operations (COOT) (figure 4.11), results in the burying of Trp away from accessible solvent. The binding of Trp appears to alter the interactions between the two subunits that make up the tetramer interface (refer to Section 4.9.2 for further detail). We propose that the tetramer is essential for either the binding of Trp and/or the allosteric response back to the active site of the protein. To investigate this further, site-directed mutagenesis studies, which disrupt the tetrameric association of *Mt*-DAH7PS, are required (this is discussed in further detail in Section 4.13).

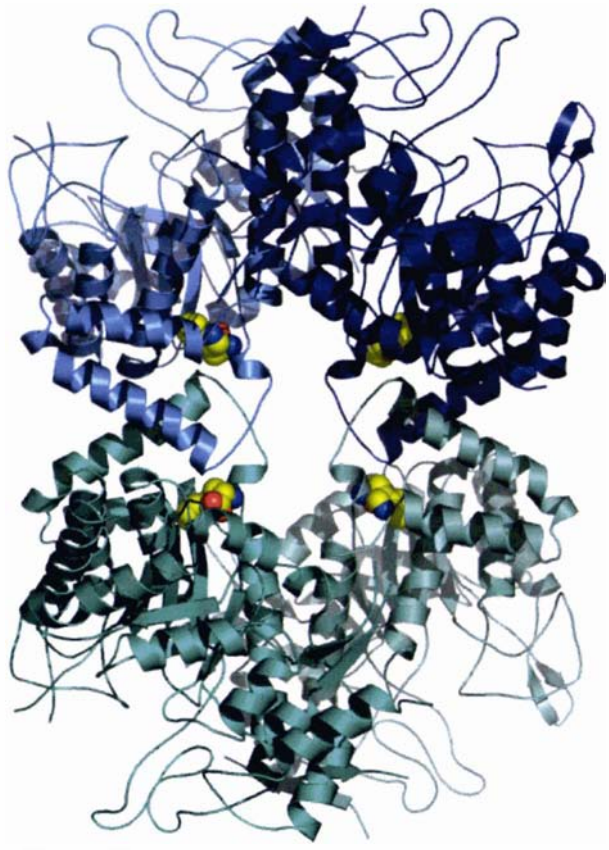


Figure 4.11 **Trp binding in *Mt*-DAH7PS tetramer.** The four molecules of Trp are shown as yellow spheres.

4.8 The Phe Binding Site

The Phe binding site in the Phe and Trp and Phe structures is essentially identical except that in the latter structure residues 11 to 15 of both subunits are disordered. Phe is located near the interface of the tight dimer, and binding is through interactions with residues from both monomers (figure 4.12 and figure 4.13). The site is ~25 Å away from Mn^{2+} at the active site, and ~27 Å away from where Trp binds (in the Trp and Phe structure). The Phe binding site consists of residues of the β1-α1 loop (Phe91), α0b (Val51), α0b-α0c loop (Pro56) and predominately the α2 helix (Arg171$_B$, Asn175 $_B$, Val170, Y173, Ala174 $_B$, Ala178 $_B$, where $_B$ indicates the other subunit of the tight dimer). The carboxylate group of Phe is coordinated by $N^{\pi 2}$ and N^{ε} of Arg171 (2.8 Å and 2.9 Å, respectively), N^{δ} of Asn175 (2.9 Å) and two water molecules (both 2.8 Å). The amino group of Phe interacts with the side chain of Asn175 (O^{δ}, 2.8 Å), the main-chain carbonyl of Phe91 (2.8 Å) and two water molecules (2.9 Å and 3.3 Å). The hydrophobic ring is buried in a hydrophobic pocket consisting of residues from both

subunits. These residues include: $C^{\varepsilon 2}$ of Tyr173 (3.7 Å), C^{δ} of Pro56 (4.5 Å), C^{γ} of Val170 (4.6 Å), C^{γ} of Val51 (4.3 Å), C^{β} of $Ala178_B$ (3.7 Å), C^{β} of $Ala174_B$ (3.8 Å) and $C^{\zeta 2}$ of $Trp3_B$ (4.38 Å). The Phe binding pocket described here is what is observed in subunit A; similar interactions are seen in subunit B.

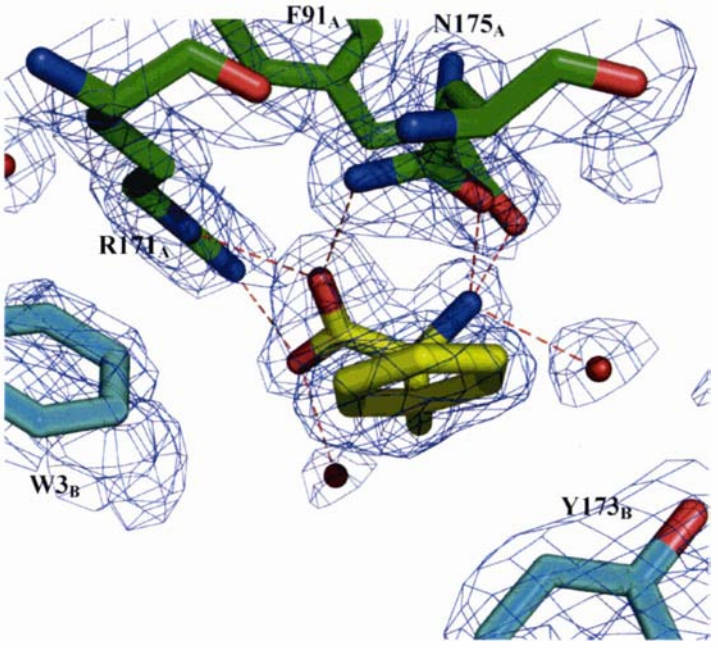


Figure 4.12 **Residues involved in the binding of Phe in the Trp and Phe structure.** 2Fo-Fc map in blue at contour level of 1.2 σ. Interactions between Phe and surrounding water molecules and protein residues are represented as red dashed lines. Residues of different subunits are shown by either an A or B next to the residue name.

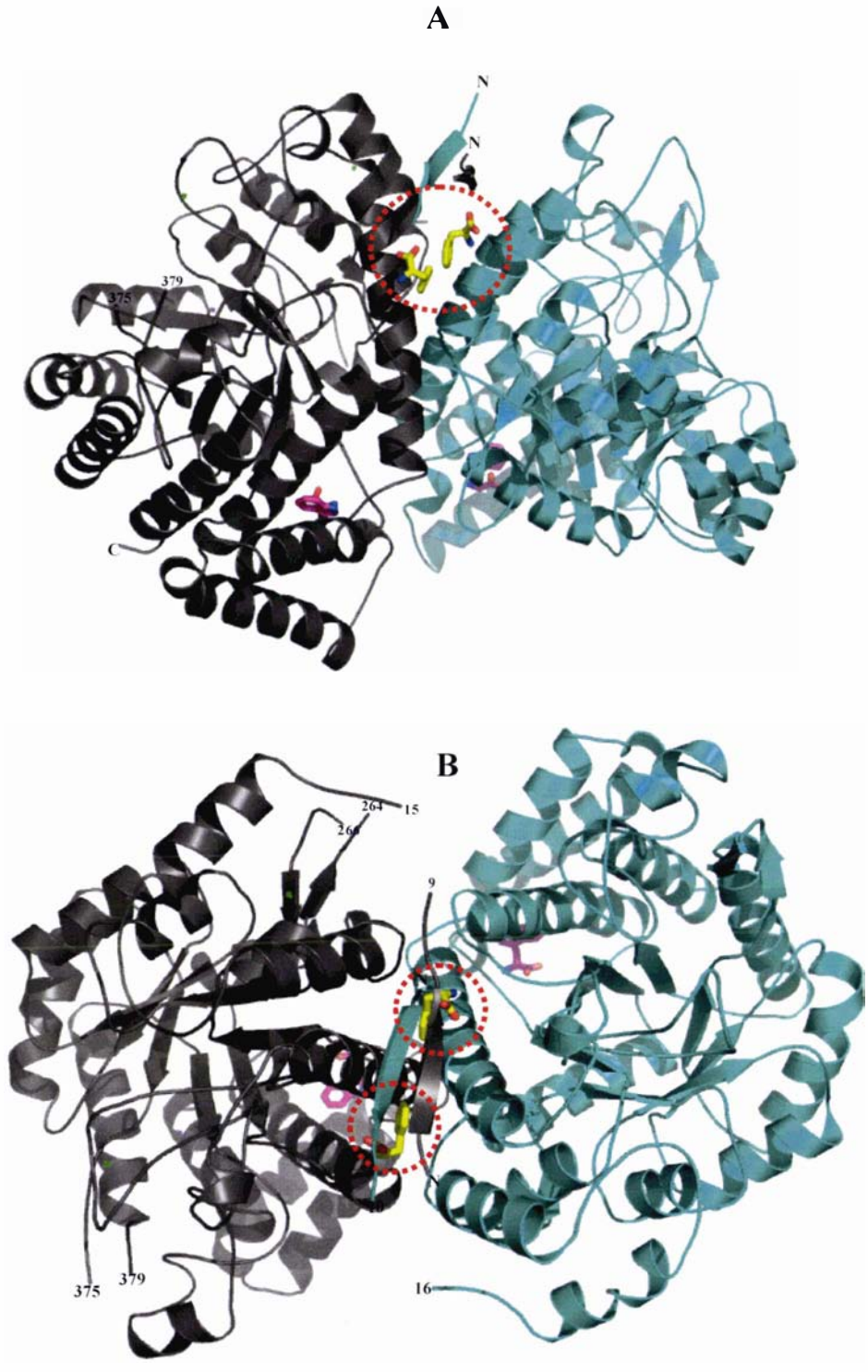


Figure 4.13 **The binding of Phe (yellow) at the dimer interface of *Mt*-DAH7PS in complex with Trp and Phe.** **A**, Side-on view and **B**, looking down the two-fold axis of symmetry. Molecule A is in cyan and molecule B in grey. Trp is shown in pink.

Upon the binding of Phe there is some movement in the residues at the N-terminus (1-13 in subunit B) (refer to the next section in this chapter for more detail), as well as the disruption of the salt-bridge between $Arg171_B$ and $Asp10_A$. There is a movement of the sidechain of $Arg171_B$ so that its N^{ε} moves 2.7 Å to interact with the carboxylate group of the Phe molecule. This is observed in both the Phe and Trp and Phe structures. In the Phe structure a second Phe molecule is built into the 'Trp' binding site, with an occupancy of 100 %, in both subunits of the asymmetric unit. The binding of Phe results in a similar movement of the residues of the α2b-β3 loop (residues 233-241) to what is observed when Trp binds in the Trp and Phe structure (figure 4.14) (described in previous section). However, in the current Phe crystal structure the position of the amino and carboxylate groups of the two Phe molecules in the 'Trp' binding is not clearly defined, making it difficult to assign any interaction between the Phe molecules and surrounding protein residues. The only significant difference between the Phe and Trp and Phe structures is that in the Phe crystal structure Arg236 in both subunits of the asymmetric unit is disordered and is modeled in as an alanine. In the Trp and Phe structure this residue interacts with Asn237 from the other subunit making up the tight dimer (refer to Section 4.9.2).

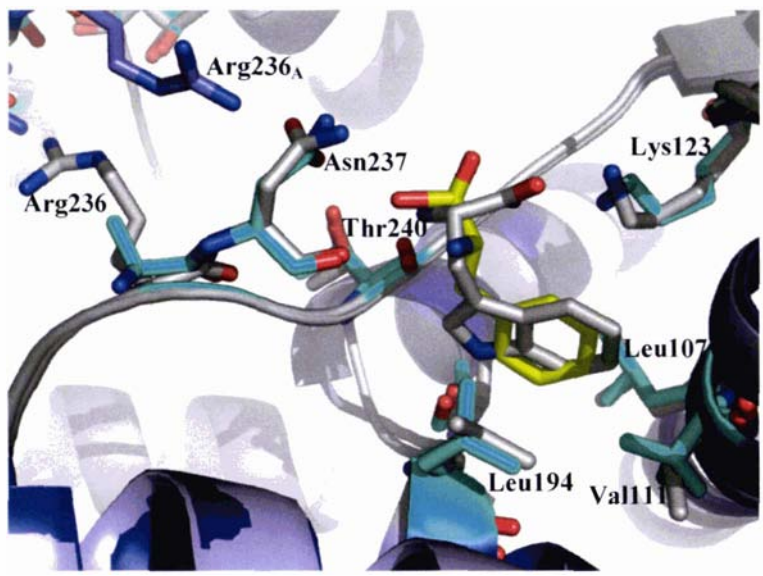


Figure 4.14 **Superposition of Phe and Trp and Phe crystal structures showing the binding of Phe (yellow) and Trp (grey) in subunit B, respectively.** Residues making up the 'Trp' binding site from both structures are shown as sticks (Phe structure is in cyan and Trp and Phe structure is in grey). Arg236 from subunit A is shown in light blue. Note that subunit A is very similar.

It should be noted that there are two Phe molecules in subunit B of the Trp and Phe structure possibly due to the high concentration of amino acid (~2 mM) used to soak crystals in. The second Phe molecule is held in place through interactions with its carboxylate group and Arg23 (α0a helix) and Arg256 (α3-β4 loop). Its amino group hydrogen bonds with N^{ε} and $N^{\eta 2}$ of Arg256 and several hydrophobic residues, including Leu18, move out of the way allowing space for the hydrophobic ring.

We propose that Tyr binds in the same place as Phe binds in *Mt*-DAH7PS due to the synergistic inhibition displayed in the presence of Trp and Tyr (refer to Section 4.3). Analysis of the Trp and Phe structure indicates several polar residues that could easily interact with the hydroxyl group of Tyr, including a water molecule that separates the two Phe molecules and is 4 Å away from the C^{ζ} atoms. The addition of a hydroxyl group to Phe would allow both amino acids to interact with this water. There are also some nearby main-chain CO groups that may also be able to interact with the hydroxyl group of a tyrosine. Interestingly, in *S. cerevisiae* a single residue dictates the enzyme's specificity for Tyr or Phe regulation. The mutation of a conserved glycine (Tyr-regulated DAH7PSs) to a serine (Phe-regulated DAH7PSs) and vice versa leads to a complete change in regulation pattern without affecting enzyme kinetics (refer to Chapter One for further detail).[49]

4.9 Structural Changes in Oligomerization Interfaces upon Trp and Phe Binding

4.9.1 Tight Dimer Interface

The tight dimer interface is made up of contributions from residues of the extended N-terminus (residues 1-13), parts of the α0b and the α0b-α0c loop, together with residues of the core α1 and α2 helices. Upon Trp and Phe binding there are no significant changes in solvent-accessible surface area buried upon formation of the tight dimer. However, there are changes in the interactions between residues of each monomer. The binding of Trp and Phe simultaneously results in a subtle movement of residues of the α2 helix. This allows Ser189 to form new interactions with Ser62 (Ser189 OH...Ser62 OH (3.0 Å) and Ser189 CO...Ser62 OH (2.8 Å) of the other subunit making up the tight

dimer. This interaction between Ser62 and Ser189 is also observed in the Trp structure but not in the Phe structure.

Another interaction gained at the dimer interface upon inhibitor binding is between Arg236 and Asn237 (Arg236 $N^{\eta 2}$...Asn237 O^{δ} (2.5 Å)) of the other monomer; the latter residue also contributes to the tetramer interface (discussed in Section 4.9.2). In the Phe and Se-Met crystal structures Arg236 is disordered in both subunits in the asymmetric unit. In the Trp structure $Arg236_A$ interacts with $Asn237_B$ and $Arg236_B$ is disordered.

The largest movement upon Trp and Phe binding is seen by residues of the extended N-terminus (1-10), which are displaced up to 6 Å in subunit B and 2.5 Å in subunit A. The difference in movement between subunits is due to the displacement of Thesit by the two Phe molecules (figure 4.15). In the Se-Met structure the N-terminal arms contribute about 40 % of the total buried surface in the dimer; however, upon the binding of Trp and Phe this is reduced to just below 30 % with the extra 10 % made up with contributions from Ser189 (α2 helix), Arg236 and Asn237 (α2b-β3 loop). Binding of the Trp and Phe also results in the loss of a salt-bridge ($Asp10_A$....$Arg171_B$) which appears to destabilize the β0-α0a loop so that residues 11-15 are disordered in the Trp and Phe structure (figure 4.15).

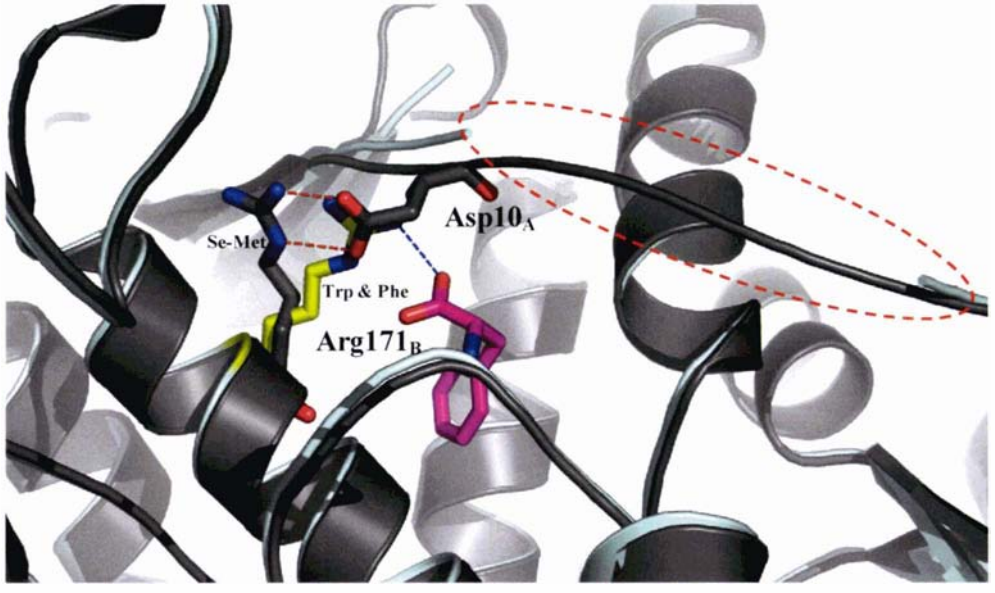


Figure 4.15 **The binding of Phe (pink) in the Trp and Phe structure (light cyan) results in a movement of the sidechain of $Arg171_B$ (from the α2 helix) (grey in Se-Met structure and yellow in Trp and Phe structure), which disrupts the**

interaction (shown in red dashed lines) between this residue and $Asp10_A$ observed in the Se-Met structure (grey). Residues 11-16 in the Trp and Phe structure are disordered and are circled in this figure.

As mentioned in the previous section, Phe is located near the interface of the tight dimer. The two Phe molecules are positioned so that the closest contact between the two amino acids is via their C^{ζ} atoms with a distance of 7 Å. Interestingly, the two Phe molecules displace the Thesit molecule seen in the Se-Met crystal structure (figure 4.16). The binding of Thesit in the Se-Met structure appears to have perturbed the crystal structure generating an unsymmetrical dimer interface. To date we have no wild-type crystal structure without Thesit bound; therefore, a crystal structure of *Mt*-DAH7PS purified without Thesit is necessary to confirm that the binding of the detergent molecule induces asymmetry in the N-terminus of the two subunits in the asymmetric unit. In the Se-Met structure, Thesit interacts with several protein residues: the hydrophilic end of the detergent molecule is found in subunit A whereas its hydrophobic end is located in subunit B. We propose that the interactions that the hydrophilic end of the molecule makes with surrounding proteins mimic the interactions that the carboxylate and amino groups of Phe have with residues in subunit A. In particular, O16 and O19 form hydrogen bonds with Asn175, Phe91 and Arg171 (Thesit O19...$O^{\delta 1}$ Asn175 (2.8 Å), Thesit O19...CO Phe91 (2.7 Å), Thesit O16...$N^{\delta 2}$ Asn175, (3.1 Å), and Thesit O16...N^{ε} Arg171 (2.8 Å)). These interactions by the detergent molecule give a plausible explanation for why we see structural changes in subunit A and not B of the tight dimer. Kinetic studies have shown that Thesit does not have any detectable effect on the activity of *Mt*-DAH7PS (refer to Section 4.3). We hypothesize that the binding of Thesit triggers an interaction with the residues of subunit A but not B and that both binding sites must be occupied to see an effect on the activity of *Mt*-DAH7PS. Upon Phe binding the N-terminus becomes more symmetrical due to both effector molecule binding sites being occupied and eliciting structural rearrangement in both monomers of the tight dimer.

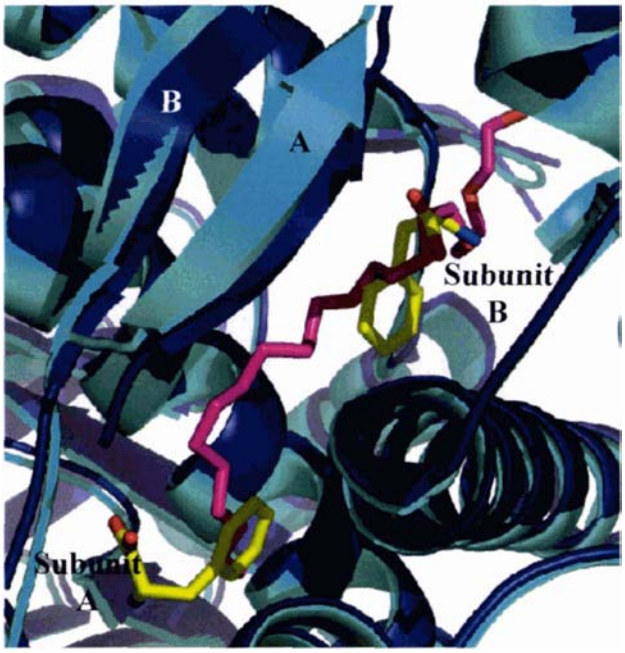


Figure 4.16 **The Thesit molecule (pink) in the Se-Met structure (blue) is displaced by two Phe molecules (yellow) in the Trp and Phe structure (cyan).** The extended β-strand (residues 3-9) in subunit A and B that make up the antiparallel β-ribbon are labeled as A and B, respectively.

4.9.2 Tetramer Interface

Upon binding of Trp and Phe there do not appear to be any significant changes in residues contributing to the tetramer interface (residues from α1 and α2b helices, and the α2b-β3 loop), although these residues undergo considerable structural changes, especially residues of the α2b-β3 loop (a list of distances moved by residues of the α2b-β3 loop upon Trp and Phe binding is in table 4.7). The total surface area buried upon tetramer formation also remains relatively unchanged. However, upon Trp and Phe binding there is a gain of a hydrogen bond ($Asn237_{A(B)}$ N^{δ}...CO $Cys231_{C(D)}$) between subunits that make up the relatively small, predominately hydrophobic tetrameric interface (refer to Chapter Three, Section 3.17.2). Asn237, whose O^{δ} hydrogen bonds to the amino group of Trp (refer to Section 4.1), also forms a new interaction with Arg236 from the other subunit of the tight dimer (CO of $Asn237_{A(B)}$...N^{η} of $Arg236_{B(A)}$) (figure 4.17) refer to Section 4.8). The Trp and Phe structure clearly indicates that Asn237 allows communication between monomers of the tight dimer upon Trp binding at the tetrameric interface. Whether Asn237 plays a crucial role in relaying the signal of

Trp binding back to the active site awaits site-directed mutagenesis of this residue along with functional and structural studies.

Residue	Movement of residues of the α2b-β3 loop upon Trp and Phe binding (Cα atoms (Å))	
	Subunit A	Subunit B
Val233	2.2	1.9
Ala234	4.3	4.6
Asp235	3.0	2.9
Arg236	0.7	1.6
Asn237	4.2	4.8
Leu238	4.0	4.3
Gln239	6.8	7.5
Thr240	3.4	3.2
Ala241	0.4	0.5
Glu242	0.3	0.3

Table 4.7 **Distances residues of the α2b-β3 loop have moved upon Trp and Phe binding, in relation to that observed in the Se-Met structure.** The slight differences between the two subunits in the asymmetric unit are listed (Subunit A and B).

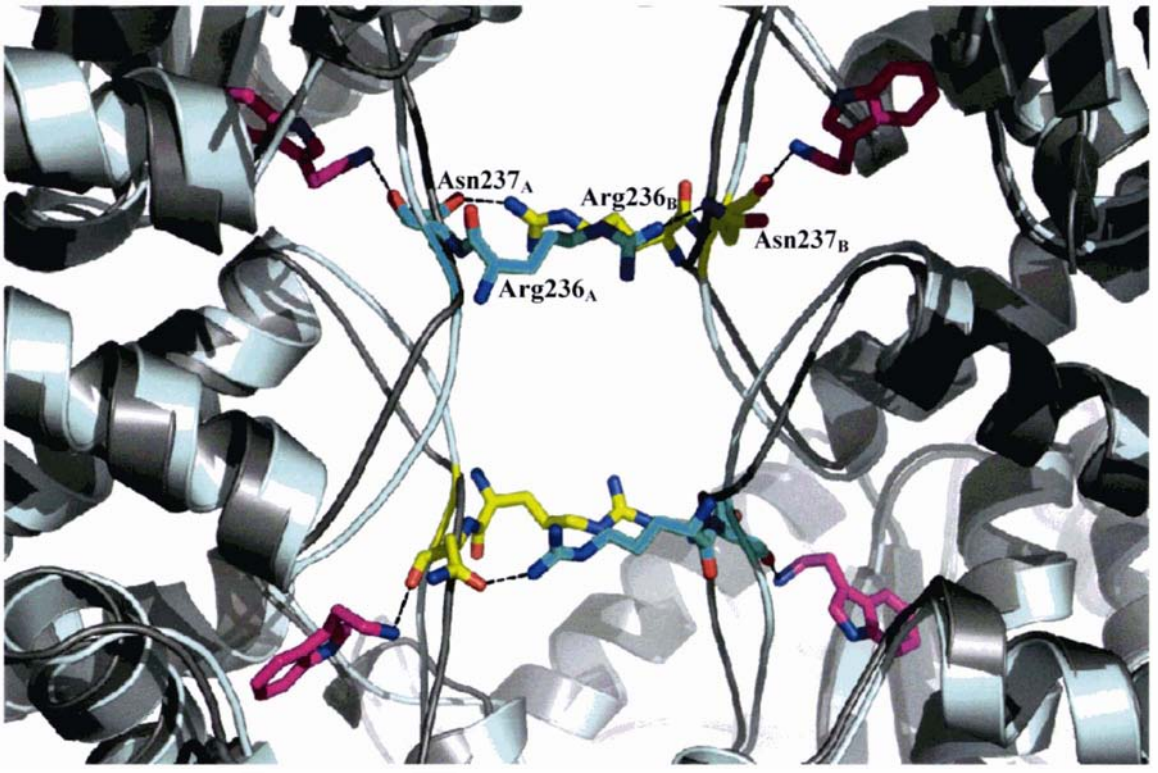


Figure 4.17 Tetramer interface of *Mt*-DAH7PS. Interaction between Asn237 and the Trp molecule and between Asn237 and Arg236 (of the other subunit making up the tight dimer) in the Trp and Phe structure. The Trp and Phe structure is in cyan and the Se-Met structure is in grey. Trp is shown in pink. Asn237 and Arg236 of chain A (and chain D) are shown in cyan and the same residues are shown in yellow for chain B (and chain C).

4.10 The Transmission of Inhibitory Signal

Allosteric Inhibition

Allosteric inhibition is defined as the binding of an inhibitor that influences conformational changes in the enzyme rather than forming a dead-end complex with the enzyme. These enthalpic or entropic changes in the enzyme may alter the binding characteristics of the enzyme for a substrate or the on-going reaction characteristics (or both). If the binding characteristics alone are affected, V_{max} will usually remain unchanged, so the pattern could be regarded as competitive. Similarly, other forms of allosteric inhibition, where V_{max} is altered, could be regarded as giving non-competitive or mixed inhibition (K_M is also altered). However, in most cases Michaelis-Menten kinetics are not obeyed in the presence of allosteric inhibitors so the terms competitive, non-competitive and mixed inhibition are not applicable.[159] Allosteric inhibition is a

form of negative heterotropic cooperativity where the binding of the inhibitor decreases the affinity of the protein for a substrate/ligand.

4.10.1 Information from Functional Studies

To investigate how the binding of Trp and Phe inhibit the activity of *Mt*-DAH7PS Michaelis-Menten plots were drawn from experiments performed at a fixed concentration of one substrate and varying concentrations of the second substrate in the presence of Trp and Phe. Figure 4.18 shows initial velocity values as a function of PEP concentration at a fixed concentration of E4P and either no aromatic amino acids, 200 μM Trp, or 100 μM Trp + 100 μM Phe. The plot shows that Trp and Phe significantly alter the maximum velocity and only slightly change the apparent K_M with respect to PEP. This would indicate that Trp and Phe inhibition is mixed with respect to PEP, which is consistent with what is observed for *Ec*-DAH7PS(Phe) (refer to Section 4.12). Figure 4.19 shows initial velocity values as a function of E4P concentration at a fixed concentration of PEP and either no aromatic amino acids, 200 μM Phe, 200 μM Trp, or 100 μM Trp + 100 μM Phe. The plot shows that at low concentrations of E4P, Phe and Trp (to a lesser extent) inhibit the rate of reaction of *Mt*-DAH7PS. The most significant finding is that the Michaelis-Menten plot becomes less hyperbolic and more sigmoidal in the presence of Trp and Phe. Although the sigmoidal nature is not obvious in the data presented in figure 4.19 the data was fitted to the Hill equation which produced a Hill coefficient of 2.3 ± 0.1. This is consistent with a sigmoidal curve and indicates cooperativity with respect to E4P. Thus the rate of reaction is significantly inhibited at low concentrations of E4P and only moderately inhibited at higher concentrations. This is different to that observed for *Ec*-DAH7PS(Phe) in which Phe inhibition is competitive with respect to E4P (refer to Section 4.12).

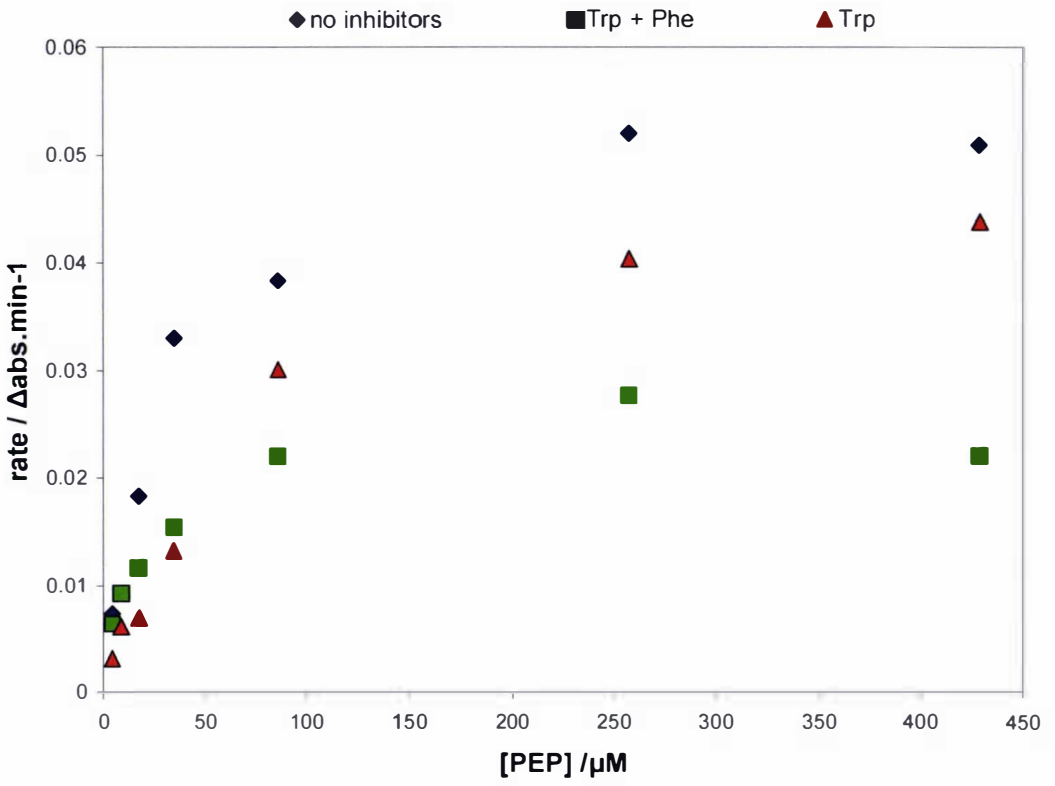


Figure 4.18 **Plot of rate (Δabs/min) against PEP concentration at a constant concentration of E4P and either no aromatic amino acids, Trp or Trp and Phe.** The reaction consisted of E4P (185 μM), $MnSO_4$ (100 μM) and PEP (4 to 429 μM), in BTP (50 mM) pH 7.5 buffer and either no aromatic amino acids, 200 μM Trp or 100 μM Trp and 100 μM Phe. The reaction was initiated by the addition of purified *Mt*-DAH7PS (2 μL, 2.6 mg/mL) and was carried out at 30 °C.

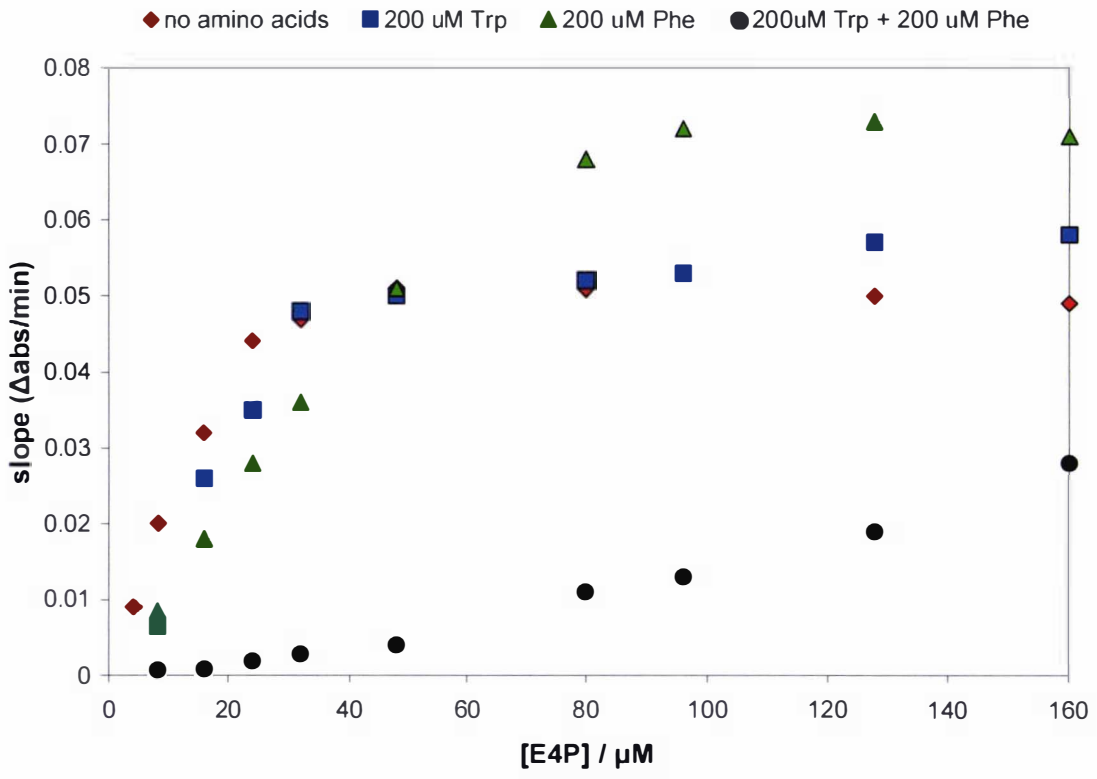


Figure 4.19 **Plot of rate (Δabs/min) against E4P concentration at a constant concentration of PEP and either no aromatic amino acids, Phe, Trp or Trp and Phe.** The reaction consisted of PEP (200 µM), $MnSO_4$ (100 µM) and E4P (8 to 160 µM), in BTP (50 mM) pH 7.5 buffer and either no aromatic amino acids, 200 µM Phe, 200 µM Trp or 100 µM Trp and 100 µM Phe. The reaction was initiated by the addition of purified *Mt*-DAH7PS (purified in the absence of Thesit) (2.5 µL, 2 mg/mL) and was carried out at 30 °C.

The crystal structures of *Mt*-DAH7PS show that the four actives sites of the tetrameric protein are clearly separated and do not appear to interact directly with each other. It would appear that the mechanism of cooperative binding must involve more general interactions between subunits and the occurrence of conformational changes. Cooperative binding in proteins can be explained by the Monod-Wyman-Changeux (MWC) or the Koshland-Némethy-Filmer (KNF) models. In the most simple of cases, a protein can exist in two conformational forms, the T-form, which predominates in the unliganded protein, and the R-form, which predominates in the protein-ligand form. In the case of the MWC model, it is assumed that all of the subunits in the protein in a given conformational state, either the R- or T-form, have identical and independent properties. The two conformational forms are in equilibrium (figure 4.20) in the

absence of ligand, and the ratio of the R- and T- forms is disturbed by the binding of ligand. The MWC equation is consistent with a sigmoidal binding curve even though its derivation assumes that the binding of one molecule of ligand does not affect the affinity for the ligand of other binding sites on the molecule. The explanation lies in the R/T equilibrium. When L_n (the equilibrium constant between R_n and T_n) is large the equilibrium is in the favor of the T-form in the absence of ligand. If ligand is introduced at low concentrations there is not enough to react significantly with the R-form; however, at higher concentrations of ligand there is enough to force the formation of significant amounts of R_{ligand} so that some R will be removed from the system, disturbing the R/T equilibrium and causing the conversion of T to R. Therefore, the overall binding curve will be sigmoidal. According to the MWC model allosteric *inhibitors* bind to the T-form of the enzyme and stabilize it and thus increase the value of L_n. Allosteric *activators* bind and stabilize the R-form and decrease value of L_n. The KNF model allows for hybrids between the two conformational forms of the protein. An extra complication is that you need to be able to identify which subunits can interact with each other. This model allows allosteric modifiers to act in a variety of different ways.[159]

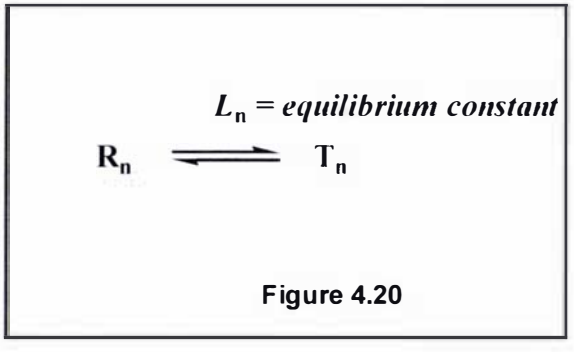


Figure 4.20

The interpretation of our kinetic results is complicated as we are dealing with a two substrate-two inhibitor system. What it does show us is that the synergistic inhibition displayed in the presence Trp and Phe is due to a change in the processing of E4P by *Mt*-DAH7PS. Using the MWC model it can be proposed that *Mt*-DAH7PS exists in two forms; an R-form, which has a high affinity for E4P, and a T-form, which has a lower affinity for E4P. If the KNF model is obeyed it is quite possible that there are more than two conformational forms. The activation observed by Phe (refer to table 4.1) can

be explained by Phe binding to and stabilizing the R-form of the enzyme. The inhibition by Trp and Phe can be explained by the two aromatic amino acids binding to and stabilizing the T-form of *Mt*-DAH7PS.

4.10.2 Information from Structural Studies

Functional studies have shown that activity of *Mt*-DAH7PS is altered by all three aromatic amino acids, with Phe and Tyr (to a lesser extent) activating the enzyme and the combination of either of these aromatic amino acids with Trp having a synergistic inhibitory effect on *Mt*-DAH7PS activity. Further kinetic studies indicated that the presence of Trp and Phe influences the processing of E4P by *Mt*-DAH7PS. The crystal structure with Trp and Phe bound clearly indicates two distinct binding sites for Trp and Phe. Analysis of the Trp, Phe, Trp and Phe, and Se-Met crystal structures has provided some clues as to the communication between the 'Phe' and 'Trp' binding sites on *Mt*-DAH7PS.

The binding of a molecule of Thesit at the tight dimer interface in the Se-Met and Trp crystal structures has made it difficult to interpret structural changes due to the binding of aromatic amino acids. The binding of the hydrophilic end of the detergent molecule in subunit A of the asymmetric unit has induced structural changes which we do not see in subunit B where the hydrophobic end of Thesit is found. As we do not have a crystal structure without Thesit we have made the assumption that any differences observed in subunit B among the four crystal structures is due to the binding of the respective aromatic amino acids.

Upon comparison of the N-terminal residues of *Mt*-DAH7PS (residues 1-20) we noticed that residues 1-17 in the Se-Met, Phe, and Trp structures are very similar but are quite different in the Trp and Phe structure (figure 4.21). It is also apparent that in the Trp and Phe structure residues 11-15 are disordered in both subunits in the asymmetric unit, which is not observed for the other three structures. Another interesting observation is that in the Trp and Phe structure there is a rearrangement in residues that contribute to the tight dimer interface (between subunit A and B) in comparison with the other three structures. This rearrangement in residues does not have a significant effect on the total surface area buried upon dimer formation. In the Se-Met, Phe and Trp structures, the N-

terminal residues (1-16) contribute ~40 % of the total buried surface in the tight dimer and residues Arg236 and Asn237 do not make a contribution. However, in the Trp and Phe structure the contribution by the N-terminal residues is reduced to ~30 % with the missing ~10 % made up by contributions from Arg236, Asn237 and Ser189.

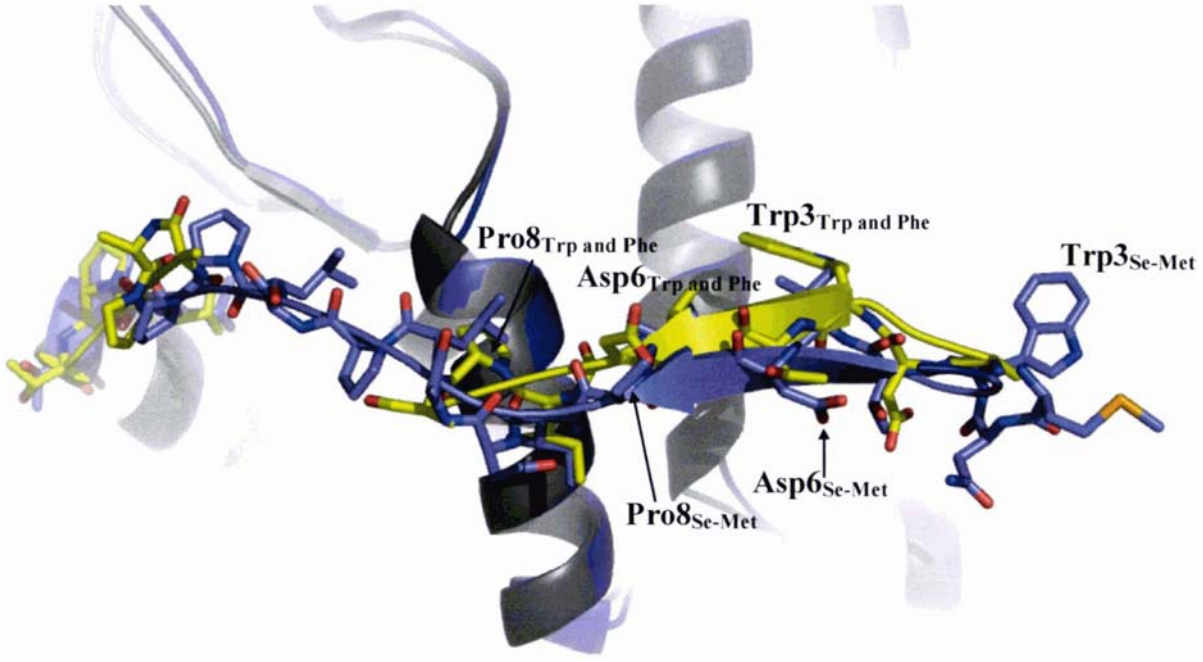


Figure 4.21 **Superposition of N-terminal residues (1-20) of subunit B in Se-Met (blue) and Trp and Phe (yellow) crystal structures.**

The inhibition of *Mt*-DAH7PS activity in the presence of Trp and Phe suggests that there is communication between the 'Trp' and 'Phe' binding sites on the enzyme. In the Trp and Trp and Phe structures, the binding of Trp results in a subtle movement of residues of the α2 helix, including Ser189, which is now able to interact with Ser62 of the other subunit of the tight dimer. Interestingly, the residues contributing to the Phe binding site are predominately from the α2 helix. We are still unsure as to how the simultaneous binding of the two amino acids induces an inhibitory signal that is relayed back to the active site of *Mt*-DAH7PS, but functional studies have indicated that the affinity of the enzyme for E4P is affected. A reason for the lack of structural changes in the active site of *Mt*-DAH7PS in complex with Trp and Phe could be because the aromatic amino acids were soaked into the crystal and if the protein was locked into a conformation due to crystal packing this may have prevented structural movements in the crystal. However, the solvent content of the crystals is high (68 %), which would

suggest that this is probably not the case. It is possible that further screening of crystallization conditions would lead to successful co-crystallization of *Mt*-DAH7PS with Trp and Phe.

The structure with both Trp and Phe bound has indicated residues, including Asn237, Arg236, Ser189 and Arg171, that may be involved in the relay of the signal back to the active site. Mutation of these residues, together with further structural and functional studies, is necessary to understand the complex regulatory mechanism of *Mt*-DAH7PS.

4.11 Allosteric Regulation of Type Iα *Ec*-DAH7PS(Phe)

A crystal structure of Se-Met substituted *Ec*-DAH7PS(Phe) in complex with Phe was determined using co-crystallization methods and refined to a resolution of 2.8 Å.[54] A model of allosteric regulation of DAH7PS from *E. coli* has been deduced, which appears to be quite different to what is observed for *Mt*-DAH7PS. Two interrelated paths of conformational changes have been proposed to transmit the inhibitory signal from the Phe binding site to the active site of *Ec*-DAH7PS. The first path involves transmission within a single subunit (repositioning of the β6a/β6b segment enhances its interaction with the β3-α3 loop, which in turn weakens the interaction of loop β3-α3 with loop β2-α2 significantly, figure 4.22). The second path involves alteration in the contacts between subunits (Phe binding weakens interactions between β4-α4 (from the other subunit making up the dimer) and β2-α2 (both contribute residues coordinating PEP and E4P), which does not directly affect the active site but instead changes its nearby residues). The combination of these two paths changes the conformation of one of the active site loops significantly and shifts the other slightly. Phe inhibition is competitive with respect to E4P and mixed with respect to PEP, explaining the lack of SO_4^{2-} in the active site (believed to mimic phosphate group of E4P) and flipped orientation of PEP observed in the crystal structure with Phe.[54] In the case of *Mt*-DAH7PS, the pathway for transmission of the signal is not as easy to deduce using the current crystal structures that we have. However, functional studies reveal that Trp and Phe inhibit the enzyme by altering the enzyme's affinity for E4P (refer to Section 4.10.1).

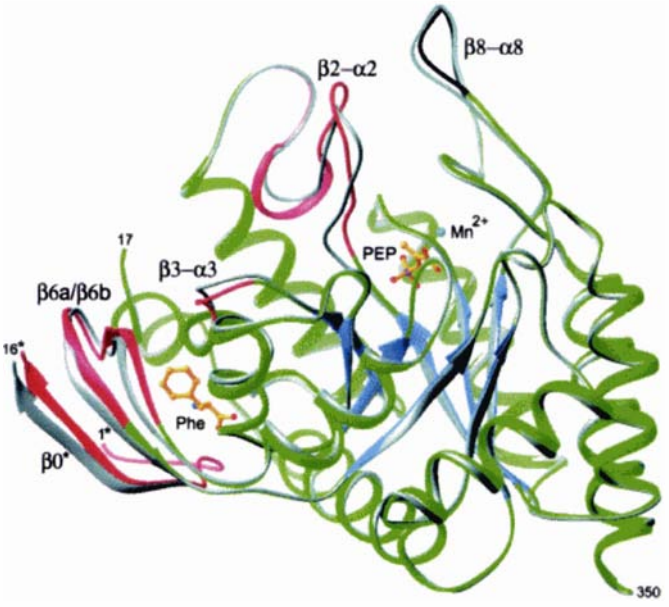


Figure 4.22 **Structural changes in *Ec*-DAH7PS(Phe) upon Phe binding.** The protein chain (green, blue, red, grey), Mn^{2+} (cyan), PEP, and Phe (both in gold) are shown for the +Phe DAH7PS. The superimposed main-chain of the –Phe enzyme is shown in black. The main-chains of the four segments colored in red shift significantly upon Phe binding. These four segments are involved in transmission of the inhibition signal within a subunit from the Phe binding site to the active site. The change in conformation of the β8-α8 loop segment, shown in dark grey, is attributed to differences in crystallization conditions.[54]

4.12 Regulation and Quaternary Structure

The quaternary structures among various DAH7PS enzymes appear to be very different. The type Iα and type Iβ DAH7PSs that are feedback-regulated by the aromatic amino acids are homotetrameric, with one monomer-monomer interface conserved between these two subfamilies. The common interface primarily involves helices α3, α4 and α5, and loops β2-α2, β3-α3, β4-α4 and β5-α5, and impacts on the positioning of the β2-α2 loop that provides part of the E4P binding site in the active site of the protein. On the other hand *Mt*-DAH7PS's monomer-monomer interface shows no commonality with any subunit interface found in any previously reported DAH7PS structure. Unlike *Ec*-DAH7PS, residues of the β2-α2 loop do not contribute to any of the inter-subunit associations, providing further evidence for a novel regulatory mechanism. The relationship between quaternary structure and feedback-inhibition needs to be investigated. DAH7PS enzymes that are feedback-regulated by the aromatic amino acids are tetrameric, as can be seen in the case of the two Iα enzymes and the type Iβ

enzyme from *T. maritima*. The *P. furiosus* enzyme is unregulated and although it crystallizes as a tetramer, it is dimeric in solution like the unregulated *H. pylori* enzyme (refer to Chapter Two, Section 2.6). Intriguingly, *Hp*-DAH7PS shares 67 % sequence similarity (45 % sequence identity) with the synergistically inhibited *Mt*-DAH7PS (figure 4.23), and comparison of the two amino acid sequences indicates that both extensions to the barrel are present in *Hp*-DAH7PS.

Figure 4.23 A ClustalW alignment of *Hp*-DAH7PS (top line) and *Mt*-DAH7PS (bottom line).

```
Mtu     MNWTVDIPIDQLPSLPPLPTDLRTRLDAALAKPAAQQPTWPADQALA-MRTVLESVPPVT 59
Hpj     ----------MSNTTWSPTSWHS-------FKIEQHPTYKDEQELERVKKELRSYPPLV 42
                  :.. .  **. ::           *:**:  :* *  ::. *.* **:.

Mtu     VPSEIVRLQEQLAQVAKGEAFLLQGGDCAETFMDNTEPHIRGNVRALLQMAVVLTYGASM 119
Hpj     FAGEARNLQERLAQVIDNKAFLLQGGDCAESFSQFSANRIRDMFKVMMQMAIVLTFAGSI 102
        ...*  .***:**** ..:***********:* : :  :**. .:.::***:***:..*:

Mtu     PVVKVARIAGQYAKPRSADIDALG---LRSYRGDMINGFAPDAAAREHDPSRLVRAYANA 176
Hpj     PIVKVGRIAGQFAKPRSNATEILDDEEVLSYRGDIINGIS--KKEREPKPERMLKAYHQS 160
        *:***.*****:*****   : *.   : *****:***::     ** .*.*:::** ::

Mtu     SAAMNLVRALTSSGLASLHLVHDWNREFVRTSPAGARYEALATEIDRGLRFMSACGVADR 236
Hpj     VATLNLIRAFAQGGLADLEQVHRFNLDFVKNNDFGQKYQQIADRITQALGFMRACGVEIE 220
         *::**:**::..***.*. ** :* :**:..  * :*: :* .* :.* ** ****  .

Mtu     N---LQTAEIYASHEALVLDYERAMLRLSDGDDGEPQLFDLSAHTVWIGERTRQIDGAHI 293
Hpj     RTPILREVEFYTSHEALLLHYEEPLVRK---DSLTNQFYDCSAHMLWIGERTRDPKGAHV 277
        .   *: .*:*:*****:*.**..::*    *.   *::* *** :*******: .***:

Mtu     AFAQVIANPVGVKLGPNMTPELAVEYVERLDPHNKPGRLTLVSRMGNHKVRDLLPPIVEK 353
Hpj     EFLRGVCNPIGVKIGPNASVSEVLELCDVLNPHNLKGRLNLIVRMGSKIIKERLPKLLQG 337
         * : :.**:***:*** : . .:*  : *:***  ***.*: ***.: ::: ** :::

Mtu     VQATGHQVIWQCDPMHGNTHESSTGFKTRHFDRIVDEVQGFFEVHRALGTHPGGIHVEIT 413
Hpj     VLKEKRHILWSIDPMHGNTVKTNLGVKTRAFDSVLDEVKSFFEIHRAEGSLASGVHLEMT 397
        *    ::::*. ******* ::. *.*** ** ::***:.***:*** *: ..*:*:*:*

Mtu     GENVTECLGGAQDISETDLAGRYETACDPRLNTQQSLELAFLVAEMLRD--- 462
Hpj     GENVTECIGGSQAITEEGLSCHYYTQCDPRLNATQALELAFLIADMLKKQRT 449
        *******:**:* *:* .*: :* * ******: *:******:*:**:.
```

* Conserved residues	: Conservative substitutions	. Semi-conservative substitutions

The difference in quaternary structure between *Hp*-DAH7PS, which is unregulated, and *Mt*-DAH7PS, which is regulated by all three of the aromatic amino acids, needs to be investigated. Analysis of the *Mt*-DAH7PS crystal structures in the presence of Trp and Phe suggests that Arg236, Asn237, Ser189, Ser62 and Arg171 may be involved in relaying the signal of Trp and Phe binding back to the active site. Interestingly, none of

these residues are conserved in *Hp*-DAH7PS (table 4.8). Of particular note are the substitutions of serine for glycine, which no longer can form hydrogen-bonds that Ser-OH atoms can form. A crystal structure of *Hp*-DAH7PS is necessary for determining why there is a difference in regulation between the dimeric and tetrameric proteins.

Residue in *Mt*-DAH7PS	**Equivalent residue in *Hp*-DAH7PS**
Arg171	Lys155
Ser62	Gly45
Ser189	Gly173
Arg 236	Glu220
Asn237	Arg221

Table 4.8 Comparison of residues that may be important for feedback-inhibition of *Mt*-DAH7PS with corresponding residues in *Hp*-DAH7PS (from ClustalW sequence alignment in figure 4.22).

4.13 Significance of Regulating DAH7PS

The regulation of DAH7PS is essential for maintaining control of cellular levels of aromatic compounds in microorganisms and plants. It has previously been shown that feedback-inhibition of DAH7PS is the main mechanism for controlling carbon flow into the shikimate pathway.[46] Structural and functional studies with the type II *M. tuberculosis* enzyme and the type I *E. coli* enzyme have shown that the binding of specific aromatic amino acids to their respective allosteric binding site on the enzyme inhibits catalysis by interfering with E4P binding. It makes physiological sense that feedback-inhibition affects the binding of E4P and not PEP because cellular levels of PEP exceed those of E4P. Furthermore, previous studies (*Ec*-DAH7PS(Phe)) have shown that, in the absence of PEP, E4P inactivates the enzyme through a reversible Schiff-base reaction with an active site lysine.[88] The intracellular concentration of PEP in the cell ranges from 70-150 μM, whereas E4P levels in the cytosol of the cell are undetectable, which is expected because of E4P's tendency to dimerize.[184,185] From the functional studies discussed in this chapter, it appears that when E4P concentrations are low the activity of *Mt*-DAH7PS is significantly inhibited by Trp and Phe or Tyr.

However, at higher E4P concentrations inhibition of this enzyme by the aromatic amino acids does not appear to be so significant.

In the case of *M. tuberculosis*, it could be argued that the regulation of aromatic amino acid biosynthesis (at the protein level) is more efficient than that observed for *E. coli*. *M. tuberculosis* possesses a single DAH7PS with a monomer molecular mass of ~50 kDa that is feedback-regulated by all three aromatic amino acids. At the other extreme *E. coli* has three proteins with monomer masses of ~30 kDa that are each regulated by one of the aromatic amino acids. The activity of *Mt*-DAH7PS is only inhibited when Trp and Phe or Tyr are present, so that if one amino acid is present in excess over the other, chorismate is still able to be synthesized as a precursor of the deficient amino acid. For this to work there would have to be additional feedback regulatory strategies for each of the aromatic amino acid biosynthetic pathways after the branch-point at the common precursor chorismate. Evidence of these regulatory systems is discussed in Section 4.3.

4.14 Conclusions and Future Studies

DAH7PS is one of the most diversely regulated metabolic pathway enzymes, with significant variations in allosteric binding sites between DAH7PSs from different organisms. In *E. coli* there are three DAH7PS isoenzymes, each sensitive to feedback-regulation by one of the three aromatic amino acids. DAH7PS from *P. furiosus*, possessing only the core $(\beta/\alpha)_8$ barrel, is reported to be unregulated and at the other extreme we have the highly decorated DAH7PS from *M. tuberculosis* which is feedback inhibited by Trp, Phe and Tyr. The determination of the crystal structure of *Mt*-DAH7PS in complex with Trp and Phe confirms the presence of two distinct regulatory binding sites formed by the two major additions that decorate the barrel. Our kinetic studies reveal a novel mechanism for allostery, involving the synergistic feedback-inhibition by Trp and Phe or Tyr. Structural studies indicate that both Trp and Phe can bind to the enzyme; however, it is only when they are present in combination that we see significant inhibition of catalytic activity. Future structural studies of *Hp*-DAH7PS are essential to shed some light into the differences in regulation of the *M. tuberculosis* and *H. pylori* enzymes, as at the primary sequence level the two proteins appear to be similar.

CHAPTER FIVE

INVESTIGATING THE ROLE *MT*-DAH7PS ARG284 AND CYS440 PLAY IN CATALYSIS

5.1 Introduction

The active site of *Mt*-DAH7PS consists of many protein residues that play an essential role in the positioning of PEP, E4P and divalent metal ion, and in enzyme catalysis. The determination of the crystal structure of *Mt*-DAH7PS has allowed us to identify key residues that contribute to the active site of the *M. tuberculosis* enzyme. Two of the active site residues that attracted our attention were Arg284 and Cys440. To investigate the role these two active site residues play in catalysis, site-directed mutagenesis studies were performed and the resulting mutant proteins were then functionally and structurally characterized.

This chapter describes the site-directed mutagenesis of R284K and C440S, enzyme purification, functional characterization and crystal structures of the two *Mt*-DAH7PS mutants. The potential role these two active sites residues play in catalysis is also discussed. Aspects of the crystal structure of native *Mt*-DAH7PS are also discussed in this chapter.

5.1.1 Choice of Residues for Mutation

Mutation of Mt-DAH7PS Arg284 (R284K)

In the Se-Met crystal structure of *Mt*-DAH7PS, Arg284 appears to bridge between the PEP and SO_4^{2-} binding sites, being hydrogen-bonded to both the PEP phosphate and SO_4^{2-}. The dual binding to both substrates and the fact that this arginine is absolutely conserved across all types of DAH7PS are the reasons why we investigated the importance this residue plays in catalysis. The mutation of the arginine to a lysine, resulting in the loss of one guanidine group, was performed to see if the removal of one hydrogen-bonding partner would influence this residue's interaction with PEP and SO_4^{2-}. The determination of the crystal structure of R284K will allow us to investigate the

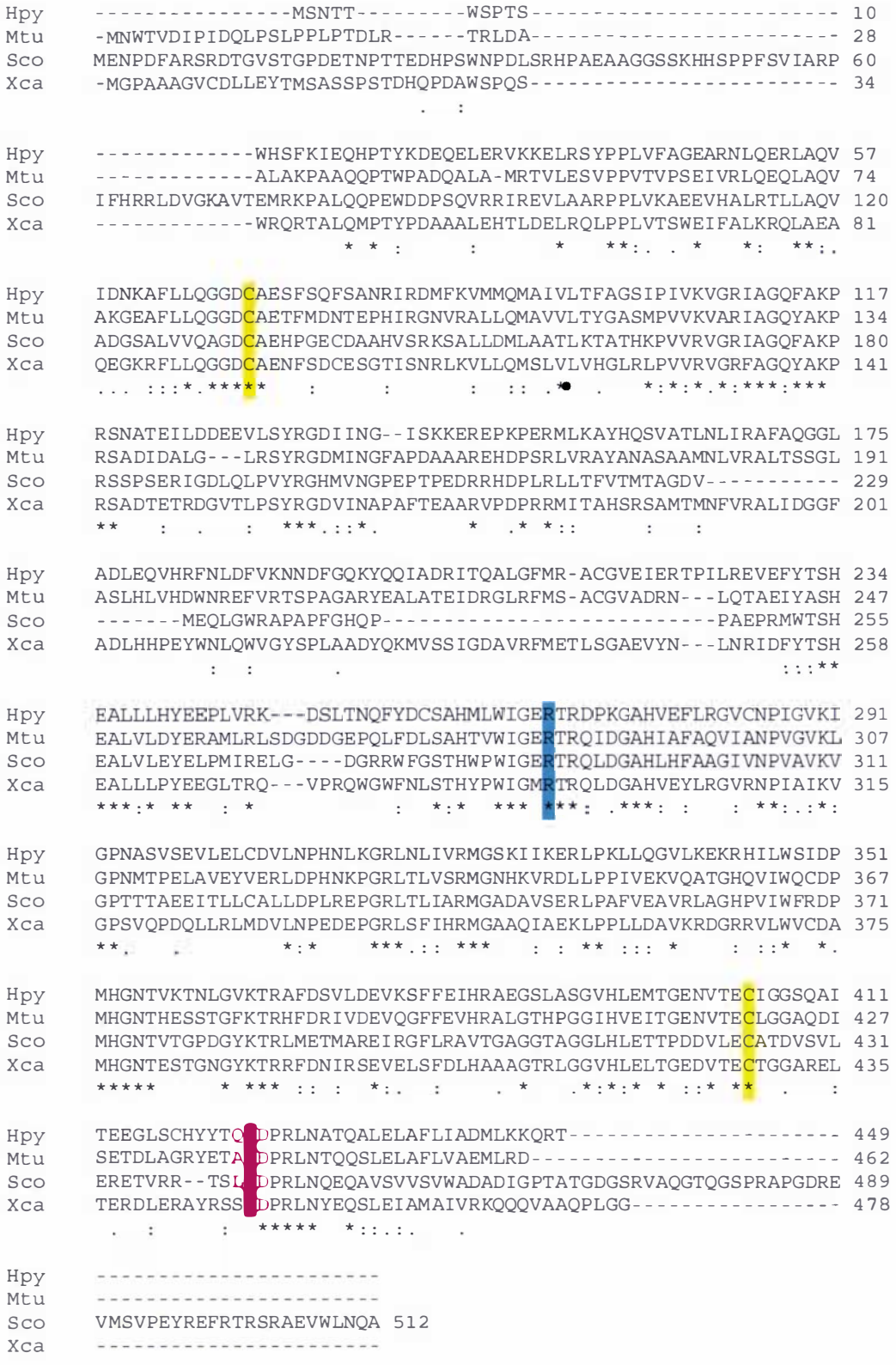
interactions R284K forms with PEP and SO_4^{2-}. Comparison of the affinity of this mutant for both substrates with that of wild-type protein will allow us to observe the influence this single amino acid substitution has on enzyme catalysis.

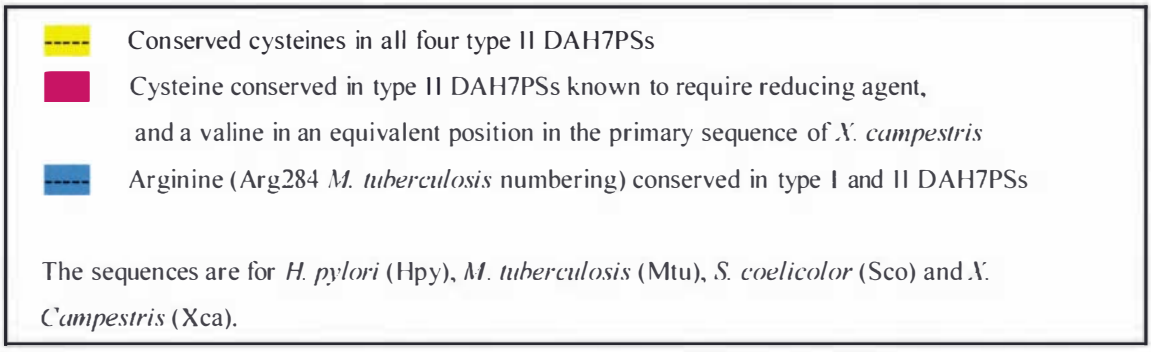
Mutation of Mt-DAH7PS Cys440 (C440S)

In Chapter Two and Three the necessity for a reducing agent to maintain fully active *Hp*-DAH7PS and *Mt*-DAH7PS activity was discussed. The crystal structure of *Mt*-DAH7PS indicates the presence of a second cysteine (Cys440) close to the metal-binding cysteine (Cys87). The disulfide formation between Cys87 and Cys440 has been captured in the crystal structure of native *Mt*-DAH7PS, described in Section 5.5. Similar to crystallization conditions of all *Mt*-DAH7PS proteins discussed in this thesis, native protein was purified in the presence of TCEP. However, the crystals were left for a while before diffraction data were collected from them, and as no reducing agent was added to the mother liquor they were able to oxidize. The comparison of the Se-Met and native *Mt*-DAH7PS structures allows us to propose that under oxidizing conditions there is a movement about Cα-Cβ of Cys440, disrupting its interaction with Lys133 to form a disulfide bond with Cys87. It is possible that this disulfide formation results in the exclusion of metal-binding, explaining the requirement of reducing agent to maintain full activity of *Mt*-DAH7PS.

Figure 5.1 shows a ClustalW alignment of type II DAH7PSs known to require reducing agent to maintain full enzyme activity, *M. tuberculosis*, *H. pylori* and *S. coelicolor*.[37] The type II DAH7PS from *X. campestris* in which reducing agent has no effect on activity is also included.[38] The alignment shows three cysteine residues are completely conserved in all these type II DAH7PSs with the exception of Cys440 (*M. tuberculosis* numbering). This residue is replaced by a valine in *X. campestris*. The *X. campestris* enzyme has been reported to be very unstable and unlike *Hp*-DAH7PS and *Mt*-DAH7PS its instability is not overcome nor is its activity restored by the inclusion of reducing agent in purification buffers.

Figure 5.1 **ClustalW alignment of type II DAH7PSs known to require reducing agent for the maintenance of full activity and DAH7PS from *X. campestris* which appears to be insensitive to reducing agent.**





The importance of Cys440 in enzyme catalysis is investigated using site-directed mutagenesis, structural and functional studies. We hypothesized that by mutating the cysteine residue we would remove the requirement of a reducing agent, as no disulfide bond with the metal-binding cysteine can occur.

The comparison of the crystal structures of *Mt*-DAH7PS, *Pf*-DAH7PS and *Tm*-DAH7PS show that Cys440 (*Mt*-DAH7PS) occupies an equivalent position to that of a serine found in the type Iβ enzymes. Therefore Cys440 was mutated to a serine, as both type Iβ DAH7PSs have not been reported to be sensitive to the presence of reducing agent.

5.2 Cloning, Expression and Purification of *Mt*-R284K and *Mt*-C440S

The R284K and C440S mutations were performed using a QuikChange® II Site-Directed Mutagenesis Kit (Stratagene), and pPro-ExHTa-*Mt*-DAH7PS as the double-stranded plasmid template. The single amino acid substitutions (either an Arg to a Lys or a Cys to a Ser) were introduced by designing synthetic oligonucleotide primers, each complementary to opposite strands of pPro-ExHTa-*Mt*-DAH7PS containing the desired mutation (primer sequences are given in Chapter Six). The resulting plasmids were then electroporated into BL21(DE3) cells that had been previously transformed with pGroESL.[170] The transformed cells were grown, induced and harvested using conditions identical to those described for wild-type *Mt*-DAH7PS protein in Chapter Three.

The purification of *Mt*-DAH7PS R284K and C440S was performed using a two-step procedure, involving passage down a Ni^{2+} affinity column, cleavage of the His_6-tag using rTEV protease, and a final passage down a size exclusion column. The purification was identical to that described for wild-type protein in Chapter Three, except buffers used for the purification of C440S contained no reducing agent. Both mutant proteins appeared to behave in a similar fashion to that observed for wild-type protein throughout the entire purification. The specific activity of R284K appeared to be very similar to that of wild-type protein. However, the specific activity of C440S was very difficult to determine due to the large apparent K_M for E4P (refer to Section 5.10).

5.3 Requirement of Reducing Agent in the Purification of C440S

When C440S was purified using buffers containing no reducing agent, the protein was active but it appeared to be unstable, and gradually lost activity over the duration of the purification. The loss of activity seemed to be a lot slower than that observed for wild-type protein. Incubation of purified C440S with TCEP (0.5 mM) for twenty-four hours increased the specific activity of the protein approximately two-fold, therefore future purifications included the addition of TCEP to all purification buffers.

The requirement of reducing agent to maintain full enzyme activity was unexpected as we had speculated that the mutation to a serine would have removed this. The plasmid was sequenced and the substitution of a cysteine to a serine was successfully achieved. There are five cysteine residues in wild-type *Mt*-DAH7PS: Cys87, a metal-binding ligand; Cys420, which is highly conserved in type II DAH7PSs; Cys231, which contributes to the tetramer interface of the protein; Cys365; and Cys440, under investigation in this study. Cys420 and Cys 365 are unlikely to form disulfides as analysis of the Se-Met *Mt*-DAH7PS crystal structure indicates that no other cysteine residues are in the vicinity of these cysteines. On the other hand, the distance between the two Cys231 residues across the tetramer interface is approximately 12 Å, a slight movement of residues of the α2b-β3 loop/α2b helix which contribute residues to the tetramer interface may allow disulfide linkage between these two cysteine residues.

Analysis of chromatograms from size exclusion columns of C440S run with and without TCEP (figure 5.2) show that in the absence of TCEP there is more protein being eluted from the column as higher molecular weight protein (>1000 kDa) than protein predicted to be tetrameric (approximately a 2:1 ratio) (for further details refer to Chapter Three, Section 3.7). However, when C440S purified in the presence of TCEP was passed down the Superdex S200 column a different ratio of tetrameric protein and higher molecular weight protein (approximately a 4:1 ratio) was observed. This provides evidence that suggests that one or more cysteine residues may be responsible for the stabilization of quaternary structure of *Mt*-DAH7PS in solution. In the absence of reducing agent the thiol groups of cysteine residues (including the metal-binding cysteine) could oxidize to sulfinic ($-SO_2H$) and sulfonic acids ($R-SO_3H$), which may change their interactions with surrounding residues. However, DTT and TCEP are unable to reduce these species[186,187] and because the loss of activity is restored by either TCEP or DTT it suggests that a second disulfide linkage between two cysteines is present in *Mt*-DAH7PS. As mentioned earlier in this section the change in destabilization of quaternary structure could be due to disulfide formation between Cys231 residues, resulting in the formation of aggregated (higher molecular weight) protein. It is clear from these studies that further investigation into the role of cysteines in *Mt*-DAH7PS (especially Cys231) and other type II proteins is necessary to understand the requirement of reducing agent in maintaining enzyme activity and stabilization of quaternary structure.

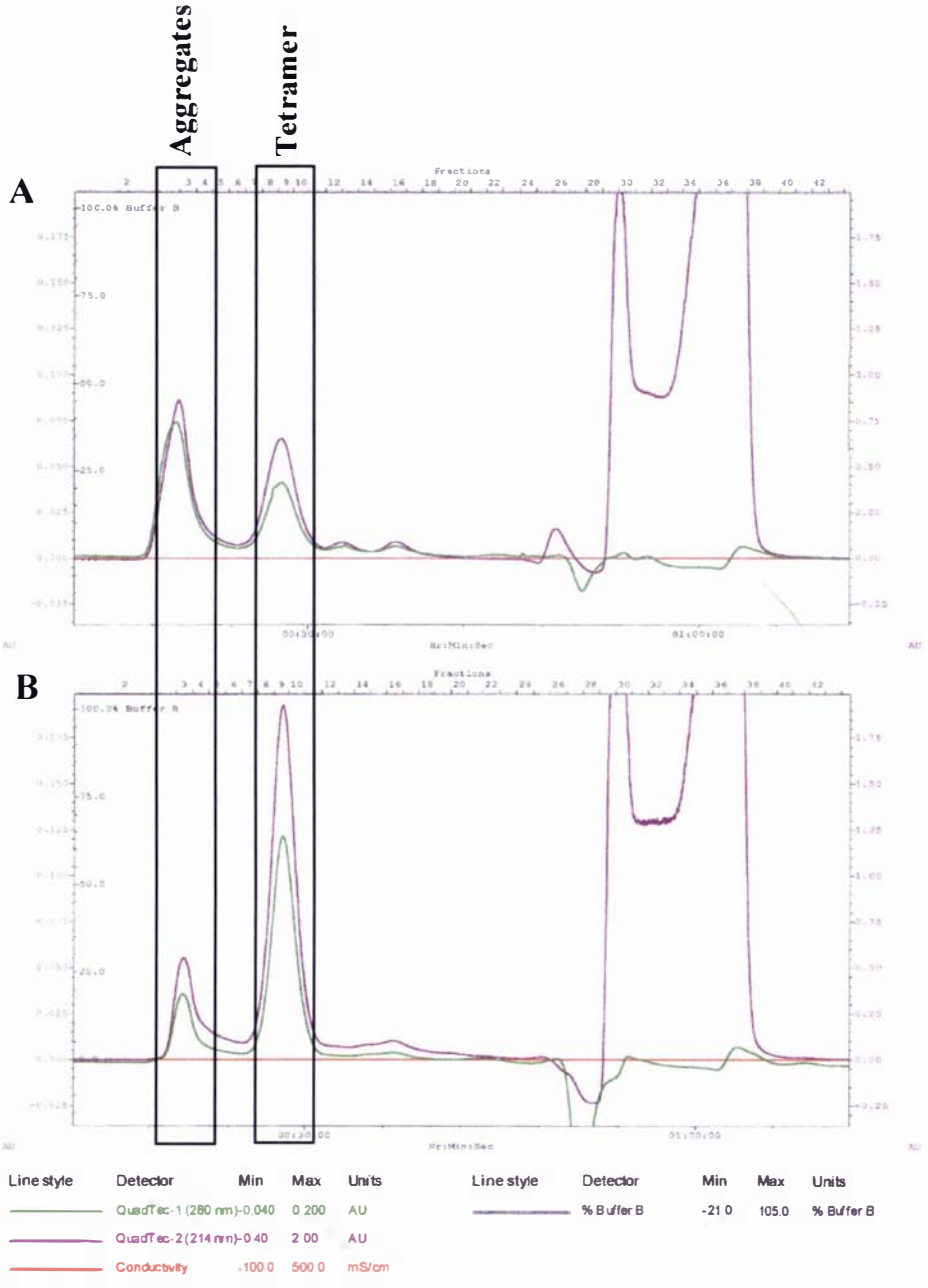


Figure 5.2 **Chromatogram trace from Superdex S200 run of C440S A, without TCEP and B, with TCEP**

5.4 Crystallization, Data Collection and Refinement of R284K and C440S

Both mutant proteins crystallized under similar conditions to that of the wild-type protein, Tris-HCl (0.1 M, pH 7.5), ammonium sulfate (1.5 M), glycerol (18 % (v/v)). Crystals were grown in a hanging-drop (1μL + 1μL) using R284K protein with a concentration of 4.1 mg/mL and a concentration of 2.9 mg/mL for C440S. Crystals of

R284K and C440S grew overnight, which is the same time frame as observed for native protein. The glycerol concentration was high enough to act as a cyro-agent so a cryoprotectant solution was not necessary. The crystals of the two mutant proteins had a different morphology to each other and wild-type protein. C440S crystals were very thin and needle-like with crystal dimensions of 0.5 x 0.02 x 0.02 mm. R284K crystals (0.5 x 0.07 x 0.07 mm) were approximately the same length but one and a half times the width of that seen for wild-type crystals.

Crystals of C440S that were grown did not diffract well and the highest resolution dataset that we were able to collect was to 2.8 Å. In contrast, the R284K crystals diffracted moderately well and a dataset was collected to a maximum resolution of 2.2 Å. Both data sets were collected in-house using a Rigaku MicroMax007 generator with Osmic blue optics and an RAxis IV++ detector. The R284K dataset was collected using non-overlapping 0.125° oscillations collected for ten minutes per frame at a crystal-detector separation of 180 mm. The dataset for C440S was collected using non-overlapping 0.2° oscillations for ten minutes per frame at the same crystal-detector distance as R284K.

The data were processed and scaled using CrystalClear 1.3.6 (Rigaku). Full details of the data collection statistics for R284K and C440S are in tables 5.1 and 5.2, respectively. Both the mutant crystals belong to the same space group as that of Se-Met protein ($P3_221$). The unit cell dimensions for both C440S and R284K were very similar to those of Se-Met *Mt*-DAH7PS. Based on this finding we were able to solve the structure by performing two rounds of rigid-body refinement calculating phase information from a water and ligand free Se-Met *Mt*-DAH7PS structure. The initial solution for both mutant proteins had an R and R_{free} value of ~30 %, and FOM of ~0.76. Optimization of the model consisted of repetitive cycles of rebuilding using COOT and refinement with REFMAC 5 (CCP4). Water molecules were added automatically in COOT and verified using 2Fo-Fc and Fo-Fc maps and by their potential to hydrogen bond to act least one protein atom or water molecule. In the R284K structure, two manganese ions, two PEP molecules and one sulfate ion (subunit B) (all identified by their shapes and environments) could be added into extra density. Subsequent refinement indicated that the occupancy of Mn^{2+} was 100 % and that of PEP and SO_4^{2-} was 75 %. Density for a molecule of detergent Thesit was also found bound at the interface between subunit A

and B with occupancy of 50 %. It was modeled in as a dodecyl ether moiety with one polyethyleneglycol unit attached. In the C440S structure two manganese ions (~75 % occupancy) and one sulfate ion (subunit B) were modeled in to both subunits. Two phosphate ions were modeled into both subunits (100 % occupancy) where PEP is found in the Se-Met and R284K structure. The 2Fo-Fc map does not show electron density for the remainder of the PEP molecule in the current C440S structure. Final refinement statistics for R284K and C440S are given in table 5.3 and 5.4, respectively.

Table 5.1 Data collection statistics for R284K

Values in parentheses are for the outermost shell	
Space group	$P3_221$
Unit cell dimensions (Å)	
$a=b$	204.44
c	66.37
Resolution range (outer shell) (Å)	31.6-2.2 (2.28-2.2)
Wavelength (nm)	1.542
No. measured reflections	337447
No. unique reflections	79940
Completeness (%)	99.1 (98.0)
Redundancy	4.2 (4.2)
$\langle I/\sigma \rangle$	9.5 (3.8)
R_{merge} (%)	10.0 (35.0)

Table 5.2 Data collection statistics for C440S

Values in parentheses are for the outermost shell	
Space group	$P3_221$
Unit cell dimensions (Å)	
$a=b$	203.1
c	66.4
Resolution range (outer shell) (Å)	39.3-2.8 (2.9-2.8)
Wavelength (nm)	1.542
No. measured reflections	123133
No. unique reflections	38616
Completeness (%)	99.3 (99.0)
Redundancy	3.2 (3.1)
$\langle I/\sigma \rangle$	4.0 (1.7)
R_{merge} (%)	18.6 (47.5)

Table 5.3 Refinement statistics for R284K

Resolution range (Å)	31.6-2.2
Number of reflections (test set)	75916 (4017)
R factor	0.213
R_{free}	0.247
Number of non-hydrogen atoms	
Protein (two molecules)	6938
Water	325
Mn^{2+}	2
PEP	40
SO_4^{2-}	5
Thesit	16
Mean *B* value (Å^2)	31.3
rms deviations from ideality	
Bond lengths (Å)	0.018
Bond angles (deg.)	1.9
Residues in the most favored region of Ramachandran plot (%)	89.8

Table 5.4 Refinement statistics for C440S

Resolution range (Å)	36.7-2.8
Number of reflections (test set)	33026 (1740)
R factor	0.236
R_{free}	0.294
Number of non-hydrogen atoms	
Protein (two molecules)	7045
Water	31
Mn^{2+}	2
PO_4^{3-}	10
SO_4^{2-}	5
Mean *B* value (Å^2)	43.7
rms deviations from ideality	
Bond lengths (Å)	0.014
Bond angles (deg.)	1.7
Residues in the most favored region of Ramachandran plot (%)	85.1

5.5 Data Collection, Refinement and Structure Determination of Native *Mt*-DAH7PS

The crystallization conditions for native *Mt*-DAH7PS are in Chapter Three, Section 3.12. A higher resolution dataset of the native protein were collected at a wavelength of 0.97929 Å on beamline 8.2 at the Advanced Light Source (Lawrence Berkeley National Laboratory, CA). The data were processed using SCALEPACK and DENZO by Minmin Yu (Berkeley, CA). Full details of the data collection statistics for native *Mt*-DAH7PS is in table 5.5. Reduction of the integrated intensities was performed using Scalepack2mtz (CCP4). The two datasets collected for the native and Se-Met substituted protein were found to be isomorphous. The native structure was kindly solved and refined by Heather Baker, Auckland University. The final refinement statistics for the native protein are in table 5.6.

Table 5.5 Data collection statistics for native *Mt*-DAH7PS

Values in parentheses are for the outermost shell	
Space group	$P3_221$
Unit cell dimensions (Å)	
$a=b$	201.83
c	65.96
Resolution range (outer shell) (Å)	43.7-2.0 (2.07-2.0)
Wavelength (nm)	0.97929
No. measured reflections	3460212
No. unique reflections	103542
Completeness (%)	99.3 (98.6)
Redundancy	5.7 (5.5)
$\langle I/\sigma \rangle$	20.4 (2.6)
R_{merge} (%)	9.6 (69.7)

Table 5.6 Refinement statistics for native *Mt*-DAH7PS

Resolution range (Å)	43.7-2.0
Number of reflections (test set)	97552(5140)
R factor (%)	18.9
R_{free} (%)	22.4
Number of non-hydrogen atoms	
Protein (two molecules)	7097
Water	531
Mn^{2+}	2
PEP	20
SO_4^{2-}	10
Thesit	16
Mean *B* value (Å^2)	31.27
rms deviations from ideality	
Bond lengths (Å)	0.016
Bond angles (deg.)	1.85
Residues in the most favored region of Ramachandran plot (%)	88.3

The asymmetric unit of native *Mt*-DAH7PS contains two molecules and in the current model, subunit A consists of the complete polypeptide chain (residues 1-462) and subunit B is equally complete. The main-chain torsion angles molecules correspond well with allowed values, with 88.3 % of residues in the most favored region of the Ramachandran plot, as defined by PROCHECK. In the current structure $Asp138_A$, $Asp199_B$, $Arg202_A$, $Cys420_A$, $Ser14_B$, $Asp263_B$ and $Phe379_B$ are in the disallowed region of the Ramachandran plot. The current model also contains two Mn^{2+} ions, two PEP molecules, two SO_4^{2-} ions, a molecule of the detergent Thesit and five hundred and thirty-one water molecules. The two SO_4^{2-} ions and a Thesit molecule (modeled in as a dodecyl ether moiety with one polyethyleneglycol unit attached) are modeled in at 100 % occupancy, and the two PEP molecules modeled in at 75 % occupancy.

The monomer-fold and tetrameric quaternary structure are similar to that observed for the Se-Met protein. The most intriguing feature of this crystal structure is the disulfide linkage between Cys440 and the metal-binding Cys87 (figure 5.3). The disulfide formation is observed in both subunits of the asymmetric unit and has been modeled in with an occupancy of 50 % in subunit A and 80 % in subunit B. A Mn^{2+} ion is modeled

into subunit A at ~50 % occupancy and ~20 % occupancy in subunit B. The metal occupancies are lower than that observed in the Se-Met *Mt*-DAH7PS crystal structure, which does not show disulfide formation between Cys440 and Cys87. There appears to be an association between disulfide formation and reduced metal ion occupancy. This allows us to suggest that the formation of a disulfide linkage between Cys440 and Cys87 results in the exclusion of metal from the enzyme active site.

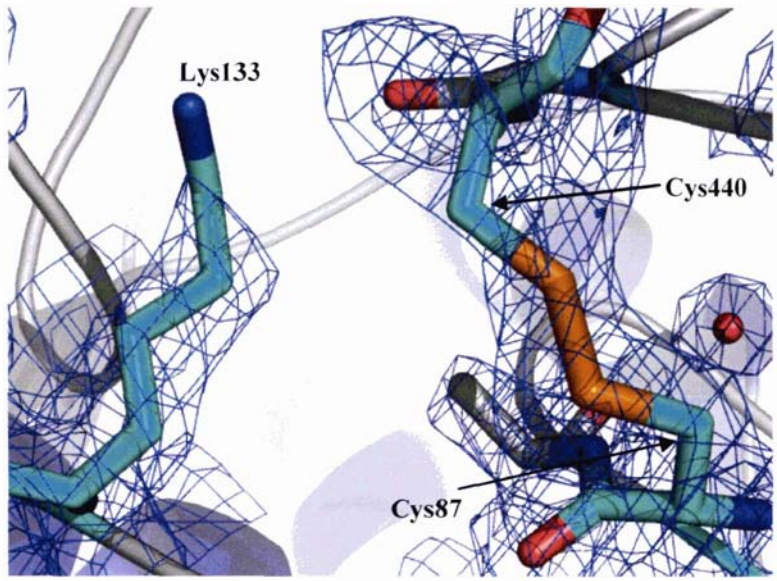


Figure 5.3 **Disulfide formation between Cys440 and Cys87 in native *Mt*-DAH7PS.** The residues here are from subunit B in the asymmetric unit. The 2Fo-Fc map (blue) is at a contour level of 1.2 σ.

5.6 Structure Determination of *Mt*-DAH7PS R284K

As was found for the Se-Met *Mt*-DAH7PS structure the asymmetric unit of the R284K crystal contains two molecules. In the current model, subunit A comprises the complete polypeptide chain (residues 1-462) and subunit B is similarly complete except for residues 10-13 and 374-379, which are disordered. The two residues (Gly-Ala) from the linker to the cleaved His_6-tag are also modeled for both subunits. The main-chain torsion angles molecules correspond well with allowed values, with 89.8 % of residues in the most favored region and Ser136A, Glu270A, Ser375A and Cys440A in the disallowed region of the Ramachandran plot, as defined by PROCHECK. The current model also contains two Mn^{2+} ions, two PEP molecules, one SO_4^{2-} ion, a molecule of the detergent Thesit and three hundred and twenty-five water molecules.

Monomer-fold and Quaternary Structure

The monomer-fold and tetrameric quaternary structure of R284K are essentially superimposable with that of Se-Met *Mt*-DAH7PS. Superposition of the two molecules in the asymmetric unit yields an rms difference of ~0.5 Å for mainchain atoms, using Lsqkab (CCP4).

Active Site

In the Se-Met *Mt*-DAH7PS structure Arg284 is 3.3 Å away from SO_4^{2-} and 3.7 Å away from the phosphate group of PEP (figure 5.4). The arginine interacts with the phosphate group of PEP via a water molecule (Arg284...(2.9 Å)...HOH...(2.6 Å)...PO_4^{3-}). The mutation of an arginine at residue 284 of *Mt*-DAH7PS to a lysine appears to have removed its interactions with both the phosphate moiety of PEP and the SO_4^{2-} group believed to represent the phosphate group of E4P. No intermediate water molecule is found in the R284K structure either. However, in one of the subunits in the asymmetric unit several residues of the β2-α2 loop that interact with E4P are partially disordered and there appears to be no SO_4^{2-} molecule present. Lys284 appears to be more disordered in this subunit as well. Mn^{2+} is modeled in both subunits with an occupancy of 100 %.

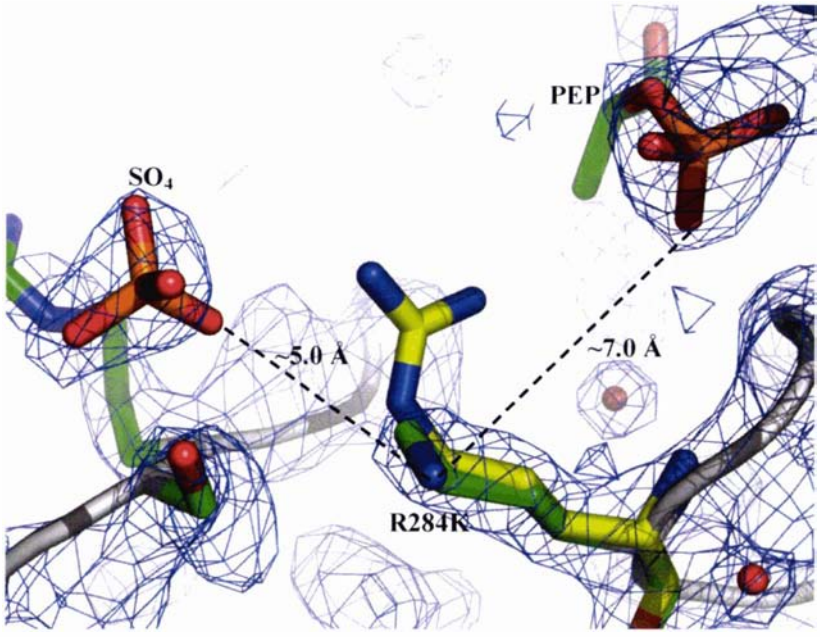


Figure 5.4 Substitution of an arginine for a lysine at residue 284 in *Mt*-DAH7PS. Arg284 in Se-Met *Mt*-DAH7PS (yellow) is superimposed onto Lys284 (green) in R284K. 2Fo-Fc map of R284K is in blue at a contour level of 1.2 σ. This figure is what is observed in subunit B, in subunit A Lys284 is more disordered.

5.7 The Role of Lys133 and Cys440 in Enzyme Catalysis

In the crystal structure of R284K a salt-bridge between Lys133 from the β2-α2 loop, with Cys440 from loop β8-α8 (figure 5.5) is more clearly defined than in the Se-Met structure, where this lysine residue appears to be more disordered (described further on in this section). Because of this, the R284K structure will be used to discuss the importance of Lys133 and its interaction with Cys440 in *Mt*-DAH7PS. Lys133 (*M. tuberculosis* numbering) is conserved across all DAH7PSs and it has been proposed that this lysine provides the proton necessary to convert the carbonyl oxygen atom of C1 of E4P to a hydroxyl group.[53] The equivalent lysine in *T. maritima* (Lys131) has been mutated and has been shown to be essential for enzyme catalysis.[51] A superposition of *Mt*-DAH7PS with the type Iα enzyme from *S. cerevisiae* and the type Iβ enzyme from *T. maritima* shows that the sidechain of Lys133 in the type II enzyme is in a different position to that observed for the type I DAH7PSs (figure 5.6). In the crystal structure of *Tm*-DAH7PS the equivalent lysine interacts with both PEP and E4P and this has been proposed for *Sc*-DAH7PS.[51,53] Interestingly, the equivalent lysine residue in *Pf*-DAH7PS (Lys60) is disordered in the crystal structure and has been built in as having two conformations.[47]

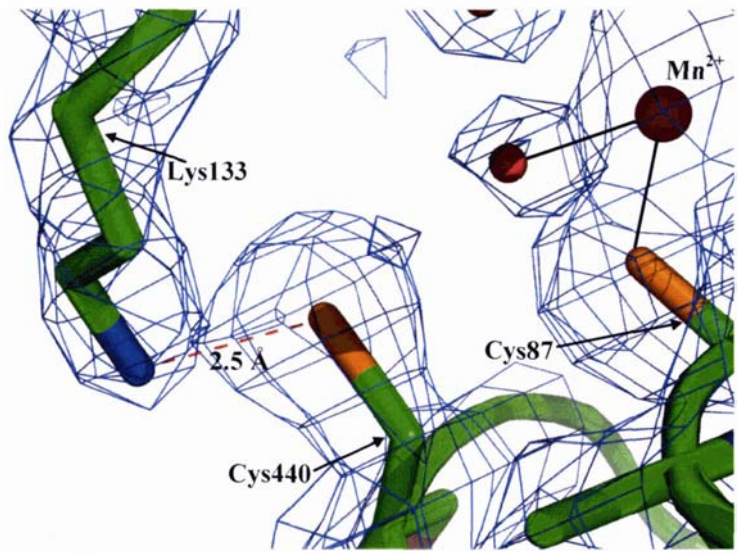


Figure 5.5 **Salt-bridge between Lys133 and Cys440 in *Mt*-R284K.** The 2Fo-Fc map is in blue at a contour level of 1.2 σ. Rotation of the Cys440 sidechain would allow disulfide formation with the metal-binding cysteine (Cys87).

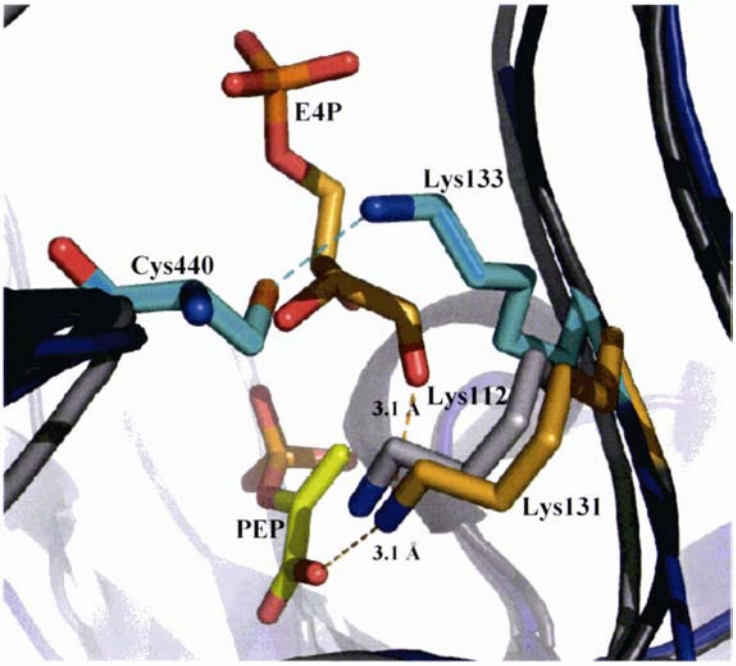


Figure 5.6 **Superposition of *Mt*-DAH7PS R284K (blue), *Tm*-DAH7PS (PDB code 1RZM, light grey) and *Sc*-DAH7PS(Tyr) (PDB code 1OFA, dark grey). Lys133 and Cys440 of *Mt*-DAH7PS R284K are in cyan and the equivalent lysine residues in *Tm*-DAH7PS (Lys131, orange) and *Sc*-DAH7PS(Tyr) (Lys112, light grey)**. The PEP and E4P molecules displayed are taken from how they are observed in the *Tm*-DAH7PS structure.

It appears that in the case of the *M. tuberculosis* enzyme, the salt-bridge between Lys133 and Cys440 may act as a "gate" to the active site, preventing E4P binding until the enzyme is ready for the second substrate (figure 5.7). This salt-bridge does not appear to be present in the type I enzymes and it could be that this interaction may only be found in type II enzymes exemplified by *Mt*-DAH7PS, which has been shown to display a novel and complex mode of feedback-regulation by the aromatic amino acids (as described in Chapter Four).

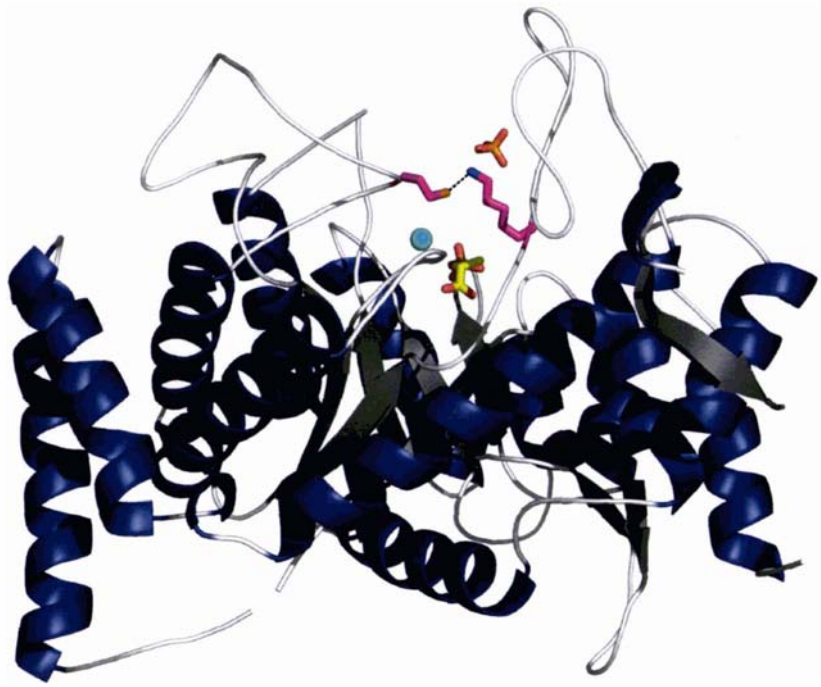


Figure 5.7 Salt-bridge between Lys133 (β2-α2 loop) and Cys440 (β8-α8 loop) acting as a gate to *Mt*-DAH7PS's active site. Lys133 and Cys440 are shown as sticks in pink. PEP (yellow) and SO_4^{2-} (orange) are shown as sticks and Mn^{2+} is shown as a sphere (cyan).

In the absence of reducing agent the interaction between Lys133 and Cys440 is disrupted due to disulfide formation between the thiol groups of Cys440 and metal-binding ligand, Cys87. In the native crystal structure this disulfide linkage is associated with the exclusion of metal (refer to Section 5.5), which explains the requirement of reducing agent to maintain full enzyme activity (refer to Section 5.3). This would suggest that Cys87 is essential for the binding of metal at the active site of *Mt*-DAH7PS. Site-directed mutagenesis studies of the metal-binding cysteine have been performed in *Pf*-DAH7PS (C31G)[188] and *Ec*-DAH7PS(Phe) (C61G)[189] and the resulting protein is able to bind metal and retains metal-dependent enzyme activity. However, when the metal-binding cysteine is mutated to either a serine, valine or alanine the mutant protein exhibits no detectable DAH7PS-activity and is unable to bind metal.[89] The loss of enzyme activity in *Mt*-DAH7PS in the absence of reducing agent is most definitely due to the formation of a disulfide linkage between Cy87 and Cys440. The involvement of the S^{γ} group of Cys87 in a disulfide bond with Cys440 results in the C^{β} atom facing towards the metal-binding site, which is equivalent to having an alanine residue at this

position. The reduction in metal occupancy and the loss of enzyme activity is consistent with results observed with the *E. coli* C61A mutant.

In the native *Mt*-DAH7PS crystal structure the formation of the disulfide appears to render the β8-α8 loop (residues 413-444) partially disordered, including Asp441, a metal-binding ligand. The comparison of the crystal structures of R284K (resolution of 2.2 Å and mean *B* value of 30.8 $Å^2$) and native *Mt*-DAH7PS (resolution of 2.0 Å and mean *B* value 31.3 $Å^2$) shows that Asp441 has a significantly higher *B* factor in the native structure (~70 $Å^2$ compared with ~40 $Å^2$ in the mutant structure) than in the mutant structure. Residues 429-441 have *B* factors of ~80-100 in the native structure and ~30-50 $Å^2$ in the R284K structure. Residues 425-428 are missing in the native structure and residues 414-424 have *B*-factors of ~70-80 $Å^2$ compared with ~30-50 $Å^2$ in the mutant structure. The modeling of E4P in *Tm*-DAH7PS indicates that the metal-binding aspartate residue (Asp309) interacts, using both of its O^{δ} groups, with the hydroxyl group of C3 of E4P.[51] The formation of a disulfide linkage between the two thiols groups of Cys87 and Cys440 in the native crystal structure appears to disorder Asp441, a residue involved in the binding of metal and proposed binding of E4P.

In the Se-Met structure (2.3 Å resolution and mean *B* value of 43.5 $Å^2$) Lys133 and Cys440 are partially disordered (*B* factors of ~60 and ~80 $Å^2$, respectively) in comparison with the R284K structure (*B* factors of ~30 and ~40 $Å^2$, respectively). A portion of the β8-α8 loop is presumed disordered in the Se-Met structure with residues 422-436 in subunit A and 418-431 in subunit B missing. We propose that the interaction between Lys133 and Cys440 is necessary for the stabilization of the β8-α8 loop, including the metal-binding ligand Asp441.

5.8 Structure Determination of *Mt*-DAH7PS C440S

As with the crystal structures discussed in this thesis, there are two molecules in the asymmetric unit. In the current model, molecule A comprises the complete polypeptide chain (residues 1-462), except residues 236-239 and molecule B is similarly complete except for residues 376-378 which are disordered. Two additional residues (Gly-Ala) from the linker to the cleaved His-tag are also modeled in for subunit B. The main-chain torsion angles molecules correspond well with allowed values, with 85.1 % of residues in the most favored region and $Ala431_A$, $Leu432_A$ and $Ser136_B$ in the disallowed region of the Ramachandran plot, as defined by PROCHECK. The final model also contains two Mn^{2+} ions, two PO_4^{3-} ions, one SO_4^{2-} ion, and thirty-one water molecules.

Monomer-fold and Quaternary Structure

The monomer-fold and tetrameric quaternary structure of C440S is essentially superimposable with the Se-Met *Mt*-DAH7PS structure; with the superposition of the two molecules in the asymmetric unit yielding an rms difference of ~0.5 Å for all mainchain atoms.

Active Site

Analysis of the current C440S structure shows differences between the subunits. In subunit B there is a clear interaction between Ser440 and Lys133 (figure 5.8) whereas in subunit A the two residues are further apart (~3.8 Å). Also, residues of the β2-α2 loop in subunit A that contribute to the binding of the phosphate group of E4P are disordered (Arg135 and Ser136 are modeled in as Ala residues and there appears to be no SO_4^{2-} ion present). In both subunits in the asymmetric unit all residues of the β8-α8 loop are present and Mn^{2+} is modeled in both subunits with a occupancy of 75 % (refer to previous section). Analysis of the current C440S structure (resolution of 2.8 Å) indicates no significant structural changes in the active site of *Mt*-DAH7PS upon the mutation of Cys440 to a serine. A higher resolution dataset of C440S may shed some light into how this conservative mutation can have such a significant influence on the enzyme's affinity for E4P (refer to Section 5.10). It will at least allow us to be able to compare *B*-factors of residues, as the current structure is at too low a resolution to be able to do this. At the moment there are two possible explanations for the dramatic increase in the binding of E4P to *Mt*-DAH7PS:

(1) It has been proposed that Lys133 (*M. tuberculosis* numbering) is required to provide the proton to C1 of E4P in the DAH7PS-catalyzed reaction (Section 5.7). The interaction between Lys133 and Cys440 may be essential for stabilizing Lys133 so that it is protonated and able to deliver a proton. A cysteine has a nominal pK_a of ~8 whereas a serine residue is less readily deprotonated with a pKa value of ~14. At physiological pH Lys133 will most likely be protonated as a lysine has a pK_a value of ~10. Given the closeness of its nominal pK_a value to the pH and the environment in which Cys440 is found it is possible that Cys440 is deprotonated. On the other hand a serine residue in its place (Ser440) will be almost certainly protonated. This change in ionization state of residue 440 in *M. tuberculosis* may alter the ability of Lys133 to donate its proton, and could explain the significant increase in the K_M for E4P. To investigate this further the pH profile of the C440S protein will need to be examined to see if this single amino acid mutation has an effect on the optimum pH of the enzyme.

(2) Alternatively, the conservative mutation may change the entropic properties of the enzyme, preventing the correct positioning of E4P for catalysis to occur. In subunit B of the C440S structure Lys133 appears to be disordered, as can be seen by the absence of density in the 2Fo-Fc map. It is quite possible that the salt-bridge between Lys133 and Cys440 is required to order the lysine residue proposed to protonate the carbonyl oxygen of E4P. As mentioned earlier a higher resolution dataset is required to investigate this theory further.

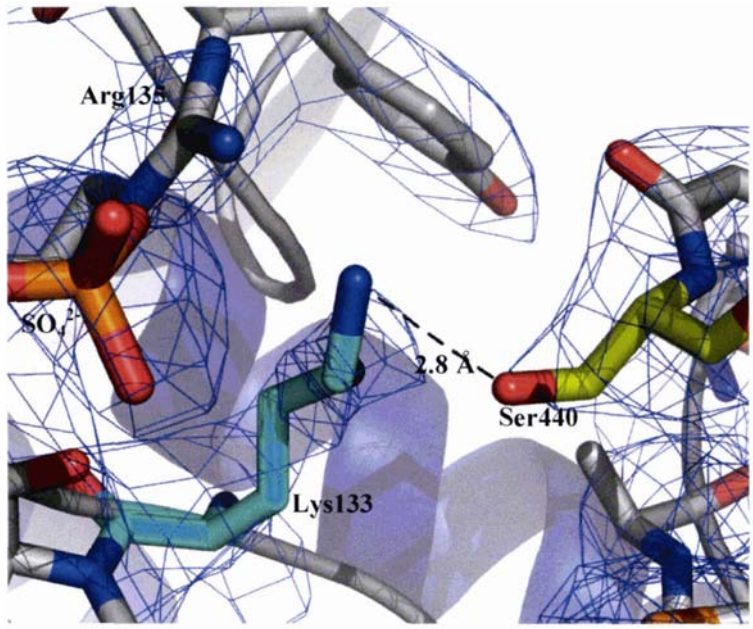


Figure 5.8 **Interaction between Lys133 (cyan) and Ser440 (yellow) in the C440S crystal structure.** This is in subunit B of the structure and the 2Fo-Fc map in blue is at a contour level of 1.1 σ. The equivalent residues in subunit A do not appear to interact with each other and the sidechain of Arg135 and Ser136, as well as a SO_4^{2-} ion are missing in this subunit.

5.9 Determination of Kinetic Parameters for *Mt*-DAH7PS R284K

The apparent K_M values for E4P and PEP were 67 ± 4 μM and 24 ± 1.7 μM, respectively and the k_{cat} value was calculated as 6.0 ± 0.1 s^{-1} (figure 5.9). It should be noted that E4P concentrations above 180 μM gave substrate inhibition of this mutant protein. These values are within the same order of magnitude as those observed for wild-type protein. This indicates that the R284K mutation has no significant effect on the binding of E4P and PEP, and the turnover of substrates by the enzyme.

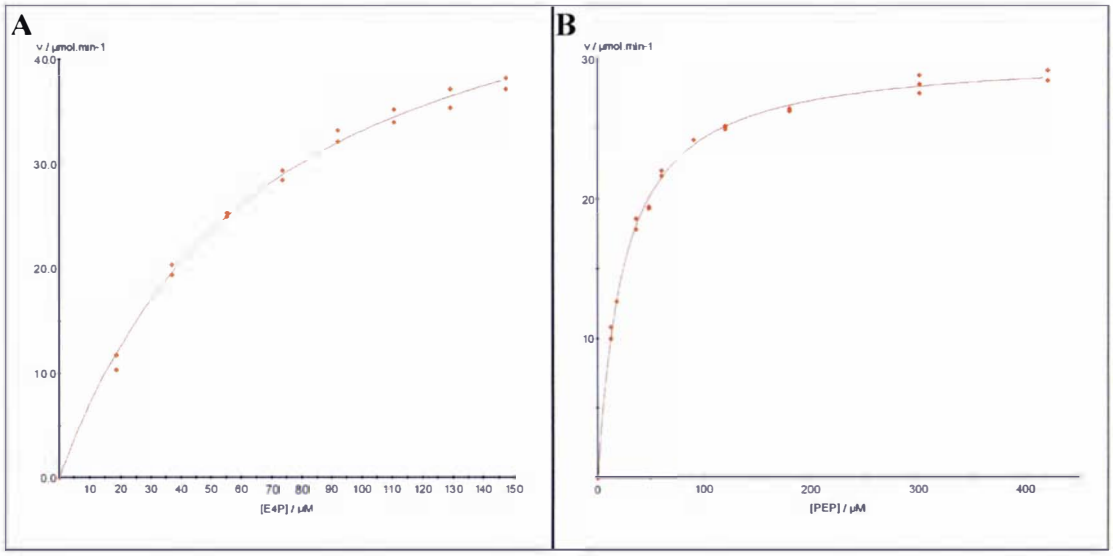


Figure 5.9 Michaelis-Menten plots for determination of K_M values for E4P A, and PEP B. The reaction mixtures for the determination of the K_M of E4P consisted of PEP (180 µM), $MnSO_4$ (100 µM) and E4P (18 to 147 µM), in BTP (50 mM) pH 7.5 buffer. The determination of the K_M of PEP consisted of reaction mixtures of E4P (147 µM), $MnSO_4$ (100 µM) and PEP (12 µM to 420 µM), in BTP (50 mM) pH 7.5 buffer. The reaction was initiated by the addition of purified *Mt*-DAH7PS (2 µL, 2 mg/mL) and was carried out at 30 °C. K_M and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using Enzfitter (Biosoft).

Both substrates are held in position in the active site via an extensive hydrogen-bonding network so the elimination of just one hydrogen bond may not be enough to alter catalysis significantly. Results show that mutating Arg284, which in the crystal structure of Se-Met *Mt*-DAH7PS appears to interact with both phosphate groups of PEP and E4P, to a lysine has little effect on the binding of PEP and E4P to the enzyme. Arg284 (*M. tuberculosis* numbering) is conserved across all DAH7PSs with the equivalent residue in *Ec*-DAH7PS(Phe) (Arg165) and *Pf*-DAH7PS (Arg115) interacting with the phosphate moiety of PEP.[47,71]

5.10 Determination of Kinetic Parameters of C440S

The apparent K_M values for E4P and PEP were 2050 ± 369 μM and 22 ± 3.0 μM, respectively and the k_{cat} value was calculated as 3.8 ± 0.05 s^{-1} (figure 5.10). The K_M of PEP is similar to that of wild-type protein, however the K_M of E4P is more than one hundred-fold higher than that observed for native *Mt*-DAH7PS.

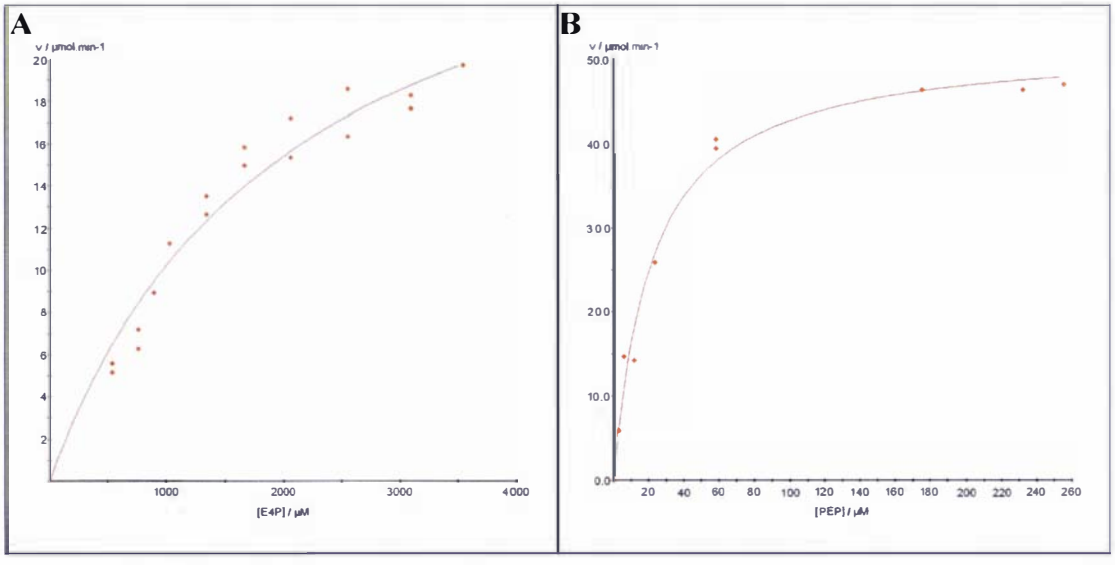


Figure 5.10 Michaelis-Menten plots for determination of K_M values for E4P A, and PEP B. The reaction mixtures for the determination of the K_M of E4P consisted of PEP (190 μM), $MnSO_4$ (100 μM) and E4P (540 μM to 3540 μM), in BTP (50 mM) pH 7.5 buffer. The determination of the K_M of PEP consisted of reaction mixtures of E4P (1780 μM), $MnSO_4$ (100 μM) and PEP (4 μM to 255 μM), in BTP (50 mM) pH 7.5 buffer. The reaction was initiated by the addition of purified *Mt*-DAH7PS (2.5 μL, 2.9 mg/mL) and was carried out at 30 °C. K_M and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using Enzfitter (Biosoft).

These values indicate that a single conservative mutation of a cysteine for a serine at residue 440 in *Mt*-DAH7PS has little effect on the binding of PEP but has a significant effect on the binding of E4P to the enzyme. It should be noted that the K_M of PEP was determined using E4P concentrations well below the K_M for E4P, which will tend to give a lower value for the K_M than the actual value. The crystal structure of C440S at the resolution we have does not appear to provide any insight into how this conservative mutation could have such a significant effect on the interaction of E4P with the protein. The k_{cat} value, or turnover of substrates does not appear to have changed significantly by this mutation.

5.11 Role of Cysteines in DAH7PS

Functional studies with both *Hp*-DAH7PS and *Mt*-DAH7PS indicate that cysteine residues play a crucial role in the regulation of enzyme activity of both these type II DAH7PSs, and a disulfide linkage between the thiol groups of Cys440 and Cys87 in *M. tuberculosis* has been captured in a crystal structure. Unlike *Ec*-DAH7PS(Phe) the *M. tuberculosis* enzyme does not appear to be stabilized upon the addition of EDTA and/or PEP alone and requires reducing agent for the maintenance of full enzyme activity (described in further detail in Chapter Three, Section 3.6).

Studies with *Ec*-DAH7PS(Phe) have shown that loss of enzyme activity is associated with the net loss of two thiols per subunit,[87] and peptide mapping revealed a disulfide linkage between Cys61 (equivalent to Cys87 in *Mt*-DAH7PS) and Cys328 (conserved in type Iα enzymes only). The loss of activity of the *E. coli* enzyme is proposed to be via metal-catalyzed oxidation of the two cysteine residues. Mutation of either of these two active-site cysteines resulted in enzyme that was insensitive to metal attack and loss of activity (refer to Chapter One, Section 1.5.1 for further detail). Cys328 was found to be nonessential, although conservative replacements at this position did have an effect on kinetic properties of the enzyme: Cys328 to Val showed a 20 % reduction in the catalytic constant and two to three-fold increases in $K_{M\text{-}PEP}$ and $K_{M\text{-}E4P}$.[89]

The comparison of the crystal structures of *Ec*-DAH7PS(Phe) and *Mt*-DAH7PS shows that the second cysteine (Cys440 in *Mt*-DAH7PS and Cys328 in *Ec*-DAH7PS(Phe)) involved in disulfide formation with the metal-binding cysteine (Cys87 in *Mt*-DAH7PS and Cys61 in *Ec*-DAH7PS(Phe)) do not occupy equivalent positions in the active site of DAH7PS (figure 5.11). Interestingly, both cysteine residues come from the β8-α8 loop of the $(\beta/\alpha)_8$ barrel. The space that Cys440 occupies in the *Mt*-DAH7PS structure is filled with a threonine or a serine in type Iα and type Iβ enzymes, respectively. Analysis of the two type Iβ crystal structures indicates that there is no second cysteine in the vicinity of the metal-binding cysteine residue. In both type Iα crystal structures Cys328 in *Ec*-DAH7PS(Phe) and Cys344 in *Sc*-DAH7PS(Tyr) do not obviously interact with any other residue. This is quite different to what is observed in *Mt*-DAH7PS where Cys440 forms a salt-bridge with Lys133 (from the β2-α2 loop). No interactions between residues of the β8-α8 and β2-α2 loops are observed in the Iα structures, and

residues 313-316 and 328-331 (part of the β8-α8 loop) are missing in the crystal structures of *Ec*-DAH7PS(Phe) and *Sc*-DAH7PS(Tyr), respectively.

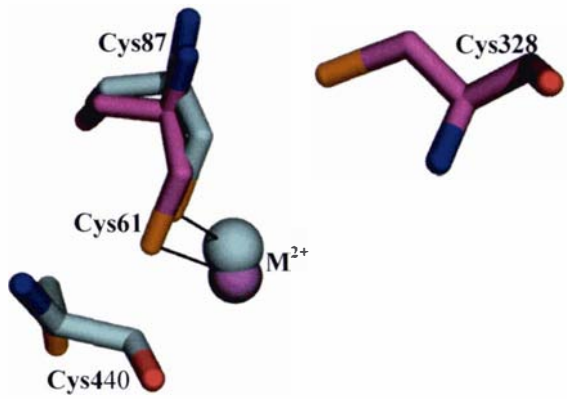


Figure 5.11 Superposition of Se-Met *Mt*-DAH7PS (cyan) and *Ec*-DAH7PS(Phe) (PDB code 1GG1) (pink) showing metal-binding Cys and second Cys involved in disulfide bond formation.

There appear to be differences in the role cysteine residues play in the regulation of enzyme activity between the three types of DAH7PSs: type Iβ enzymes have been reported to not require reducing agent for enzyme activity;[188] the type Iα enzyme from *E. coli* requires the presence of either EDTA or PEP to prevent disulfide formation between Cys61 and Cys328 and loss of enzyme activity; and the type II enzyme from *M. tuberculosis* absolutely requires reducing agent for the maintenance of full enzyme activity.

Comparison of the amino acid sequences of *Mt*-DAH7PS with type II enzymes from higher plants show that Cys440 (*M. tuberculosis* numbering) is also conserved in a number of plant DAH7PSs. Plant DAH7PSs have been shown to be hysteretically activated by DTT, and recently it has been reported that an *Arabidopsis* isoenzyme requires reduced thioredoxin or reducing agent for activity.[67] Investigation of the genes up and down stream of *Hp*-DAH7PS in the *H. pylori* genome indicates the presence of a peroxiredoxin (two genes upstream of DAH7PS) and a putative ferredoxin (four genes upstream of DAH7PS). In the case of the genome of *M. tuberculosis*, genes upstream and down stream of the DAH7PS gene are yet to be annotated with an actual or putative

function. Interestingly, three disulfide reductase activities have been identified in *H. pylori* whose reactions contribute to the redox balance of the cell.[190] Enzyme activity has also been identified that uses cystine (Cys-Cys) as a substrate. Therefore, it is conceivable that the activity of type II DAH7PSs from *M. tuberculosis* and *H. pylori* is regulated by the intracellular redox potential of their respective bacterium. Further investigations are required to confirm this.

5.12 Conclusions and Future Studies

The roles of several active-site residues of *Mt*-DAH7PS that attracted our attention were examined using site-directed mutagenesis, enzyme kinetics and X-ray crystallography. Investigation of the role of Arg284 led us to propose that this residue is not essential for the correct positioning of PEP and E4P for enzyme catalysis to occur. Instead we believe that Arg284 provides several of the numerous interactions between both substrates and the protein. On the other hand, Cys440 in *Mt*-DAH7PS appears to play an important role in the binding of E4P as can be seen by a significant increase in the K_M of E4P upon mutation of this residue to a serine. The requirement of C440S for a reducing agent suggests that cysteine residues play an important role in the stabilization or regulation of *Mt*-DAH7PS. Future studies are needed to investigate the role of cysteines in this type II enzyme. The importance of the salt-bridge formed between Lys133 and Cys440 on enzyme activity also needs to be looked into.

CHAPTER SIX SUMMARY OF THESIS

6.1 Overall Conclusions

The studies described in this thesis cover the functional characterization of two type II DAH7PSs from *H. pylori* and *M. tuberculosis*, and the structural characterization of the *M. tuberculosis* enzyme. The crystal structure of *Mt*-DAH7PS reported in this thesis is the first solved structure of a type II DAH7PS. The extensive characterization of both type II enzymes is the result of the successful solubilization of both enzymes by the co-expression with the *E. coli* molecular chaperones, GroEL and GroES.

The functional studies of *Hp*-DAH7PS described in Chapter Two of this thesis show that this type II enzyme catalyzes a metal-dependent, ordered sequential reaction with defined stereochemistry as seen for type I DAH7PSs. The results are consistent with a reaction mechanism where E4P binds to the PEP-DAH7PS complex and phosphate is released from the enzyme before DAH7P. Investigation of the substrate specificity of *Hp*-DAH7PS indicates that this type II enzyme is able to utilize, although relatively poorly, five carbon sugars, in which the C2 hydroxyl group is either absent or present in either possible configuration, as alternative substrates to E4P. This parallels observations made with the type I enzymes from *E. coli* and *P. furiosus*.[77,47] Substrate specificity studies have not been performed on any other microbial type II enzymes to date. Although the overall sequence identity between type I and type II enzymes is relatively low the comparison of a ClustalX alignment of ninety type II DAH7PSs with sequence and structural information available for *Ec*-DAH7PS shows that key residues implicated in metal, PEP and E4P binding are conserved between the two enzyme families. These residues are found in the same order and relative spacing in the primary sequence in both enzyme types, and are found in regions of relatively high sequence similarity, or conservation. These results suggest that catalysis by type I and type II enzymes occurs on a similar active site scaffold and that the two DAH7PS families may indeed be distantly related.

To identify the evolutionary relationship between the two DAH7PS families structural information for type II enzymes is required. As part of the studies in this thesis the crystal structure of *Mt*-DAH7PS was determined. Analysis of the crystal structure

revealed significant similarities and differences between the two types of DAH7PS, as detailed in Chapter Three. The comparison of the *Mt*-DAH7PS structure with the crystal structures of the type I enzymes reveal a remarkable similarity in fold and function between both families of enzyme. Despite their minimal sequence identity the key residues that interact with PEP and the divalent metal ion are completely conserved and positioned almost identically in the two DAH7PS types. The common active site architecture and chemistry, housed within a shared protein fold, suggest very strongly that the type I and II enzymes have arisen from a common DAH7PS ancestor. Interestingly, while the core $(\beta/\alpha)_8$ barrel is shared between both enzymes the quaternary structure and extensions to the basic barrel are significantly different. While the four type I enzymes (*Ec*-DAH7PS, *Sc*-DAH7PS, *Tm*-DAH7PS and *Pf*-DAH7PS) and *Mt*-DAH7PS crystallize as homotetramers, the association of the tetramers use completely different structural elements. While the type Iα and Iβ enzymes share a common dimer, the monomer-monomer interface of *Mt*-DAH7PS shows no commonality with any subunit interface in any previously reported DAH7PS structure.

The second important difference between the DAH7PS enzymes is the additional structural elements that decorate the core $(\beta/\alpha)_8$ barrel. The studies described in Chapter Four provide further evidence that suggests that these extra structural motifs are involved in allosteric regulation by the aromatic amino acids. The type Iβ *Pf*-DAH7PS is the most stripped-down DAH7PS, its monomer comprising just the core $(\beta/\alpha)_8$ barrel, is not regulated allosterically as it lacks these extra structural motifs.[47] The type Iβ *Tm*-DAH7PS, which is very similar in the core $(\beta/\alpha)_8$ barrel structure to *Pf*-DAH7PS, has a ferredoxin-like domain attached to the N-terminus of the barrel which has been implicated in Phe and Tyr inhibition.[51,95] The crystal structures of type Iα *Ec*-DAH7PS(Phe) and *Sc*-DAH7PS(Tyr), in complex with Phe and Tyr, respectively, reveal that both amino acids binding sites involve extensions of the N-terminus and extended α5-β6 loop.[49,52-54] In these studies the crystal structure of *Mt*-DAH7PS was determined in complex with Trp and Phe which revealed that the two extra structural elements decorating the barrel, an extended β-strand followed by three helices (α0a, α0b and α0c) at the N-terminus, and a pair of helices (α2a and α2b) that extend the α2-β3 connecting loop are involved in the binding of Phe and Trp, respectively. This is consistent with the synergistic inhibition displayed in the presence of Trp and either Phe or Tyr. While the nature of the communication between both aromatic amino acid

binding sites and the transmission of inhibition back to the active site still remains unclear; functional studies of *Mt*-DAH7PS indicate that the presence of Trp and Phe influences the processing of E4P by *Mt*-DAH7PS. Interestingly, whereas the tetrameric *Mt*-DAH7PS is feedback-regulated by all three aromatic amino acids *Hp*-DAH7PS, which exists as a dimer in solution, is unregulated by Trp, Phe and Tyr. Future structural studies of *Hp*-DAH7PS are essential to shed some light into the differences in regulation of the *M. tuberculosis* and *H. pylori* enzymes, as at the primary sequence level the two proteins appear to be similar.

An intriguing observation of both the *H. pylori* and the *M. tuberculosis* enzymes is the absolute requirement for a reducing agent. Without reducing agent both proteins rapidly lost activity over time. This loss in enzyme activity could then be restored upon the addition of reducing agent. The crystal structure of *Mt*-DAH7PS indicated the presence of a second cysteine residue (Cys440) close to the metal-binding cysteine (Cys87). A simple rotation about Cys440 C^{α}-C^{β} to another rotamer would allow formation of a disulfide bond with Cys87, precluding metal binding and explaining the need for a reducing agent to maintain full activity of the enzyme. To investigate the importance of this second cysteine (Cys440) on enzyme activity this residue was mutated to a serine. The conservative mutation appeared to have little effect on the binding of PEP but increased the K_M of E4P by approximately 100-fold. A crystal structure of the C440S *Mt*-DAH7PS mutant, although at poor resolution, revealed no significant changes in the active site to explain this dramatic change in the binding affinity of *Mt*-DAH7PS for E4P. The most plausible explanation is that the mutation of Cys440 to a serine one has altered the interaction between this residue and Lys133, an essential residue speculated to provide a proton for the reaction. Further studies are required to understand the role Cys440 plays in catalysis.

A second active-site residue that attracted our attention was Arg284 which in the Se-Met *Mt*-DAH7PS crystal structure interacts with both phosphate groups of PEP and E4P. This residue is also absolutely conserved across all types of DAH7PS. The mutation of this arginine to a lysine appeared to have little effect on the K_M of PEP and E4P, and k_{cat}. The crystal structure of the *Mt*-DAH7PS mutant also revealed no significant changes in the active site which suggests that this residue is not essential for the correct positioning

of PEP and E4P for enzyme catalysis to occur. Instead it is speculated that Arg284 provides several of the numerous interactions between both substrates and the protein.

6.2 Important Notes about Data Reported in this Thesis

It should be noted that there are several deviations between data presented in the article published in the Journal of Molecular Biology[181] and in this thesis. The first deviation is that *Mt*-DAH7PS is reported in the paper to exist as a dimer in solution and crystallographically. After more extensive investigations this was found to be incorrect and *Mt*-DAH7PS is in fact tetrameric in solution and in the crystal-form. The second discrepancy is that in the published article Trp, Tyr and chorismate are reported to inhibit *Mt*-DAH7PS activity while Phe is stated to have no effect on enzyme activity. However, further inhibition studies with the aromatic amino acids and chorismate indicated that both Phe and Tyr activate *Mt*-DAH7PS; whereas Trp and chorismate have no notable effect on *Mt*-DAH7PS activity. The most plausible explanation for the differences between the two sets of data is that the experiments in the paper were performed on enzyme that had been pre-treated with EDTA for twenty-four hours prior to inhibition assays with the aromatic amino acids. Studies in this thesis show that incubation with EDTA for twenty-four hours induces protein to convert from its tetrameric form to an aggregated state. This can then be reversed by overnight incubation with PEP and $MnSO_4$. It is possible that reaction rates performed on EDTA-treated protein are affected by the equilibrium between the two states of protein in the assay system. In order to clarify this, one would need to compare *Mt*-DAH7PS treated and untreated with EDTA keeping all other assay variables constant.

6.3 Possible Future Experiments

- To determine the crystal structure of *Hp*-DAH7PS which may help explain why this type II enzyme, which appears to be very similar to *Mt*-DAH7PS at the primary sequence level, is dimeric in solution and unregulated by the aromatic amino acids and chorismate. Structural information may allow one to identify putative small molecule/s responsible for regulation of this enzyme.

- The disruption of the *Mt*-DAH7PS tetramer interface would be interesting to see if tetramer formation was associated with feedback-inhibition by the aromatic amino acids. A possible mutation would be to mutate Phe227, which contributes 14 % of the interface buried upon tetramer formation, to a charged residue, for example, an aspartate.

- A crystal structure of *Mt*-DAH7PS without the presence of Thesit is necessary to eliminate any movements in the structure induced by the binding of this non-ionic detergent molecule. This may allow one to extract more information out of the crystal structures with and without Trp and Phe bound, in order to identify the transmission of inhibition from Trp and Phe binding, to altered processing of E4P at the enzyme's active-site. Determination of the binding constants of Trp and Phe are also required to help explain this novel mode of allosteric inhibition. This could be achieved by labeling Trp and Phe and using NMR spectroscopy to look at the effect on the signal in the presence of *Mt*-DAH7PS. A series of experiments at different concentrations of labeled amino acids may allow a binding constant to be determined.

- The finding that the mutation of Cys440 to a serine resulted in a dramatic increase in the K_M of E4P indicates that this residue plays an important role in enzyme catalysis. To investigate whether the mutation of a cysteine to a serine results in an altered relative pK_a of Lys133 a titration curve of pH against enzyme activity could be performed.

CHAPTER SEVEN EXPERIMENTAL PROCEDURES

7.1 GENERAL METHODS

Agarose Gel Electrophoresis

DNA fragments were separated on the basis of size using agarose gel electrophoresis. Agarose gels (0.8 % (w/v)) were prepared by adding 0.4 g agarose to 50 mL 1X TAE buffer (40 mM Tris.HCl, 20 mM acetic acid and 2 mM EDTA at pH 8.0) and heating until dissolved. All gels were run using a Sub-Cell® GT Agarose Gel electrophoresis system (Bio-Rad) in 1X TAE buffer. DNA samples were premixed with 6X loading buffer ((0.2 % (w/v) bromophenol blue in 50 % (v/v) glycerol) and then loaded into the wells. Electrophoresis was performed at 80 V until the dye front had migrated to the other end of the gel. The DNA on the gels was visualized by staining the gel in ethidium bromide (~0.5 μg/mL), followed by exposure to ultraviolet light (302 nm). Pictures of the gels were taken using an Alpha Imager gel documentation system (Alpha Innotech Corporation, USA).

Quantification and Size Determination of DNA Fragments

The concentration of DNA samples was estimated by comparison with DNA samples of known concentration run alongside on an agarose gel. The approximate size of a DNA band was estimated by comparing its migration through the gel against that of DNA standards with known size (1 Kb Plus DNA ladder, Invitrogen) which were run alongside the DNA sample of interest.

DNA Sequencing

DNA sequencing services were provided by the Massey University Allan Wilson Center for Molecular Evolution and Ecology Genome Service. DNA sequencing was carried out on either an ABI Prism 377-64 sequencer or an ABI Prism 3730 capillary sequencer, using BIGDYE labeled dideoxy chain termination chemistries (Applied Biosystems).

Transformation of pGroESL into Competent *E. coli* BL21(DE3) and DL41(DE3) Cells

Transformation into competent BL21(DE3) (kindly donated by Trevor Loo) and DL41(DE3) cells (donated by Dr Shaun Lott) was performed as described by Inoue *et al.*[191] This consisted of thawing 50 μL aliquots of cells on ice and incubating for 30 minutes with pGroESL (5 μL of ~10 ng/μL) (kindly donated by Dr Mark Patchett) on ice. The cells were heat-shocked at 42 °C for 60 seconds and returned to ice for a further 10 minutes. The cells were plated onto pre-warmed (37 °C) LB agar plates containing chloramphenicol (25 μg/mL) and incubated at 37 °C for ~16 hours.

Preparation of Electro-Competent *E. coli* BL21(DE3)-pGroESL and DL41(DE3)-pGroESL Cells

Electro-competent *E. coli* BL21(DE3)-pGroESL cells were prepared using the protocol described by Dower *et al.* [192] 3 mL of LB (25 μg/ mL chloramphenicol) was inoculated with a scraping from a glycerol stock of BL21(DE3)-pGroESL and left to grow overnight. This was then used to inoculate 100 mL LB (25 μg/ mL chloramphenicol). Cells were grown until the cell density reached an OD_{600} of ~0.5 and the culture was then centrifuged (6400*g*) for 5 minutes at 4 °C. The supernatant was discarded and the pellet re-suspended in 100 mL cold 1 mM HEPES pH 7.0 buffer (0.2 μM filtered). The re-suspended cells were centrifuged again, supernatant discarded and pellet re-suspended in 50 mL HEPES buffer. The re-suspended cells were then centrifuged again, pellets re-supended in 4 mL cold sterile glycerol (10 % (v/v)). The cells were centrifuged (7700*g*) for the final time for 10 minutes at 4 °C and re-suspended in 500 μL glycerol (10 % (v/v)). They were then pipetted out into 50 μL aliquots on ice, snap-frozen in liquid nitrogen and stored at -80 °C.

Transformation of Competent *E. coli* BL21(DE3)-pGroESL and DL41(DE3)-pGroESL Cells

The transformation of the DAH7PS plasmid into electro-competent BL21(DE3)-pGroESL cells was performed by electroporation following the protocol described by Dower *et al.* 50 μL of cells were thawed on ice and added to a electroporation cuvette (Bio-Rad, 0.2 cm gap) along with 2 μL of plasmid preparation (10-50 ng/μL). The cuvette was flicked to mix and incubated on ice for 2 minutes. The cuvette was removed from the ice, sides wiped down, and the cells were electroporated using a Bio-Rad Gene Pulser (200 Ω, 25 μF, 2.5 kV) with a time constant of ~4.6 msec. 500 μL of SOC was immediately added to the cuvette, mixed and the cells transferred to a sterile 1.5 mL Eppendorf tube and incubated at 37 °C for 1 hour. The cells were plated onto pre-warmed LB agar plates (100 μg/mL Amp and 25 μg/mL chloramphenicol) undiluted and 5-fold diluted with SOC media (e.g 10 μL cells added to 90 μL SOC).

Media

All *E. coli* cultures grown up in this project were grown in Luria Broth (LB) base unless otherwise stated. LB broth (Invitrogen) was made up (25 g/L) with Milli-Q water and sterilized by autoclaving at 121 °C and 15 psi for 20 minutes. Agar media for agar plates was prepared by adding 1.5 % agar (Oxoid) to the liquid media prior to autoclaving.

SOC media was used to incubate freshly transformed *E. coli* cells after electroporation. SOC media consists of 20 g/L bacto-tryptone (Merck), 5 g/L bacto-yeast extract (Invitrogen), 0.5 g/L NaCl (Univar), 2.5 mM KCl (Merck), 10 mM $MgCl_2$ (BDH) and 10 mM glucose (Invitrogen). All components were added to Milli-Q water, except glucose, and autoclaved. Solutions of glucose (100 mM) and $MgCl_2$ (1 M) were filter-sterilized (0.2 μM) and added immediately prior to use.

Antibiotic Stocks

Stock solutions of ampicillin (100 mg/mL) in Milli-Q water were filter sterilized (0.2 μM) and stored at -20 °C. Stock solutions of chloramphenicol (50 mg/mL) were made up in ~95 % ethanol and stored at -20 °C.

Storage of Cultures

All strains of *E. coli* were stored at -80 °C as frozen glycerol stocks. An aliquot of each culture was taken when the optical density at 600 nm (OD_{600}) reached ~0.6. Glycerol was added to the culture to give a final concentration of 10 % (v/v) and the aliquot was snap-frozen in liquid nitrogen and stored at -80 °C.

Centrifugation

Centrifugation in this project was performed in one of three centrifuges: a SORVALL Evolution RC centrifuge, a SORVALL Heraeus multifuge® 1S/1S-R or a MiniSpin® centrifuge (Eppendorf).

Sonication

Sonication was performed using a VirTis VirSonic digital 475 ultrasonic cell disrupter using an 1/8 inch probe at ~60 Watts.

pH Measurement

The pH of buffers used in this project was measured using a Model 20 pH/Conductivity Meter (Denver Instrument Company) and a Sartorius Professional Meter. The pH of the solution was adjusted using 10 M NaOH and ~12 M HCl.

Growth of *E. coli* Cells

All 1 L cell cultures were grown from an overnight 25 mL culture of LB (Amp/Chloramphenicol) which had been inoculated with a scraping of a glycerol stock of the appropriate strain of *E. coli* cells. *E. coli* cells containing pGroESL and the DAH7PS-plasmid construct were grown in LB containing ampicillin (100 μg/mL) and chloramphenicol (25 μg/mL) in baffled flasks (1 L media per 5 L flask) at 37 °C with shaking (150 rpm) until mid-logarithmic phase (OD_{600} ~ 0.4-0.6). The growth temperature was then lowered to 25 °C before induction.

Induction of Protein Expression

Expression of genes inserted into the multiple cloning site of pET-32a(+) and pProEX-HTa are under the control of the *lac* promoter and is therefore induced by the addition of lactose or isopropyl-β-D-thiogalactopyranoside (IPTG) (Applichem). IPTG is a non-physiological analogue of lactose which activates gene expression but is not metabolized by the cell like lactose. IPTG was added to cultures to a final concentration of 1 mM (unless otherwise stated) to induce protein expression.

Harvesting and Lysis of Cells

The cells were harvested approximately 6 hours after induction by centrifuging at 6400*g* for 20 minutes at 4 °C. The cell pellets were then stored at -80 °C until lysis.

The cell pellets were re-suspended in the appropriate lysis buffer (on ice). Volumes smaller than 5 mL were lysed by sonication on ice and larger volumes were lysed using a French Press (8000 psi) or by passage through a cell disrupter (Constant Cell Disrupter Systems) at 97 MPa. The DNA was broken up by sonication on ice and the cell debris removed by centrifugation (27,000*g* at 4 °C for 15 minutes).

The lysis buffer used for *Hp*-DAH7PS consisted of 10 mM BTP pH 7.5, 2 mM DTT, 0.005 % (v/v) Thesit and 200 μM PEP. The lysis buffer used for *Mt*-DAH7PS consisted of 20 mM BTP pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.005 % (v/v) Thesit, 200 μM PEP and 100 μM $MnSO_4$. The BTP buffer was made up and treated with Chelex (1-2 hrs at room temperature) prior to addition of additives.

Polyacrylamide Gel Electrophoresis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed using the method of Laemmli[193] with a 4 % (w/v) stacking gel and a 12 % (w/v) separating gel, using a Mini Protean III cell (Bio-Rad). All samples were prepared in SDS loading buffer without boiling except for whole cell samples, which were boiled for one minute. Low range SDS-PAGE molecular weight standards (Bio-Rad) were used. A constant voltage of 260 V was applied across the electrodes until the dye front reached the bottom of the separating gel. After electrophoresis, gels were

stained for protein using Coomassie Blue. The staining solution consisted of 1 g/L Coomassie Brilliant Blue R 250 (Park Scientific) in 50 % (v/v) methanol, 10 % acetic acid (v/v) in water for approximately 20 minutes. To remove excess dye the gel was then destained in an identical solution without the dye.

Storage of Enzymes

All enzyme preparations were stored as aliquots of volumes no greater than 200 μL, snap-frozen in liquid nitrogen and stored at -80 °C.

Determination of Protein Concentration

Protein concentrations were determined by the method of Bradford,[194] using bovine serum albumin as a standard. Bradford reagent (Bio-Rad) was diluted 5-fold with Milli-Q water and filtered (0.2 μM). The BSA protein standards were made from dilution of a stock (Bio-Rad, 20 mg/mL). The assays were performed by thoroughly mixing 100 μL of protein standard solution or sample with 1 mL of Bradford reagent. Absorbance readings (595 nm) were then taken of all the standards and samples and the concentration (mg/mL) was determined using Cary UV software.

Purified Water

Water was purified by passage through a Sybron/Barnstead NANOpure II filtration system (Maryland, USA), containing two ion-exchange and two organic filters. This water is referred to as Milli-Q water throughout this thesis. Double distilled water was generated by filtering Milli-Q water through a layer of charcoal powder and then distilling over charcoal. All Milli-Q water used in the cloning of genes was autoclaved prior to use. Chelexed water in this thesis refers to Milli-Q water which has been treated with Chelex (Bio-Rad) by stirring for 1-2 hours before it is removed by filtration (0.2 μM).

Fast Protein Liquid Chromatography (FPLC)

FPLC was carried out using a Bio-Rad Biologic protein chromatography system at 4 °C. All buffers and solvents for FPLC were filtered using a 0.2 μM filter (Millipore).

Immobilized Metal Affinity Chromatography

Immobilized Metal Affinity Chromatography (IMAC) is the most commonly used procedure in purifying recombinant protein with either a C- or an N- terminal poly-His tag. It allows a one-step purification of the protein of interest to be achieved. In this project the metal-chelating resin used consisted of nitrilo triacetic acid (NTA) (a metal-chelator) immobilized onto a polysaccharide matrix through a spacer arm. The resin was charged with Ni^{2+} ions which allowed the retention of recombinant protein through coordination of the poly-His tag to the metal ions. Proteins were eluted from the column using an imidazole gradient which competes with the His residues for a coordination site of Ni^{2+}. Proteins with few surface histidines are eluted with low concentrations of imidazole, while proteins with a poly-His tag are retained until high concentrations of imidazole are reached.

In this study ~6 mL of Chelating Sepharose resin (Amersham Biosciences) was packed into a low pressure 1 cm x 10 cm Econo-column (Bio-Rad). The resin was charged with Ni^{2+} ions using a $NiCl_2$ solution (~0.1 M) and the excess unbound metal ions were removed from the column by extensive washing with Milli-Q water. The loading buffer (equivalent to lysis buffer) consisted of 20 mM BTP pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.005 % v/v Thesit, 200 μM PEP, 100 μM $MnSO_4$. The column was equilibrated with loading buffer, then crude lysate that had been filtered through a 0.8 μm filter was loaded onto the column at a flow rate of 0.4-0.5 mL/min. The column was washed with 5 column volumes of loading buffer before bound proteins were eluted with a linear gradient of 0-500 mM imidazole (over ~20 column volumes (130 mL)) in loading buffer. The purification was performed at 4 °C and the buffers were applied to the column using an Econo peristaltic pumping system (Bio-Rad) and gradient mixer. The flow rate was set at 1 mL/min, and 2 mL fractions were collected using a Econo fraction collector. Fractions were analyzed for DAH7PS activity and by SDS-PAGE. The resin was regenerated by stripping Ni^{2+} ions with 0.5 M EDTA, pH 8.0 then washed extensively with Milli-Q water.

Cleavage of N-terminal His_6 Tag by rTEV

The fractions exhibiting DAH7PS activity were pooled and washed with loading buffer to dilute the concentration of imidazole to ~200 mM. This involved diluting with loading buffer and then concentrating using a spin concentrator (10 kDa molecular-weight cutoff (MWCO) filter (Vivascience)) repeatedly. The protein was then incubated for up to 15 hours at 22 °C with recombinant tobacco etch virus (rTEV) protease (0.1 mg per litre of original culture) (kindly donated by Trevor Loo). The completeness of cleavage was monitored by SDS-PAGE. The enzyme preparation was then concentrated to no more than 5 mg/mL using a spin concentrator (10 kDa MWCO filter). The rTEV was successfully removed in the final step of the purification by size exclusion chromatography.

Size Exclusion Chromatography (SEC)

SEC or gel filtration chromatography separates proteins on the basis of molecular mass. The matrix is porous acting like a molecular sieve that excludes large proteins which are eluted in the void volume and retains small proteins that are eluted from the column between the void volume and total column volume in order of decreasing molecular mass.

SEC using a Superdex S200 HR 10/300 column (Amersham Biosciences) was used as a final purification step in the purification of all the proteins discussed in this thesis. This step removed rTEV protease from *Mt*-DAH7PS proteins cleaved from the His_6 tag. This step was also used to exchange *Hp*-DAH7PS and *Mt*-DAH7PS proteins into more suitable buffers for crystallization trials. ~400-450 μL of protein was injected onto the column with an isocratic flow of 0.4 mL/mL for 40 mL. The Gel filtration buffer used for the purification of *Hp*-DAH7PS consisted of BTP (10 mM, pH 7.0) supplemented with PEP (200 μM), $MnSO_4$ (100 μM) and DTT (1 mM). The purification of the *Mt*-DAH7PS proteins used a similar buffer but with the addition of 150 mM NaCl and the substitution of TCEP for DTT. All buffers were made up using charcoal distilled Chelexed Milli-Q water. Fractions that gave a UV absorbance (280 and 214 nm) were analyzed by SDS-PAGE.

Determination of Molecular Mass using SEC

Size exclusion chromatography (Superdex S200, Amersham Biosciences) was used to determine the molecular mass of *Hp*-DAH7PS and *Mt*-DAH7PS in solution by comparing elution times with those for molecular mass standards (MW-GF-200, Sigma). The molecular mass standards were cytochrome C (12.5 kDa), carbonic anhydrase (29 kDa), bovine serum albumin (66 kDa), alcohol dehydrogenase (150 kDa) and β-amylase (200 kDa). The gel filtration buffers used are described in the previous section. The same buffer was used to elute both DAH7PS and the standards. All SEC was performed at a flow rate of 0.4 mL/min over 40 minutes.

For the determination of the molecular mass of *Hp*-DAH7PS ~2 mg of bovine serum album, ~1 mg of β-amylase and ~1 mg of cytochrome C was dissolved in 300 μL gel filtration buffer, centrifuged at max-speed on a Eppendorf bench-top centrifuge for 2 minutes to remove any small solid particles and injected onto the column. A second run was performed by injecting ~1 mg of alcohol dehydrogenase in 300 μL gel filtration buffer. A standard curve was plotted using the known protein standards and the molecular mass of *Hp*-DAH7PS was estimated using its elution time from the column. Elution times for *Hp*-DAH7PS and standards are in Chapter Two, Section 2.6.

For the determination of the molecular mass of *Mt*-DAH7PS ~2 mg of carbonic anhydrase and ~3 mg of alcohol dehydrogenase were dissolved in 500 μL gel filtration buffer, centrifuged at max-speed on a Eppendorf bench-top centrifuge for 2 minutes to remove any small solid particles and injected onto the column. A second run was performed by injecting ~1 mg of cytochrome C and ~2 mg β-amylase in 500 μL gel filtration buffer. A standard curve was plotted using the known protein standards and the molecular mass of *Mt*-DAH7PS was estimated using its elution time from the column. The elution times for the four protein standards and *Mt*-DAH7PS from the gel filtration column were: cytochrome C, 44.4 mins; carbonic anhydrase, 40.7 mins; alcohol dehydrogenase, 31.6 mins; β-amylase, 28.5 mins and *Mt*-DAH7PS, 29.0 mins.

Standard Enzyme Assays and Kinetic Measurements

The assay system for DAH7PS was a modified form of the assay used by Schoner and Herrmann.[75] The consumption of PEP was monitored at 232 nm ($\varepsilon = 2.8 \times 10^3\ M^{-1}\ cm^{-1}$ at 30 °C) using either a Varian Cary 1 or 100 UV Visible spectrophotometer.

Measurements were made using 1 cm path length quartz cuvettes. Standard reaction mixtures contained PEP (~150 μM) (Research Chemicals), E4P (~150 μM) (Sigma), and $MnSO_4$ (~100 μM) (Sigma) in BTP buffer (50 mM, pH 7.5). The BTP buffer was treated with Chelex. PEP and E4P solutions were made up in buffer and then treated with Chelex. The $MnSO_4$ solution was made up with BTP buffer that had been pre-treated with Chelex. The reaction was initiated by the addition of enzyme to give a final volume of 1 mL. All assays were performed at 30 °C. Initial rates of reaction were determined by a least-squares fit of the initial rate data. One unit (1 U) of enzyme activity is defined as the loss of 1 μmol of PEP per minute at 30 °C. Specific activity is defined as the loss of 1 μmol of PEP per minute at 30 °C per mg of protein (U/mg). K_m and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using Enzfitter (Biosoft, 1999). Errors between measurements were no more than 10 %.

The concentration of PEP, E4P and E4P analogues was determined using this assay system. For example, to determine the concentration of E4P a reaction mixture consisting of >200 μM PEP, ~100 μM $MnSO_4$ and 5 μL of E4P stock solution (of unknown concentration) is made up and initiated with enzyme to give a final volume of 1 mL. The reaction is allowed to go to completion and the difference between the absorbance before enzyme initiation and the absorbance at completion of the reaction is measured (ΔA_1). This change in absorbance does not take into account the increase in absorbance due to the addition of enzyme so a correction factor is needed. To determine the correction factor an identical assay is performed in the absence of E4P, so that no reaction takes place. This allows the determination of the increase in absorbance due to the addition of enzyme to the cuvette (ΔA_2). The corrected change in absorbance is then given by $\Delta A_1 + \Delta A_2$. To convert absorbance into concentration Beer's Law is applied ($A=\varepsilon.c.l$, where l=1 cm and $\varepsilon = 2.8 \times 10^3\ M^{-1}\ cm^{-1}$ at 30 °C).

Crystallization Trials

Screening for crystallization conditions was achieved using sitting-drop vapour diffusion in 96-well Intelliplates (Hampton Research) at 18 °C by mixing 100 nL of protein solution with 100 nL of well solution. The initial screens were set down using the Centre for Molecular Biodiscovery Crystallization facility, which consists of a

modified Perkin Elmer Multiprobe robot for the transfer of mother liquor from pre-made stock solutions into Intelliplates and a Cartesian Honeybee robot for assembling the nanolitre crystallization experiments. Initial screens included the Hampton Crystal Screens I and II (Hampton Research), a systematic PEG-pH screen, a PEG/Ion screen (Hampton Research) and the footprint Screen. Screens using native and Se-Met *Mt*-DAH7PS used protein concentrated to ~3 mg/mL. Screening of *Hp*-DAH7PS involved using protein with a concentration of ~20 mg/mL.

Further manual screening was performed using hanging-drop vapor diffusion in 24-well plates (1 μL + 1 μL drops and 400 μL of appropriate well solution). Crystal trays were set up at room temperature and then transferred to 10 °C. Crystal trays, pipette tips and coverslips were blown, using compressed air, to remove any dust.

Data Collection

Crystals were transferred straight from the mother liquor drop, if the concentration of glycerol was >15 %, onto the goniometer using an appropriate sized loop attached to a crystal mount (Hampton). If a cryo-protectant was necessary the crystals were transferred into a ~ 5 μL drop of cryo-solution, left for ~ 1 minute and then mounted. In the case of crystals that were soaked, the crystals were transferred into a second hanging-drop in ~ 2 μL of appropriate soaking solution. Crystals were frozen upon mounting onto the goniometer by a stream of gaseous nitrogen at 110 K. Crystals that were stored for transfer to a synchrotron were frozen in the same manner but then transferred to a dewar filled with liquid nitrogen. Crystals were transported overseas in a dry-dewar.

Once the crystals were mounted onto the goniometer two images were taken 90° to each other (exposure time of ~10 minutes) to check the level of diffraction by the crystal. The spacing between spots on the image plates was also checked and because the spots overlapped the detector was moved out to 180 mm. If diffraction was good the two images were indexed and strategy was run to determine the optimum starting angle and required rotation for a complete dataset with sufficient redundancy (CrystalClear 1.3.2).

X-ray Diffraction Apparatus

The apparatus used in the collection of crystal data in-house was a Rigaku MicroMax-007 rotating anode X-ray generator (λ=1.5418 nm, 800 W, 40 kV), with Osmic Blue confocal optics, R-Axis IV++ image plate detector and an Oxford Series 700 cryostream. Data collected at the Advanced Light Source Source (Berkeley, CA) was performed using beamline 8.2 at a wavelength of 0.9796 Å (the peak of the Se absorption edge). Data that was collected at the Stanford Synchrotron Radiation Laboratory was perfomed at a wavelength of 0.97929 Å.

Processing of Data

Data that was collected in-house was processed using CrystalClear 1.3.6 (Rigaku) and d*TREK software. The processing of the data collected at the Advanced Light was performed by Minmin Yu using DENZO and SCALEPACK.[173] The processing of data at the Stanford Synchrotron Radiation Laboratory was performed using DENZO and SCALEPACK.

Other Crystallographic Methods

The stereochemical quality of the protein model was monitored with PROCHECK.[176] Protein-protein interactions were analyzed using the Protein-Protein Interaction Server (http://www.biochem.ucl.ac.uk/bsm/PP/server), based on principles described by Jones and Thornton.[195] Hydrogen bonds were identified following the criteria of Baker and Hubbard.[196] Structural figures were prepared with PYMOL (http://www.pymol.org). Superposition of molecules was performed by Lsqkab (CCP4[183]) and distances reported in this thesis were measured in COOT[178]and PYMOL.

Amino Acid Sequence Alignments

Amino acid sequence alignments displayed in this thesis were performed by copying and pasting the desired primary sequences into ClustalW (EMBL-EBI) (http://www.ebi.ac.uk/ClustalW) and performing a multiple sequence alignment.

7.2 EXPERIMENTAL FOR CHAPTER TWO

Expression, Solubilization and Biochemical Characterization of Type II DAH7PS from *H. pylori*

Amplification of the *Hp*-DAH7PS gene by PCR

DNA corresponding to the open reading frame (ORF) of *H. pylori* strain J99 DAH7PS (GenBank accession number Q9ZMU5) was amplified from *H. pylori* strain J99 genomic DNA using *Pfu* Turbo DNA polymerase (Stratagene). The primers used were: *forward* 5'-GATTATACATATGTCAAACACAACCTGGTCGC incorporating a *Nde*I restriction site and *reverse* 5'-TGGGATCCATTAAGTGCGTTGTTTTTTAAGC incorporating a *Bam*HI restriction site. The PCR reaction mixture consisted of ~7 ng of genomic DNA, 6 pmol of both primers, 250 µM of each of the deoxyribonucleotide triphosphates (dNTPs), 1x Mg^{2+} free polymerase buffer and 1 mM $MgCl_2$, and 2 units of *Pfu* Turbo DNA polymerase in a total reaction volume of 20 µL. The thermo-cycling program consisted of an initial denaturation at 95 °C for 3 minutes and this was followed by 35 cycles of denaturation for 30 seconds, primer annealing at 57 °C for 30 seconds with extension at 72 °C for 2 minutes. The 1.3 kbp PCR product was then purified directly using a Nucleospin Extract Kit (Machrey-Nagel).

Digestion with Restriction Endonucleases

The *Hp*-DAH7PS gene was digested with 20 units of *Nde*I in a total volume of 56 µL in 1x SuRE/cut Buffer H (50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 100 µM DTT, pH 7.5, Roche) at 37 °C for ~2 hours. The DNA was precipitated with 0.2 volumes of 3 M sodium acetate pH 5.2, 2 volumes of ethanol, stored at -80 °C overnight and then centrifuged (14,000 *g*, 15 minutes) at 4 °C. The supernatant was removed and the DNA pellet rinsed with 70 % ethanol, centrifuged again (14,000 *g*, 5 minutes) at 4 °C and then the pellet was air-dried and re-suspended in 20 µL Milli-Q water. Digestion with *Bam*HI involved incubation with 10 units of *Bam*HI in a total volume of 24 µL of 1x SuRE/cut Buffer B (10 mM Tris-HCl, 100 mM NaCl, 5 mM $MgCl_2$, 1 mM 2-mercaptoethanol, pH 8.0, Roche) for ~2 hours at 37 °C. The digested *Hp*-DAH7PS gene was then gel purified using a Nucleospin Extract Kit with a final concentration of ~100 ng/uL, as estimated by agarose gel.

Ligation of *Hp*-DAH7PS Gene into pET-32a(+)

The *Hp*-DAH7PS gene was ligated into pET-32a(+) previously digested with *Nde*I and *Bam*HI (kindly donated by Dr Mark Patchett) using T4 DNA ligase (Roche). The ligation reaction consisted of ~6 ng of pre-cut pET-32a(+), 0.2 units of ligase and ~25 ng of DNA insert made up to 5 μL with 1x ligation buffer (66 mM Tris-HCl, 5 mM $MgCl_2$, 5 mM DTT, 1 mM ATP, pH 7.5) and incubated overnight at 16 °C.

Transformation into *E. coli* XL1-Blue Cells

The ligation mixture was used to transform competent *E. coli* XL1-Blue (Stratagene). 50 μL of competent cells were thawed and incubated on ice for 30 minutes with 2 μL of ligation mixture. The cells were then heated shocked at 42 °C for 30 seconds and then returned to ice before 250 μL of SOC was added. The cells were incubated at 37 °C for 1 hour with shaking (220 rpm). The cells were plated onto LB-Amp plates following an appropriate dilution (1-100 fold) and incubated overnight at 37 °C. Plasmid DNA was isolated from 5 mL cultures in LB-Amp using a Quantum Prep plasmid mini-prep kit (Bio-Rad) and used as the template in PCR sequencing reactions to identify plasmids containing the DAH7PS ORF. The sequence of the insert DNA in these recombinant plasmids was identical to the *H. pylori* DAH7PS ORF. A sequenced pET-32a(+)-*Hpy*DAH7PS plasmid was used to transform, by electroporation, *E. coli* BL21(DE3) cells already containing the pGroESL plasmid (described in the General Methods Section).

Purification by Ion Exchange Chromatography (IEC)

IEC separates proteins based on their different net charges at a given pH. The sidechains of surface amino acids can be either protonated or deprotonated depending on the pH of the environment, so that the overall charge on the protein can dictate the protein's behavior on an ion-exchange matrix. This type of chromatography allows purification of a given protein based on its isoelectric point (pI). Stationary phase matrices can carry charged groups that are either negatively charged (Cation Exchange Chromatography (CEC)) or positively charged (Anion Exchange Chromatography (AEC)). The charges on these matrices are balanced by counter-ions such as Cl^- in the case of anion exchange and Na^+ for cation exchange. The net charge of the protein of

interest is the same as that of the counter-ions and therefore the protein competes with the counter-ions to bind to the charged matrix. If the protein carries no charge or the opposite charge of the counter-ions it will pass through the column in the mobile phase. In the case of *Hp*-DAH7PS the predicted pI is 7.5. At 1 pH unit above this (e.g pH 8.5) the overall charge on the protein is negative therefore AEC can be used as a purification step. The second step in the purification of *Hp*-DAH7PS involved decreasing the pH to 6.5 and using CEC.

Purification using AEC and CEC

The supernatant fraction collected after cell lysis and centrifugation was filtered using a 0.2 μM filter (Millipore) and diluted five-fold with buffer A (50 mM BTP, pH 8.5, 1 mM DTT) and loaded onto a column (8 mL) of Source 15Q® (10/10) (Amersham) equilibrated in buffer A at 4 °C. The unbound protein was removed with two column volumes of buffer A, and the bound protein was eluted with a 50 mL linear gradient between buffer A and buffer A + 0.33 M NaCl, at a flow rate of 1.5 mL/min. Active fractions (1 mL) were pooled, concentrated using a 20 mL Vivaspin 10 kDa MWCO concentrator (Vivascience), diluted five-fold with buffer B (50 mM BTP, pH 6.5, 1 mM DTT), and loaded onto either a Mono® S (5/5) column (Amersham) or a Source 15S® (5/10) column (Amersham) equilibrated in buffer B at 4 °C. The unbound protein was removed with two column volumes (16 mL) of buffer B, and the bound protein was eluted with a 45 mL linear gradient between buffer B and buffer B + 1 M NaCl, at a flow rate of 0.7 mL/min. Active fractions (1 mL) were pooled and concentrated using a 2 mL Vivaspin 10 kDa MWCO concentrator. The enzyme preparation could be further purified using size exclusion chromatography (refer to General Methods Section).

Michaelis-Menten Kinetics

For the determination of kinetic mechanism, PEP (9. 9 to 32.5 μM) and E4P (20 to 200 μM) concentrations were varied at different fixed concentrations of the other substrate (E4P: 13.5, 16.5, 22.5 and 43. 5 μM) (PEP: 20, 50, 100 and 200 μM) with $MnSO_4$ (94 μM) in BTP (50 mM, pH 7.5) buffer. Reactions were initiated by the addition of purified *Hp*-DAH7PS (2 μL, 10 mg/mL). All assays were performed in duplicate.

DAH7P for product inhibition studies was obtained by purification of a large-scale reaction between E4P and PEP catalyzed by *E. coli* DAH7PS as previously described.[197] DAH7P generated in this way was purified by AEC (Source 15Q®) eluting with ammonium bicarbonate (0 - 500 mM, pH 8.0). Fractions containing DAH7P were pooled and lyophilised. The assay mixtures used to determine DAH7P inhibition with respect to PEP contained E4P (120 μM), PEP (10, 20, 50 or 100 μM), $MnSO_4$ (100 μM) and DAH7P (0, 1 or 2 mM). Mixtures used to determine inhibition with respect to E4P contained PEP (250 μM), E4P (10, 20, 50 & 100 μM), $MnSO_4$ (100 μM) and DAH7P (0, 1, or 2 mM). Reactions were initiated by the addition of purified *Hp*-DAH7PS (2.5 μL, 12.3 mg/mL) and duplicate assays were performed for all reaction conditions.

Metal Activation

Hp-DAH7PS apoenzyme was prepared by incubation with EDTA (0.5 mM, pH 8.0) for 30 minutes at 4 °C prior to the assay. The final concentration of EDTA in the reaction mixture was 1 μM. The assay mixture contained PEP (196 μM), E4P (135 μM), and 100 μM metal salt (except $CdSO_4$, 300 μM) in BTP buffer (50 mM, pH 7.5, with EDTA (10 μM)). All solutions (except the metal salts) were pre-treated with Chelex. The reaction was initiated by the addition of apoenzyme (2 μL, 5 mg/mL). The metal salts used were $MnSO_4$ (Sigma), $ZnSO_4$ (BDH), $MgCl_2$ (BDH), $CdSO_4$ (May and Baker), $CoSO_4$ (May and Baker), $NiCl_2$ (May and Baker), $CuSO_4$ (May and Baker) and $CaCl_2$ (Prolabo). All metal solutions were made up in Chelexed Milli-Q water.

Effect of DTT

Hp-DAH7PS (0.2 mg/mL) was partially inactivated by incubation at 4 °C for 24 hours in BTP buffer (50 mM, pH 7.5) retaining approximately 20 % of its original activity. Reactivation with DTT was investigated by pre-incubating the enzyme in the presence and absence of 1 mM DTT at 30 °C for 10 minutes prior to initiation of an assay. Standard assay conditions were used (refer to General Methods Section) with the addition of enzyme to the cuvette followed by initiation with E4P.

Substrate Specificity

Assays to determine activity with DL-glyceraldehyde 3-phosphate (G3P) and D-glucose 6-phosphate (G6P) contained 250 μM PEP and 6 mM G3P (Sigma) or 1 mM G6P (Sigma). Activity on a range of five-carbon sugars was also investigated. To determine kinetics for PEP the reaction mixture contained 4 – 505 μM PEP and either 10 mM A5P (Sigma) or 8 mM R5P (Research Organics Inc.) or 5.5 mM 2dR5P (Sigma). To determine kinetics for the phosphate sugar the reaction mixture contained either 300 μM PEP and 1 - 15 mM A5P or 202 μM PEP and 1 - 15 mM R5P, or 253 μM PEP and 0.5 – 30 mM 2dR5P. The reaction was initiated by the addition of *Hp*-DAH7PS (2 μL, ~2 mg/mL). To confirm the eight-carbon sugar formed from PEP and R5P, a large-scale reaction was carried out. PEP and R5P were dissolved in water (2.5 mL) and $MnSO_4$ was added (to give a final concentration of 100 μM). The pH of the solution was adjusted to 7.1 with 10 M NaOH. The total volume was made up to 3 mL, *Hp*-DAH7PS was added (40 μL, 9 mg/mL), and the reaction followed at 232 nm. After the loss of PEP had ceased the mixture was filtered to remove enzyme (10 kDa MWCO ultrafiltration device) and the product was purified by AEC (Source 15Q®, eluted with a 175 mL linear gradient of 0 - 500 mM ammonium bicarbonate). Fractions containing product were pooled and lyophilized. The white solid was redissolved in D_2O and analyzed by ^1H NMR (Bruker Avance 400).

Facial Selectivity of the *Hp*-DAH7PS Catalyzed Reaction

(*Z*)-3-FluoroPEP (>95 % stereochemically pure) was synthesized according to reported procedures.[198] The reaction mixture initially contained (*Z*)-3-fluoroPEP (1 mM), E4P (100 μM) and $MnSO_4$ (100 μM), in buffer (50 mM BTP, pH 7.5). *Hp*-DAH7PS (5 μL, 20 mg/mL) was added and the disappearance of 3-fluoroPEP was followed at 232 nm ($\varepsilon = 1.25 \times 10^3\ M^{-1}cm^{-1}$). The reaction was performed at 30 °C. As the reaction rate slowed due to the depletion of substrates more E4P and 3-fluoroPEP were added (which would have been equivalent to an initial concentration of 4 mM for 3-fluoroPEP and 1.5 mM for E4P). When the reaction had ceased, the reaction mixture was treated with Chelex overnight at room temperture, filtered (2 mL, 10 kDa MWCO) dissolved in D_2O and analyzed by ^{19}F NMR (Bruker Avance 400). ^{19}F chemical-shifts were referenced to $CFCl_3$.

7.3 EXPERIMENTAL FOR CHAPTER THREE

Structural & Functional Characterization of the Type II DAH7PS from *M. tuberculosis*

Cloning of the Gene for *Mt*-DAH7PS

The cloning of the gene for *Mt*-DAH7PS into pProEx-HTa (Novagen) expression vector was performed by Dr Shaun Lott prior to the start of this project. The DNA encoding the opening reading frame encoding *Mt*-DAH7PS (Rv2178c) was amplified from *M. tuberculosis* H37Rv genomic DNA using the PCR with primers designed to introduce *Nco*I and *Sac*I restriction sites at the 5' and 3' ends, respectively. The PCR product was cloned into pProEx-HTa, yielding a protein containing an N-terminal tobacco etch virus (TEV) protease-cleavable His tag upstream of the predicted translation start site.

Production of Se-Met substituted *Mt*-DAH7PS Protein

For production of selenomethionine (Se-Met)-substituted protein, the pProEx-HTa-*Mt*-DAH7PS plasmid construct was transformed into cells of the methionine auxotroph DL41 (DE3) that also contained pGroESL. DL41(DE3) cells were transformed with pGroESL, electro-competent DL41(DE3)-pGroESL cells prepared and pProEx-HTa-*Mt*-DAH7PS transformed into the cells using electroporation (protocols described in the General Methods Section). The transformed cells were grown in minimal media (500mL) (described below) with Se-Met (1 mL of a 12.5 mg/mL stock solution) the only methionine source. The Se-Met stock solution was made up with autoclaved Milli-Q water and filter-sterilized (0.2 μM). Cells were otherwise grown following the method described in the General Methods Section.

LeMaster Medium for Se-Met Protein Production:

LeMaster solution I

Chemical	Amount needed to make 1 L (g)
Ala	0.5
Arg HCl	0.58

Asp	0.4
Cys	0.03
Glu	0.67
Gln	0.33
Gly	0.54
His	0.06
Iso	0.23
Leu	0.23
Lys HCl	0.42
Phe	0.13
Pro	0.1
Ser	2.08
Thr	0.23
Tyr	0.17
Val	0.23
Adenine	0.5
Guanosine	0.67
Thymine	0.17
Uracil	0.5
Sodium acetate	1.5
Succinic Acid	1.5
Ammonium chloride	0.75
K_2HPO_4	10.5
Total	**23.02**

To make up the LeMaster solution I 11.5 g of the powder mix and 0.5 g of NaOH were made up to 500 mL with Milli-Q water and autoclaved.

LeMaster solution II

To make up the carbon source solution, the following was dissolved in 10 mL of Milli-Q water and then filter-sterilized (0.2 μM):

Ingredient	Per 10 mL
Glucose	1 g
$MgSO_4.7H_2O$	0.025 g
$FeSO_4.7H_2O$	0.00043 g
Conc. H_2SO_4	0.83 µL
Thiamine (vitamin B2)	0.0001 g

To make up the complete medium 50 mL of carbon source was aded to 450 mL of LeMaster solution I and finally 1 mL of 12.5 mg/mL of Se-Met was added.

Effect of EDTA and PEP on the Stability of *Mt*-DAH7PS Activity

A NAP-5 column (Pharmacia) desalting/buffer exchange column pre-packed with Sephadex G-25 media was equilibrated with approximately 10 mL buffer. The buffer contained BTP (20 mM, pH 7.0), NaCl (150 mM), PEP (200 µM) and $MnSO_4$ (200 µM). ~100 µL purified *Mt*-DAH7PS was loaded onto the column and allowed to completely enter the gel bed before 400 µL buffer was added. The protein was eluted with ~1 mL of buffer and concentrated to ~100 µL using a 2 mL 10 kDa MWCO concentrator. The enzyme sample was then assayed using the following assay conditions: PEP (175 µM), E4P (80 µM) $MnSO_4$ (100 µM) made up to 1 mL with BTP (50 mM, pH 7.5) and initiated by the addition of *Mt*-DAH7PS (2 µL, 4.6 mg/mL). Duplicate columns were performed simultaneously; one with TCEP (0.5 mM) and the other without TCEP in the buffer. Columns were run with and without TCEP under several further conditions; (1) with EDTA substituting $MnSO_4$ and (2) no PEP. Specific activity of *Mt*-DAH7PS after buffer exchange using the NAP-5 column was determined. To determine whether loss of specific activity could be regenerated upon addition of TCEP, both enzyme preparations (with and without TCEP) were then put down a second NAP-5 column (using the above protocol) in the presence of TCEP.

Metal Activation of *Mt*-DAH7PS

For analysis of divalent metal activation of *Mt*-DAH7PS the enzyme was treated with EDTA (250 mM, pH 8.0) for 24 hours at 4 °C. The EDTA concentration was then reduced to 50 mM, by dilution and concentration using a 2 mL 10 kDa MWCO concentrator, prior to assay. The final concentration of EDTA in the reaction mixture was 100 μM. The activity was determined upon addition of metal salt. The reaction mixture contained PEP (174 μM), E4P (178 μM), and 200 μM metal salt in BTP (50 mM) pH 7.5. The reaction was initiated by the addition of enzyme (2.5 μL, 2.1 mg/mL). The metal salts were $MnSO_4$, $ZnSO_4$, $MgCl_2$, $CdSO_4$, $CoSO_4$, $NiCl_2$, $CuSO_4$, $CaCl_2$, and $FeCl_2$ (Riedel-de Haën).

pH Profile of *Mt*-DAH7PS

To determine the pH profile of *Mt*-DAH7PS assays were performed over a pH range from 3.8-9.2. Buffers used were 100 mM BTP for pHs in the range 6.3-9.2 and 100 mM acetate buffer for pHs in the range 3.8-5.7. The acetate buffer was prepared from two solutions of 100 mM sodium acetate and 100 mM acetic acid. The assays used to determine activity consisted of PEP (220 μM), E4P (180 μM), $MnSO_4$ (100 μM) made up to 1 mL in a quartz cuvette. The reaction was initiated by the addition of *Mt*-DAH7PS (3 μL, 3.6 mg/mL).

Solving the Crystal Structure of Se-Met *Mt*-DAH7PS using SAD Methods

Single wavelength anomalous dispersion (SAD) methods were used to solve the structure of Se-Met *Mt*-DAH7PS. SAD requires the presence of heavy and/or anomalous scattering atoms. The most popular is Se-Met which is introduced into proteins in place of normally occurring methionine by genetic engineering. Anomalous scattering of X-ray radiation by heavy or anomalous scattering atoms leads to a breakdown of Friedels' law, which states that the intensity of the reflection *hkl* is identical to the intensity of the reflection *–h-k-l*.[199] The ability of the heavy atom to absorb X-rays of a specified wavelength results in small differences in intensity between reflections *hkl* and *–h-k-l* (referred to as Bijvoet differences). You can use the nonequivalence of Friedal pairs in the anomalous scattering data to establish phases from the anomalous scattering atoms (e.g Se), and use it to provide phase information

for the determination of the structure of the macromolecule. The availability of accurate detectors and improved data processing algorithms now makes it possible to precisely measure the small Bijvoet differences which can be used for SAD phasing by new efficient phasing programs.[200] A description of the methods used to determine the structure of Se-Met *Mt*-DAH7PS is in Chapter Three, Section 3.16. The structure determination of Se-Met *Mt*-DAH7PS was performed by Minmin Yu at the Advanced Light Source (Berkeley, CA).

7.4 EXPERIMENTAL FOR CHAPTER FOUR

Feedback Regulation of Type II DAH7PS from *M. tuberculosis* and *H. pylori*

Feedback-inhibition Studies with *Hp*-DAH7PS

Solutions of Phe (20 mM) (Sigma), Tyr (10 mM) (Sigma) and Trp (20 mM) (Sigma) in water alone or in combination were added to standard assay reaction mixtures (PEP (75 μM), E4P (61μM) and $MnSO_4$ (100 μM) in BTP buffer (50 mM, pH 7.5)) to give a final total aromatic amino acid concentration of 250 μM. The reaction was initiated by the addition of *Hp*-DAH7PS (1 μL, 10 mg/mL). A freshly prepared solution of chorismic acid (10 mM) in water was added to an assay system (PEP (300 μM), E4P (100 μM) and $MnSO_4$ (100 μM)) to give final concentrations of 200 μM or 400 μM.

Feedback-inhibition Studies with *Mt*-DAH7PS

Solutions of Phe, Tyr, Trp and chorismic acid were made up as described in the previous section. The aromatic amino acids and chorismic acid were added alone or in combination to standard assay reaction mixtures to give a final concentration of 200 μM for each aromatic amino acid or chorismic acid. The assay solution contained PEP (200 μM), E4P (142 μM) and $MnSO_4$ (100 μM) in BTP buffer (50 mM, pH 7.5). Assays performed using different ratios of Trp and Phe, or Trp and Tyr are described in Chapter Four. The total concentration of aromatic amino acid in the reaction mixture was 200 μM for all assays. Assays were also performed in the presence of either Trp (0-200 μM), Phe (0-200 μM), or Tyr (0-200 μM). All assay solutions contained PEP (250 μM), E4P (140 μM) and $MnSO_4$ (100 μM) in BTP buffer (50 mM, pH 7.5). All reactions

were initiated by the addition of *Mt*-DAH7PS (2 μL, 4.6 mg/mL) to give a total reaction volume of 1 mL. All reactions were performed in duplicate.

Soaking of Native Crystals in Trp and Phe

To obtain a crystal structure of *Mt*-DAH7PS in the presence of Trp and Phe alone and in combination, native crystals were grown in 0.1 M Tris-HCl pH 8.0, 1.5 M ammonium sulfate, 12 % (v/v) glycerol overnight at 10 °C. The crystals were then transferred to a second hanging drop consisting of mother liquor supplemented with cryo-agent (15-18 % glycerol) and the appropriate aromatic amino acid (~2 mM) (referred to as the soaking solution). This hanging-drop consisted of a 2 μL drop above a reservoir of 400 μL soaking solution. The crystals were allowed to equilibrate in the soaking solution at room temperature for the desired period of time before they were mounted onto the goniostat.

Trp and Phe Inhibition of *Mt*-DAH7PS with Respect to PEP and E4P

The assay mixtures used to determine the effect of the aromatic amino acids on the Michaelis-Menten plot with respect to PEP contained PEP (4-429 μM), E4P (185 μM), $MnSO_4$ (100 μM) in BTP (50 mM, pH 7.5). The effect of Trp and Phe on the Michaelis-Menten plot with respect to E4P used assay mixtures that contained E4P (8-160 μM), PEP (200 μM), $MnSO_4$ (100 μM) in BTP (50 mM, pH 7.5). Reaction mixtures either had no aromatic amino acids, 200 μM Trp or 100 μM Trp + 100 μM Phe. The reaction was initiated by the addition of *Mt*-DAH7PS (2 μL, 2.6 mg/mL) and duplicate assays were performed for all reaction conditions.

7.5 EXPERIMENTAL FOR CHAPTER FIVE

Investigating the Role Arg284 and Cys440 play in Catalysis by *Mt*-DAH7PS

Cloning of the *Mt*-DAH7PS R284K and C440S Genes

The R284K and C440S mutations were performed using the QuikChange® II Site-Directed Mutagenesis Kit (Stratagene), and pPro-ExHTa-*Mt*-DAH7PS as the double-stranded plasmid template. The single amino acid substitutions (either an Arg to a Lys

or a Cys to a Ser) were introduced using synthetic oligonucleotide primers, each complementary to opposite strands of pPro-ExHTa-*Mt*-DAH7PS, containing the desired mutation. The primers used for C440S were: *forward* 5'- CTATGAGACGGCAAGTG ATCCGCGGCT and *reverse* 5'- AGCCGCGGATCACTTGCCGTCTCATAG incorporating a Cys to Ser mutation (TGT to AGT in the *forward* primer and ACA to ACT in the *reverse* primer). The primers used for R284K were: *forward* 5'- GGATCG GCGAGAAGACACGACAAATCGATGG and *reverse* 5'- CCATCGATTTGTCGTG TCTTCTCGCCGATCC incorporating a Arg to Lys mutation (CGG to AAG in the *forward* primer and CCG to CTT in the *reverse* primer). The procedure included three steps (described in the QuikChange® manual): the primers with the desired mutation were extended during temperature cycling by *PfuUltra* high fidelity DNA polymerase, generating a mutated plasmid containing sticky ends. The product was then treated with *Dpn* I to digest the parental DNA template and the pPro-ExHTa-mutant construct was transformed into XL1-Blue supercompetent cells supplied in the kit. Plasmid mini-prep DNA was isolated from transformants using a HighPure Plasmid Isolation Kit (Roche) and the insert DNA sequenced to verify the sequence was correct before the plasmid-mutant construct was used to transform, by electroporation, *E. coli* BL21(DE3) already containing the pGroESL plasmid (described in the General Methods Section).

REFERENCES

(1) Bentley, R., The shikimate pathway-a metabolic tree with many branches. *Crit. Rev. Biochem. Mol. Biol.* **1990,** 25, 307-384.

(2) Haslam, E., *Shikimic Acid-Metabolism and Metabolites.* 1994; 57, p 1470.

(3) Walsh, C. T.; Liu, J.; Rusnak, F.; Sakaitani, M., Molecular studies on enzymes in chorismate metabolism and the enterobactin biosynthetic pathway. *Chem. Rev.* **1990,** 90, (7), 1105-29.

(4) Hudson, A. T.; Campbell, I. M.; Bentley, R., Biosynthesis of 6-methylsalicylic acid by *Mycobacterium phlei. Biochemistry* **1970,** 9, (20), 3988-92.

(5) Kishore, G. M.; Shah, D. M., Amino acid biosynthesis inhibitors as herbicides. *Annu. Rev. Biochem.* **1988,** 57, 627-63.

(6) Steinrucken, H. C.; Amrhein, N., The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid-3-phosphate synthase. *Biochem. Biophys. Res. Commun.* **1980,** 94, (4), 1207-1212.

(7) Schonbrunn, E.; Eschenburg, S.; Shuttleworth, W. A.; Schloss, J. V.; Amrhein, N.; Evans, J. N.; Kabsch, W., Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. *Proc. Natl. Acad. Sci.* **2001,** 98, (4), 1376-1380.

(8) Eykmann, J. F., *Recl. Trav. Chim* **1885,** 4, 32.

(9) Haslam, E., *The Shikimate Pathway.* Butterworth & Co (Publishers) Ltd: London, 1974.

(10) Fischer, H. O. L.; Dangschat, G., Quinic acid and derivatives. VIII. Configuration of shikimic acid and degradation to glucodesonic acid. (Relation of cyclic plant acids to carbohydrates.). *Helv. Chim. Acta* **1937,** 20, 705-16.

(11) Fischer, H. O. L.; Dangschat, G.; Taube, C.; Radt, F.; Stettiner, H., Quinic acid and derivatives. II. Constitution and configuration of quinic acid. *Chem. Ber.* **1932,** 65, 1009-31.

(12) Fischer, H. O. L.; Dangschat, G., Quinic acid and derivatives. VI. Degradation of shikimic acid to aconitic acid. *Helv. Chim. Acta* **1935,** 18, 1204-6.

(13) Fischer, H. O. L.; Dangschat, G., Quinic acid and derivatives. V. Constitution of shikimic acid. *Helv. Chim. Acta* **1934,** 17, 1200-7.

(14) Davis, B. D., Aromatic biosynthesis. I. The role of shikimic acid. *J. Biol. Chem.* **1951,** 191, (1), 315-25.

(15) Davis, B. D., Aromatic biosynthesis. V. Antagonism between shikimic acid and its precursor, 5-dehydroshikimic acid. *J. Bacteriol.* **1952,** 64, (5), 749-63.

(16) Weiss, U.; Davis, B. D.; Mingioli, E. S., Aromatic biosynthesis. X. Identification of an early precursor as 5-dehydroquinic acid. *J. Am. Chem. Soc.* **1953,** 75, 5572-6.

(17) Salamon, I. I.; Davis, B. D., Aromatic biosynthesis. IX. The isolation of a precursor of shikimic acid. *J. Am. Chem. Soc.* **1953,** 75, 5567-71.

(18) Gibson, F.; Jackman, L. M., Structure of chorismic acid, a new intermediate in aromatic biosynthesis. *Nature* **1963,** 198, 388-9.

(19) Davis, B. D.; Mingioli, E. S., Aromatic biosynthesis. VII. Accumulation of two derivatives of shikimic acid by bacterial mutants. *J. Bacteriol.* **1953,** 66, 129-36.

(20) Coggins, J. R.; Boocock, M. R.; Chaudhuri, S.; Lambert, J. M.; Lumsden, J.; Nimmo, G. A.; Smith, D. D., The arom multifunctional enzyme from *Neurospora crassa*. *Methods Enzymol.* **1987,** 142, 325-41.

(21) Srinivasan, P. R.; Shigeura, H. T.; Sprecher, M.; Sprinson, D. B.; Davis, B. D., The biosynthesis of shikimic acid from D-glucose. *J. Biol. Chem.* **1956,** 220, 477-97.

(22) Srinivasan, P. R.; Sprinson, D. B.; Kalan, E. B.; Davis, B. D., The enzymic conversion of sedoheptulose-1,7-diphosphate to shikimic acid. *J. Biol. Chem.* **1956,** 223, 913-20.

(23) Berlyn, M. B.; Giles, N. H., Organization of enzymes in the polyaromatic synthetic pathway: separability in bacteria. *J. Bacteriol.* **1969,** 99, (1), 222-30.

(24) Pittard, J.; Wallace, B. J., Gene controlling the uptake of shikimic acid by *Escherichia coli*. *J. Bacteriol.* **1966,** 92, (4), 1070-5.

(25) Larimer, F. W.; Morse, C. C.; Beck, A. K.; Cole, K. W.; Gaertner, F. H., Isolation of the ARO1 cluster gene of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **1983,** 3, (9), 1609-14.

(26) Ahmed, S. I.; Giles, N. H., Organization of enzymes in the common aromatic synthetic pathway: evidence for aggregation in fungi. *J. Bacteriol.* **1969,** 99, (1), 231-7.

(27) Giles, N. H.; Case, M. E.; Partridge, C. W.; Ahmed, S. I., A gene cluster in *Neurospora crassa* coding for an aggregate of five aromatic synthetic enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **1967,** 58, (4), 1453-60.

(28) Llewellyn, D. J.; Daday, A.; Smith, G. D., Evidence for an artificially evolved bifunctional 3-deoxy-D-*arabino*heptulosonate-7-phosphate synthase-chorismate mutase in *Bacillus subtilis*. *J. Biol. Chem.* **1980,** 255, (5), 2077-84.

(29) Wu, J.; Sheflyan, G. Y.; Woodard, R. W., Bacillus subtilis 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase revisited: resolution of two long-standing enigmas. *Biochem. J.* **2005,** 390, (2), 583-590.

(30) Wu, J.; Woodard, R. W., New Insights into the Evolutionary Links Relating to the 3-Deoxy-D-*arabino*-heptulosonate 7-Phosphate Synthase Subfamilies. *J. Biol. Chem.* **2006,** 281, (7), 4042-4048.

(31) Mousdale, D. M.; Campbell, M. S.; Coggins, J. R., Purification And Characterization Of Bifunctional Dehydroquinase-Shikimate: NADP Oxidoreductase From Pea Seedlings. *Phytochem.* **1987,** 26, (10), 2665-2670.

(32) Coggins, J. R.; Boocock, M. R.; Campbell, M. S.; Chaudhuri, S.; Lambert, J. M.; Lewendon, A.; Mousdale, D. M.; Smith, D. D., Functional domains involved in aromatic amino acid biosynthesis. *Biochem. Soc. Trans.* **1985,** 13, (2), 299-303.

(33) Fowden, L., The chemical approach to plants. *Sci. Progr. (London)* **1965,** 53, (212), 583-99.

(34) Herrmann, K. M.; Weaver, L. M., The shikimate pathway. *Annu. Rev. Plant Physiol. Mol. Biol.* **1999,** 50, 473-503.

(35) Morris, P. F.; Doong, R.-L.; Jensen, R. A., Evidence from *Solanum tuberosum* in Support of the Dual-Pathway Hypothesis of Aromatic Biosynthesis. *Plant Physiol.* **1989,** 89, 10-14.

(36) Jensen, R. A.; Xie, G.; Calhoun, D. H.; Bonner, C. A., The correct phylogenetic relationship of KdsA (3-deoxy-*manno*-octulosonate 8-phosphate synthase) with one of two independently evolved classes of AroA (3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase). *J. Mol. Evol.* **2002,** 54, 416-423.

(37) Walker, G. E.; Dunbar, B.; Hunter, I. S., Evidence for a novel class of microbial 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase in *Streptomyces coelicolor* A#(2), *Streptomyces rimosus* and *Neurospora crassa*. *Microbiology* **1996,** 142, 1973-1982.

(38) Gosset, G.; Bonner, C. A.; Jensen, R. A., Microbial origin of plant-type 2 keto-3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthases, exemplified by the chorismate- and tryptophan-regulated enzyme from *Xanthomonas campestris*. *J. Bacteriol.* **2001,** 183(13), (July), 4061-4070.

(39) Subramaniam, P. S.; Xie, G.; Xia, T.; Jensen, R. A., Substrate ambiguity of 3-Deoxy-D-*manno*-Octulosonate 8-Phosphate Synthase from *Neisseria gonorrhoeae* in the Context of Its Membership in a Protein family containing a Subset of 3-deoxy-D-*arabino*-heptulosonate 7-Phosphate Synthases. *J. Bacteriol.* **1998,** 180, (1), 119-127.

(40) Raetz, C. R. H., Biochemistry of endotoxins. *Annu. Rev. Biochem.* **1990,** 59, 129-70.

(41) Kloosterman, H.; Hessels, G. I.; Vrijbloed, J. W.; Euverink, G. J.; Dijkhuizen, L., (De)regulation of key enzyme steps in the shikimate pathway and phenylalanine-specific pathway of the actinomycete *Amycolatopsis methanolica*. *Microbiology* **2003,** 149, (11), 3321-3330.

(42) Silakowski, B.; Kunze, B.; Muller, R., Stigmatella aurantiaca Sg a15 carries genes encoding type I and type II 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthases: involvement of a type II synthase in aurachin biosynthesis. *Arch. Microbiol.* **2000,** 173, (5-6), 403-11.

(43) Guo, J.; Frost, J. W., Kanosamine biosynthesis: a likely source of the aminoshikimate pathway's nitrogen atom. *J. Am. Chem. Soc.* **2002,** 124, (36), 10642-10643.

(44) Gourley, D. G.; Shrive, A. K.; Polikarpov, I.; Krell, T.; Coggins, J. R.; Hawkins, A. R.; Isaacs, N. W.; Sawyer, L., The two types of 3-dehydroquinase have distinct structures but catalyze the same overall reaction. *Nat. Struct. Biol.* **1999,** 6, (6), 521-5.

(45) Harris, J.; Kleanthous, C.; Coggins, J. R.; Hawkins, A. R.; Abell, C., Different mechanistic and stereochemical courses for the reactions catalyzed by type I and type II dehydroquinases. *J. Chem. Soc., Chem. Commun.* **1993,** (13), 1080-1.

(46) Ogino, T.; Garner, C.; Markley, J. L.; Herrmann, K. M., Biosynthesis of aromatic compounds: ^{13}C NMR spectroscopy of whole *Escherichia coli* cells. *Proc. Natl. Acad. Sci. U. S. A.* **1982,** 79, (19), 5828-32.

(47) Schofield, L. R.; Anderson, B. F.; Patchett, M. L.; Norris, G. E.; Jameson, G. B.; Parker, E. J., Substrate Ambiguity and Crystal Structure of *Pyrococcus furiosus* 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate Synthase: An Ancestral 3-Deoxyald-2-ulosonate-phosphate Synthase? *Biochemistry* **2005,** 44, 11950-11962.

(48) Tribe, D. E.; Camakaris, H.; Pittard, J., Constitutive and repressible enzymes of the common pathway of aromatic biosynthesis in *Escherichia coli* K-12: regulation of enzyme synthesis at different growth rates. *J. Bacteriol.* **1976,** 127, (3), 1085-97.

(49) Hartmann, M.; Schneider, T. R.; Pfeil, A.; Heinrich, G.; Lipscomb, W. N.; Braus, G. H., Evolution of feedback-inhibited β/α barrel isoenzymes by gene duplication and a single mutation. *Proc. Natl. Acad. Sci.* **2003,** 100, (3), 862-867.

(50) Euverink, G. J.; Hessels, G. I.; Franke, C.; Dijkhuizen, L., Chorismate mutase and 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase of the methylotrophic actinomycete *Amycolatopsis methanolica*. *Appl. Environ. Microbiol.* **1995,** 61, (11), 3796-803.

(51) Shumilin, I. A.; Bauerle, R.; Wu, J.; Woodard, R. W.; Kretsinger, R. H., Crystal structure of the reaction complex of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Thermotoga maritima* refines the catalytic mechanism and indicates a new mechanism of allosteric regulation. *J. Mol. Biol.* **2004,** 341, (2), 455-466.

(52) Hu, C.; Jiang, P.; Xu, J.; Wu, Y.; Huang, W., Mutation analysis of the feedback inhibition site of phenylalanine-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase of *Escherichia coli*. *J. Basic Microbiol.* **2003,** 43, (5), 399-406.

(53) König, V.; Pfeil, A.; Braus Gerhard, H.; Schneider Thomas, R., Substrate and metal complexes of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Saccharomyces cerevisiae* provide new insights into the catalytic mechanism. *J. Mol. Biol.* **2004,** 337, (3), 675-90.

(54) Shumilin, I., A.; Zhao, C.; Bauerle, R.; Kretsinger, R., H., Allosteric inhibition of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase alters the coordination of both substrates. *J. Mol. Biol.* **2002,** 320, (5), 1147-56.

(55) Jossek, R.; Bongaerts, J.; Sprenger, G. A., Characterization of a new feedback-resistant 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase AroF of *Escherichia coli*. *FEMS Microbiol. Lett.* **2001,** 202, (1), 145-8.

(56) Akowski, J. P.; Bauerle, R. H., Steady-state kinetics and inhibitor binding of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (tryptophan sensitive) from *Escherichia coli*. *Biochemistry* **1997,** 36, 15817-15822.

(57) Im, S. W. K.; Davidson, H.; Pittard, J., Phenylalanine and tyrosine biosynthesis in *Escherichia coli* K-12. Mutants derepressed for 3-deoxy-D-*arabino*heptulosonic acid 7-phosphate synthetase (phe), 3-deoxy-D-*arabino*heptulosonic acid 7-phosphate synthetase (tyr), chorismate mutase T-prephenate dehydrogenase, and transaminase A. *J. Bacteriol.* **1971,** 108, (1), 400-9.

(58) Brown, K. D.; Somerville, R. L., Repression of aromatic amino acid biosynthesis in *Escherichia coli* K-12. *J. Bacteriol.* **1971,** 108, (1), 386-99.

(59) Grove, C. L.; Gunsalus, R. P., Regulation of the aroH operon of *Escherichia coli* by the tryptophan repressor. *J. Bacteriol.* **1987,** 169, (5), 2158-64.

(60) Arvidson, D. N.; Bruce, C.; Gunsalus, R. P., Interaction of the *Escherichia coli* trp aporepressor with its ligand, L-tryptophan. *J. Biol. Chem.* **1986,** 261, (1), 238-43.

(61) Zurawski, G.; Gunsalus, R. P.; Brown, K. D.; Yanofsky, C., Structure and regulation of aroH, the structural gene for the tryptophan-repressible 3-deoxy-D-*arabino*-heptulosonic acid-7-phosphate synthetase of *Escherichia coli*. *J. Mol. Biol.* **1981,** 145, (1), 47-73.

(62) Gunsalus, R. P.; Yanofsky, C., Nucleotide sequence and expression of *Escherichia coli* trpR, the structural gene for the trp aporepressor. *Proc. Natl. Acad. Sci. U. S. A.* **1980,** 77, (12), 7117-21.

(63) Hinnebusch, A. G., Mechanisms of gene regulation in the general control of amino acid biosynthesis in *Saccharomyces cerevisiae*. *Microbiol. Rev.* **1988,** 52, (2), 248-73.

(64) Pereira, S. A.; Livi, G. P., Cloning and expression of the *ARO3* gene encoding DAHP synthase from *Candida albicans*. *Gene* **1993,** 132, 159-165.

(65) Schnappauf, G.; Hartmann, M.; Kunzler, M.; Braus, G. H., The two 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase isoenzymes from *Saccharomyces cerevisiae* show different kinetic modes of inhibition. *Arch. Microbiol.* **1998,** 169, 517-524.

(66) Kunzler, M.; Paravicini, G.; Egli, C. M.; Irniger, S.; Braus, G. H., Cloning, primary structure and regulation of the ARO4 gene, encoding the tyrosine-inhibited 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Saccharomyces cerevisiae*. *Gene* **1992,** 113, (1), 67-74.

(67) Entus, R.; Poling, M.; Herrmann Klaus, M., Redox regulation of *Arabidopsis* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *Plant Physiol.* **2002,** 129, (4), 1866-71.

(68) Suzich, J. A.; Dean, J. F. D.; Herrmann, K. M., 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from carrot root (*Daucus carota*) is a hysteretic enzyme. *Plant Physiol.* **1985,** 79, (3), 765-70.

(69) Pinto, J. E. B. P.; Dyer, W. E.; Weller, S. C.; Herrmann, K. M., Glyphosate induces 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase in potato (*Solanum tuberosum L.*) cells grown in suspension culture. *Plant Physiol.* **1988,** 87, (4), 891-3.

(70) Ray, J. M.; Yanofsky, C.; Bauerle, R., Mutational Analysis of the Catalytic and Feedback Sites of the Tryptophan-Sensitive 3-Deoxy-D-*arabino*-Heptulosonate-7-Phosphate Synthase of *Eschericha coli*. *J. Bacteriol.* **1988,** 170, (12), 5500-5506.

(71) Shumilin, I. A.; Kretsinger, R. H.; Bauerle, R. H., Crystal structure of phenylalanine-regulated 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*. *Structure* **1999,** 7, 865-875.

(72) Abell, C., Enzymology and Molecular Biology of the Shikimate Pathway. *Comprehensive Natural Products Chemistry* **1999,** 1, 573-607.

(73) Shultz, J.; Hermodson, M. A.; Garner, C. C.; Herrmann, K. M., The nucleotide sequence of the aroF gene of *Escherichia coli* and the amino acid sequence of the encoded protein, the tyrosine-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *J. Biol. Chem.* **1984,** 259, (15), 9655-61.

(74) McCandliss, R. J.; Poling, M. D.; Herrmann, K. M., 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. Purification and molecular characterization of the phenylalanine-sensitive isoenzyme from *Escherichia coli*. *J. Biol. Chem.* **1978,** 253, (12), 4259-65.

(75) Schoner, R.; Herrmann, K. M., 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. Purification, properties, and kinetics of the tyrosine-sensitive isoenzyme from *Escherichia coli*. *J. Biol. Chem.* **1976,** 251, 5440-5447.

(76) Ray, J. M.; Bauerle, R., Purification and properties of tryptophan-sensitive 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*. *J. Bacteriol.* **1991,** 173, (6), 1894-901.

(77) Sheflyan, G. Y.; Howe, D. L.; Wilson, T. L.; Woodard, R. W., Enzymatic Synthesis of 3-Deoxy-D-*manno*-octulosonate 8-Phosphate, 3-Deoxy-D-*altro*-octulosonate 8-Phosphate, 3,5 -Dideoxy-D-*gluco(manno)*-octulosonate 8-Phosphate by 3-Deoxy-D-*arabino*-heptulosonate 7-Phosphate Synthase. *J. Am. Chem. Soc.* **1998,** 120, (43), 11027-11032.

(78) Sundaram, A. K.; Woodard, R. W., *Neisseria gonorrhoeae* 3-Deoxy-D-*arabino*-heptulosonate 7-Phosphate Synthase Does not Catalyse the Formation of the *ribo* Analogue. *Org. Letters* **2001,** 3, (1), 21-24.

(79) Pietersma, A. L.

(80) Ahn, M.

(81) Stephens, C. M.; Bauerle, R., Analysis of the Metal Requirement of 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Eschericha coli*. *J. Biol. Chem.* **1991,** 266, (31), 20810-20817.

(82) Baasov, T.; Knowles, J. R., Is the First Enzyme of the Shikimate Pathway, 3-Deoxy-D-*arabino*-Heptulosonate-7-Phosphate Synthase (Tyrosine Sensitive), a Copper Metalloenzyme? *J. Bacteriol.* **1989,** 171, (11), 6155-6160.

(83) Staub, M.; Denes, G., Purification and properties of the 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (phenylalanine sensitive) of *Escherichia coli* K12. II. Inhibition of activity of the enzyme with phenylalanine and functional group-specific reagents. *Biochim. Biophys. Acta* **1969,** 178, (3), 599-608.

(84) Staub, M.; Denes, G., Purification and properties of the 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (phenylalanine sensitive) of *Escherichia coli* K12. I. Purification of enzyme and some of its catalytic properties. *Biochim. Biophys. Acta* **1969,** 178, (3), 588-98.

(85) Simpson, R. J.; Davidson, B. E., Studies on 3-deoxy-D-*arabino*heptulosonate-7-phosphate synthetase(phe) from *Escherichia coli* K12. 1.Purification and subunit structure. *Eur. J. Biochem.* **1976,** 70, (2), 493-500.

(86) Paravicini, G.; Schmidheini, T.; Braus, G., Purification and properties of the 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (phenylalanine-inhibitable) of *Saccharomyces cerevisiae*. *Eur. J. Biochem.* **1989,** 186, (1-2), 361-6.

(87) Park, O. K.; Bauerle, R., Metal-catalyzed oxidation of phenylalanine-sensitive 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from Escherichia coli:

inactivation and destabilization by oxidation of active-site cysteines. *J. Bacteriol.* **1999,** 181, (5), 1636-1642.

(88) Parker, E. J.; Bulloch, E. M. M.; Jameson, G. B.; Abell, C., Substrate deactivation of phenylalanine senstive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate by erythrose 4-phosphate. *Biochemistry* **2001,** 40(49), 14821-14828.

(89) Stephens, C. M.; Bauerle, R. H., Essential cysteines in 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*. *J. Biol. Chem.* **1992,** 267(9), (March 25), 5762-5767.

(90) Howe, D. L.; Duewel, H. S.; Woodard, R. W., Histidine 268 in 3-Deoxy-D-*arabino*-heptulosonic acid 7-phosphate synthase plays the same role as histidine 202 in 3-deoxy-D-*manno*-octulosonic acid 8-phosphate synthase. *J. Biol. Chem.* **2000,** 275 (51), (December 22), 40258-40265.

(91) Xu, J.; Hu, C.; Shen, S.; Wang, W.; Jiang, P.; Huang, W., Requirement of the N-terminus for dimer formation of phenylalanine-sensitive 3-deoxy-D-*arabino*-heptulosonate synthase AroG of *Escherichia coli*. *J. Basic Microbiol.* **2004,** 44, (5), 400-406.

(92) Shumilin, I., A.; Bauerle, R.; Kretsinger Robert, H., The high-resolution structure of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase reveals a twist in the plane of bound phosphoenolpyruvate. *Biochemistry* **2003,** 42, (13), 3766-76.

(93) Wagner, T.; Shumilin, I. A.; Bauerle, R.; Kretsinger, R. H., Structure of 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate Synthase from *Escherichia coli*: Comparison of the Mn^{2+} 2-Phosphoglycolate and the Pb^{2+} 2-Phosphoenolpyruvate Complexes and Implications for Catalysis. *J. Mol. Biol.* **2000,** 301, (2), 389-399.

(94) Schofield, L. R.; Patchett, M. L.; Parker, E. J., Expression, purification, and characterization of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Pyrococcus furiosus*. *Protein Expr. Purif.* **2004,** 34, 17-27.

(95) Wu, J.; Howe, D. L.; Woodard, R. W., *Thermotoga maritima* 3-Deoxy-D-*arabino*-heptulosonate 7-Phosphate (DAHP) Synthase: The ancestral eubacterial dahp synthase? *J. Biol. Chem.* **2003,** 278, (30), 27525-27531.

(96) Jensen, R. A.; Nester, E. W., Regulatory enzymes of aromatic amino acid biosynthesis in *Bacillus subtilis*. 1. Purification and properties of 3-deoxy-D-

arabino-heptulosonate 7-phosphate synthetase. *J. Biol. Chem.* **1966,** 241, (14), 3365-72.

(97) Birck, M. R.; Woodard, R. W., *Aquifex aeolicus* 3-deoxy-D-*manno*-2-octulosonic acid 8-phosphate synthase:a new class of KDO 8-P synthase? *J. Mol. Evol.* **2001,** 52, 205-214.

(98) Sau Apurba, K.; Li, Z.; Anderson Karen, S., Probing the role of metal ions in the catalysis of *Helicobacter pylori* 3-deoxy-D-*manno*-octulosonate-8-phosphate synthase using a transient kinetic analysis. *J. Biol. Chem.* **2004,** 279, (16), 15787-94.

(99) Wu, J.; Patel, M.; Sundaram, A. K.; Woodard, R. W., Functional and biochemical characterization of a recombinant *Arabidopsis thaliana* 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase. *Biochem. J.* **2004,** 381, 185-193.

(100) Duewel, H. S.; Sheflyan, G. Y.; Woodard, R. W., Functional and Biochemical Characterization of a Recombinant 3-Deoxy-D-*manno*-octulosonic Acid 8-Phosphate Synthase from the Hyperthermophilic Bacterium *Aquifex aeolicus*. *Biochem. Biophys. Res. Commun.* **1999,** 263, (2), 346-351.

(101) Ray, P. H., Purification and characterization of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthetase from *Escherichia coli*. *J. Bacteriol.* **1980,** 141, 635-644.

(102) Taylor, W. P.; Sheflyan, G. Y.; Woodard, R. W., A single point mutation in 3-deoxy-D-*manno*-octulosonate-8-phosphate synthase is responsible for temperature sensitivity in a mutant strain of *Salmonella typhimurium*. *J. Biol. Chem.* **2000,** 275, (41), 32141-32146.

(103) Radaev, S.; Dastidar, P.; Patel, M.; Woodard, R. W., Structure and mechanism of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase. *J. Biol. Chem.* **2000,** 275(13), (March 31), 9476-9484.

(104) Wang, J.; Duewel, H. S.; Woodard, R. W.; Gatti, D. L., Structures of *Aquifex aeolicus* KDO8P synthase in complex with R5P and PEP, and with a bisubstrate inhibitor: role of active site water in catalysis. *Biochemistry* **2001,** 40, (51), 15676-83.

(105) Kohen, A.; Jakob, A.; Baasov, T., Mechanistic studies of 3-deoxy-D-*manno*-2-octulosonate-8-phosphate synthase from *Escherichia coli*. *Eur. J. Biochem.* **1992,** 208, 443-449.

(106) Howe, D. L.; Sundaram, A. K.; Wu, J.; Gatti, D., L; Woodard, R. W., Mechanistic insight into 3-Deoxy-D-*manno*-octulosonate-8-phosphate Synthase

and 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate Synthase Utilizing Phosphorylated Monosaccharide Analogues. *Biochemistry* **2003,** 42, 4843-4854.

(107) Ahn, M.; Pietersma Amy, L.; Schofield Linley, R.; Parker Emily, J., Mechanistic divergence of two closely related aldol-like enzyme-catalysed reactions. *Org. Biomol. Chem.* **2005,** 3, (22), 4046-9.

(108) Sheflyan, G. Y.; Sundaram, A. K.; Taylor, W. P.; Woodard, R. W., Substrate ambiguity of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* revisited. *J. Bacteriol.* **2000,** 182, (17), 5005-5008.

(109) Duewel, H. S.; Woodard, R. W., A metal bridge between two enzyme families. 3-Deoxy-D-*manno*-octulosonate-8-phosphate synthase from Aquifex aeolicus requires a divalent metal for activity. *J. Biol. Chem.* **2000,** 275, (30), 22824-22831.

(110) Shulami, S.; Yaniv, O.; Rabkin, E.; Shoham, Y.; Baasov, T., Cloning, expression, and biochemical characterization of 3-deoxy-D-*manno*-2-octulosonate-8-phosphate (KDO8P) synthase from the hyperthermophilic bacterium *Aquifex pyrophilus*. *Extremophiles : life under extreme conditions* **2003,** 7, (6), 471-81.

(111) Krosky, D. J.; Alm, R.; Berg, M.; Carmel, G.; Tummino, p. J.; Xu, B.; Yang, W., *Helicobacter pylori* 3-deoxy-D-*manno*-octulosonate-8-phosphate (KDO-8-P) synthase is a zinc-metalloenzyme. *Biochim. Biophys. Acta* **2002,** 1594, 297-306.

(112) Li, J.; Wu, J.; Fleischhacker, A. S.; Woodard, R. W., Conversion of Aquifex aeolicus 3-Deoxy-D-*manno*-octulosonate 8-Phosphate Synthase, a Metalloenzyme, into a Nonmetalloenzyme. *J. Am. Chem. Soc.* **2004,** 126, (24), 7448-7449.

(113) Shulami, S.; Furdui, C.; Adir, N.; Shoham, Y.; Anderson, K. S.; Baasov, T., A Reciprocal Single Mutation Affects the Metal Requirement of 3-Deoxy-D-*manno*-2-octulosonate-8-phosphate (KDO8P) Synthases from *Aquifex pyrophilus* and *Escherichia coli*. *J. Biol. Chem.* **2004,** 279, (43), 45110-45120.

(114) Oliynyk, Z.; Briseno-Roa, L.; Janowitz, T.; Sondergeld, P.; Fersht, A. R., Designing a metal-binding site in the scaffold of *Escherichia coli* KDO8PS. *Protein Engineering, Design & Selection* **2004,** 17, (4), 383-390.

(115) Duewel, H. S.; Radaev, S.; Wang, J.; Woodard, R. W.; Gatti, D. L., Substrate and metal complexes of 3-deoxy-D-manno-octulosonate-8-phosphate synthase from

Aquifex aeolicus at 1.9-.ANG. resolution. Implications for the condensation mechanism. *J. Biol. Chem.* **2001,** 276, (11), 8393-8402.

(116) Liang, P.-H.; Lewis, J.; Anderson, K. S., Catalytic mechanism of KDO8P Synthase: transient kinetic studies and evaluation of a putative reaction intermediate. *Biochemistry* **1998,** 37, 16390-16399.

(117) Stuart, F.; Hunter, I. S., Purification and characterization of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Streptomyces rimosus*. *Biochim. Biophys. Acta* **1993,** 1161, (2-3), 209-15.

(118) Yoo, J. C.; Hee-Jae, B.; Eun-Ha, L.; Sung-Jun, K.; Jung-Jun, L., Purification and Characteristics of 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate Synthetase from *Streptomyces caespitosus*. *Korean J. Microbiol.* **1993,** 31, (4), 340-345.

(119) Nimmo, G. A.; Coggins, J. R., The purification and molecular properties of the tryptophan-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Neurospora crassa*. *Biochem. J.* **1981,** 197, 427-436.

(120) Nimmo, G. A.; Coggins, J. R., Some kinetic properties of the tryptophan-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Neurospora crassa*. *Biochem. J.* **1981,** 199, (3), 657-65.

(121) Gorisch, H.; Lingens, F., 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate synthase of *Streptomyces aureofaciens*. II. Repression and inhibition by tryptophan and tryptophan analogues. *Biochim. Biophys. Acta* **1971,** 242, 630-636.

(122) Doong, R. L.; Jensen, R. A., Synonymy of the three apparent isoenzymes of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase in *Pisum sativum* L. with 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase and the DS-Co/DS-Mn isoenzyme pair. *New Phytol.* **1992,** 121, 165-171.

(123) Doong, R. L.; Ganson, R. J.; Jensen, R. A., Plastid-localized 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DS-Mn): the early-pathway target of sequential feedback inhibition in higher plants. *Plant, Cell Environ.* **1993,** 16, (4), 393-402.

(124) Herrmann, K. M., The shikimate pathway: early steps in the biosynthesis of aromatic compounds. *Plant Cell* **1995,** 7, (7), 907-19.

(125) Doong, R. L.; Gander, J. E.; Ganson, R. J.; Jensen, R. A., The cytosolic isoenzyme of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase in

Spinacia oleracea and other higher plants: extreme substrate ambiguity and other properties. *Physiol. Plant.* **1992,** 84, (3), 351-60.

(126) Weaver, L. M.; Pinto, J. E. B. P.; Herrmann, K. M., Expression of potato 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthase in *Escherichia coli*. *Bioorg. Med. Chem. Lett.* **1993,** 3, (7), 1421-8.

(127) Onderka, D. K.; Floss, H. G., Steric course of the chorismate synthetase reaction and the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthetase reaction. *J. Am. Chem. Soc.* **1969,** 91, (21), 5894-5896.

(128) DeLeo, A. B.; Sprinson, D. B., Mechanism of 3-deoxy-D-*arabino*heptulosonate-7-phosphate (DAHP) synthetase. *Biochem. Biophys. Res. Commun.* **1968,** 32, (5), 873-7.

(129) Floss, H. G.; Onderka, D. K.; Carroll, M., Stereochemistry of the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthetase reaction and the chorismate synthetase reaction. *J. Biol. Chem.* **1972,** 247, (3), 736-744.

(130) Sundaram, A. K.; Woodard, R. W., Probing the stereochemistry of *E. coli* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (phenylalanine-sensitive)-catalysed synthesis of KDO 8-P analogs. *J. Org. Chem.* **2000,** 65, 5891-5897.

(131) Staub, M.; Denes, G., A kinetic study of the mechanism of action of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase in *Escherichia coli* K 12. *Biochim. Biophys. Acta* **1967,** 132, 528-530.

(132) DeLeo, A. B.; Dayan, J.; Sprinson, D. B., Purification and kinetics of tyrosine-sensitive 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate synthetase from *Salmonella*. *J. Biol. Chem.* **1973,** 248, (7), 2344-2353.

(133) Lambert, J. M.; Boocock, M. R.; Coggins, J. R., The 3-dehydroquinate synthase activity of the pentafunctional arom enzyme complex of *Neurospora crassa* is Zn^{2+}-dependent. *Biochem. J.* **1985,** 226, (3), 817-29.

(134) Hedstrom, L.; Abeles, R., 3-Deoxy-D-*manno*-octulosonate-8-phosphate synthase catalyzes the C-O bond cleavage of phosphoenolpyruvate. *Biochem. Biophys. Res. Commun.* **1988,** 157, (2), 816-20.

(135) Dotson, G. D.; Nanjappan, P.; Reily, M. D.; Woodard, R. W., Stereochemistry of 3-Deoxyoctulosonate 8-phosphate synthase. *Biochemistry* **1993,** 32, 12392-12397.

(136) Asojo, O.; Friedman, J.; Adir, N.; Belakhov, V.; Shoham, Y.; Baasov, T., Crystal structures of KDOP synthase in its binary complexes with the substrate

phosphoenolpyruvate and with a mechanism-based inhibitor. *Biochemistry* **2001,** 40, (21), 6326-34.

(137) Furdui, C. M.; Sau, A. K.; Yaniv, O.; Belakhov, V.; Woodard, R. W.; Baasov, T.; Anderson, K. S., The Use of (E)- and (Z)-Phosphoenol-3-fluoropyruvate as Mechanistic Probes Reveals Significant Differences between the Active Sites of KDO8P and DAHP Synthases. *Biochemistry* **2005,** 44, (19), 7326-7335.

(138) Wang, J.; Duewel, H. S.; Stuckey, J. A.; Woodard, R. W.; Gatti, D. L., Function of His185 in *Aquifex aeolicus* 3-Deoxy-D-*manno*-octulosonate 8-Phosphate Synthase. *J. Mol. Biol.* **2002,** 324, (2), 205-214.

(139) Kim, C.-G.; Kirschning, A.; Bergon, P.; Zou, P.; Su, E.; Sauerbrei, B.; Ning, S.; Ahn, Y.; Breuer, M.; et al., Biosynthesis of 3-amino-5-hydroxybenzoic acid, the precursor of mC7N units in ansamycin antibiotics. *J. Am. Chem. Soc.* **1996,** 118, (32), 7486-7491.

(140) Guo, J.; Frost, J. W., Biosynthesis of 1-deoxy-1-*imino*-D-erythrose 4-phosphate: a defining metabolite in the aminoshikimate pathway. *J. Am. Chem. Soc.* **2002,** 124, (4), 528-9.

(141) Chen, S.; von Bamberg, D.; Hale, V.; Breuer, M.; Hardt, B.; Muller, R.; Floss, H. G.; Reynolds, K. A.; Leistner, E., Biosynthesis of ansatrienin (mycotrienin) and naphthomycin. Identification and analysis of two separate biosynthetic gene clusters in *Streptomyces collinus* Tu 1892. *Eur. J. Biochem.* **1999,** 261, (1), 98-107.

(142) Floss, H. G.; Yu, T. W., Lessons from the rifamycin biosynthetic gene cluster. *Curr. Opin. Chem. Biol.* **1999,** 3, (5), 592-7.

(143) Guo, J.; Frost, J. W., Synthesis of Aminoshikimic Acid. *Org. Letters* **2004,** 6, (10), 1585-1588.

(144) August, P. R.; Tang, L.; Yoon, Y. J.; Ning, S.; Muller, R.; Yu, T.-W.; Taylor, M.; Hoffmann, D.; Kim, C.-G.; Zhang, X.; Hutchinson, C. R.; Floss, H. G., Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of Amycolatopsis mediterranei S699. *Chem. Biol.* **1998,** 5, (2), 69-79.

(145) Guillemin, K. J.; Salama, N. R., *Helicobacter pylori* Functional Genomics. *Methods Microbiol* **2002,** 33, 291-315.

(146) Alm, R. A.; Ling, L.-S. L.; Moir, D. T.; King, B. L.; Brown, E. D.; Doig, P. C.; Smith, D. R.; Noonan, B.; Guild, B. C.; DeJonge, B. L.; Carmel, G.; Tummino, P.

J.; Caruso, A.; Uria-Nickelsen, M.; Mills, D. M.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S. D.; Jiang, Q.; Taylor, D. E.; Vovis, G. F.; Trust, T. J., Genomic sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature (London)* **1999,** 397, (6715), 176-180.

(147) Murphy, M. F.; Katz, E., Regulatory control of 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate synthetase in *Streptomyces antibioticus*. *Can. J. Microbiol.* **1980,** 26, 874-880.

(148) Ip, K.-M.; Doy, C. H., The kinetics of a purified form of 3-deoxy-D-*arabino* heptulosonate-7-phosphate synthase (tryptophan) from *Neurospora crassa*. *Eur. J. Biochem.* **1979,** 98, (2), 431-40.

(149) Kam-ming, I.; Doy, C. H., The kinetics of a purified form of 3-deoxy-D-*arabino* heptulosonate-7-phosphate synthase (tryptophan) from *Neurospora crassa*. *Eur. J. Biochem.* **1979,** 98, 431-440.

(150) Baneyx, F.; Palumbo, J. L., Improving heterologous protein folding via molecular chaperone and foldase co-expression. *Methods Mol. Biol.* **2003,** 205, (E. coli Gene Expression Protocols), 171-197.

(151) Houry, W. A.; Frishman, D.; Eckerskorn, C.; Lottspeich, F.; Hartl, F. U., Identification of *in vivo* substrates of the chaperonin GroEL. *Nature* **1999,** 402, (November), 147-154.

(152) Cliff, M. J.; Limpkin, C.; Cameron, A.; Burston, S. G.; Clarke, A. R., Elucidation of Steps in the Capture of a Protein Substrate for Efficient Encapsulation by GroE. *J. Biol. Chem.* **2006,** 281, (30), 21266-21275.

(153) Martin, J.; Mayhew, M.; Langer, T.; Hartl, F. U., The reaction cycle of GroEL and GroES in chaperonin-assisted protein folding. *Nature* **1993,** 366, (6452), 228-33.

(154) Ashiuchi, M.; Yoshimura, T.; Kitamura, T.; Kawata, Y.; Nagai, J.; Gorlatov, S.; Esaki, N.; Soda, K., In vivo effect of GroESL on the folding of glutamate racemase of *Escherichia coli*. *J. Biochem. (Tokyo).* **1995,** 117, (3), 495-8.

(155) Yanase, H.; Moriya, K.; Mukai, N.; Kawata, Y.; Okamoto, K.; Kato, N., Effects of GroESL coexpression on the folding of nicotinoprotein formaldehyde dismutase from *Pseudomonas putida* F61. *Biosci. Biotechnol. Biochem.* **2002,** 66, (1), 85-91.

(156) Widersten, M., Heterologous expression in *Escherichia coli* of soluble active-site random mutants of haloalkane dehalogenase from *Xanthobacter autotrophicus*

Gj10 by coexpression of molecular chaperonins GroEL/ES. *Protein Expr. Purif.* **1998,** 13, (3), 389-395.

(157) Kim David, E.; Chivian, D.; Baker, D., Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* **2004,** 32, (Web Server issue), W526-31.

(158) Gough, J.; Karplus, K.; Hughey, R.; Chothia, C., Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J. Mol. Biol.* **2001,** 313, (4), 903-919.

(159) Palmer, T., *ENZYMES Biochemistry, Biotechnology, Clinical Chemistry.* Horwood Publishing Limited: England, 2001; p 402.

(160) Cleland, W. W., The kinetics of enzyme-catalyzed reactions with two or more substrates or products. III. Prediction of initial velocity and inhibition patterns by inspection. *Biochim. Biophys. Acta* **1963,** 67, 188-96.

(161) Parker, E. J. Mechanistic Studies on Shikimate Pathway Enzymes. PhD, University of Cambridge, Cambridge, 1996.

(162) Yoo, J. C.; Sung-Jun, K.; Jung-Jun, L., Inhibitors of 3-Deoxy-D-arabino-heptulosonate-7-phosphate Synthetase from *Streptomyces caespitosus*. *Korean J. Microbiol.* **1993,** 31, (6), 550-553.

(163) Williamson, R. M.; Pietersma, A. L.; Jameson, G. B.; Parker, E. J., Stereospecific deuteration of 2-deoxyerythrose 4-phosphate using 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *Bioorg. Med. Chem. Lett.* **2005,** 15, (9), 2339-2342.

(164) Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson Toby, J.; Higgins Desmond, G.; Thompson Julie, D., Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **2003,** 31, (13), 3497-500.

(165) Whitaker, R. J.; Fiske, M. J.; Jensen, R. A., *Pseudomonas aeruginosa* possesses two novel regulatory isozymes of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *J. Biol. Chem.* **1982,** 257, (21), 12789-94.

(166) Stokstad, E., Infectious disease. Drug-resistant TB on the rise. *Science* **2000,** 287, (5462), 2391.

(167) Rattan, A.; Kalia, A.; Ahmad, N., Multidrug-resistant *Mycobacterium tuberculosis*: molecular perspectives. *Emerg. Infect. Dis.* **1998,** 4, (2), 195-209.

(168) Parrish, N. M.; Dick, J. D.; Bishai, W. R., Mechanisms of latency in *Mycobacterium tuberculosis*. *Trends Microbiol.* **1998,** 6, (3), 107-12.

(169) Parish, t.; Stoker, N. G., The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*. *Microbiology* **2002,** 148, 3069-3077.

(170) Goloubinoff, P.; Gatenby, A. A.; Lorimer, G. H., GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli*. *Nature* **1989,** 337, 44-47.

(171) Burns, J. A.; Butler, J. C.; Moran, J.; Whitesides, G. M., Selective Reduction of Disulfides by Tris (2-carboxyethyl) phosphine. *J. Org. Chem.* **1991,** 56, 2648-2650.

(172) Hendrickson, W. A.; Horton, J. R.; LeMaster, D. M., Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* **1990,** 9, (5), 1665-72.

(173) Otwinowski, Z.; Minor, W., Processing of x-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **1997,** 276, (Macromolecular Crystallography, Part A), 307-326.

(174) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M., PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **1993,** 26, (2), 283-91.

(175) Uson, I.; Sheldrick, G. M., Advances in direct methods for protein crystallography. *Curr. Opin. Struct. Biol.* **1999,** 9, (5), 643-648.

(176) Sheldrick, G. M., Macromolecular phasing with SHELXE. *Zeitschrift fuer Kristallogr.* **2002,** 217, (12), 644-650.

(177) Terwilliger, T. C., Maximum-likelihood density modification. *Acta Crystallog. sect. D* **2000,** D56, (8), 965-972.

(178) Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallog. sect. D* **2004,** D60, (12, Pt. 1), 2126-2132.

(179) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D* **1997,** D53, (3), 240-255.

(180) Sassetti, C. M.; Boyd, D. H.; Rubin, E. J., Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* **2003,** 48, (1), 77-84.

(181) Webby, C. J.; Baker, H. M.; Lott, J. S.; Baker, E. N.; Parker, E. J., The Structure of 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate Synthase from *Mycobacterium*

tuberculosis Reveals a Common Catalytic Scaffold and Ancestry for Type I and Type II Enzymes. *J. Mol. Biol.* **2005,** 354, (4), 927-939.

(182) Kwak, J. H.; Hong, K. W.; Lee, S. H.; Hong, J. H.; Lee, S. Y., Identification of amino acid residues involved in feedback inhibition of the anthranilate synthase in *Escherichia coli*. *J. Biochem. Mol. Biol.* **1999,** 32, (1), 20-24.

(183) Collaborative Computational Project, Number 4, The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D* **1994,** 50, (Pt 5), 760-3.

(184) Oldiges, M.; Kunze, M.; Degenring, D.; Sprenger, G. A.; Takors, R., Stimulation, Monitoring, and Analysis of Pathway Dynamics by Metabolic Profiling in the Aromatic Amino Acid Pathway. *Biotechnology* **2004,** 20, 1623-1633.

(185) Williams, J. F.; Blackmore, P. F.; Duke, C. C.; MacLeod, J. K., Fact, uncertainty and speculation concerning the biochemistry of D-erythrose-4-phosphate and its metabolic roles. *Int. J. Biochem.* **1980,** 12, (3), 339-44.

(186) Aziz, Q. H.; Partridge, C. J.; Munsey, T. S.; Sivaprasadarao, A., Depolarization Induces Intersubunit Cross-linking in a S4 Cysteine Mutant of the Shaker Potassium Channel. *J. Biol. Chem.* **2002,** 277, (45), 42719-42725.

(187) Messens, J.; Martins, J. C.; Zegers, I.; Van Belle, K.; Brosens, E.; Wyns, L., Purification of an oxidation-sensitive enzyme, pI258 arsenate reductase from *Staphylococcus aureus*. *J. Chromatogr. B. Biomed. Sci. App.* **2003,** 790, (1-2), 217-227.

(188) Schofield, L.

(189) Sundaram, A. K.; Howe, D. L.; Sheflyan, G. Y.; Woodard, R. W., Probing the potential metal binding site in Escherichia coli 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (phenylalanine-sensitive). *FEBS Lett.* **1998,** 441, (2), 195-199.

(190) Kaakoush Nadeem, O.; Mendz George, L., *Helicobacter pylori* disulphide reductases: role in metronidazole reduction. *FEMS Immunol. Med. Microbiol.* **2005,** 44, (2), 137-42.

(191) Inoue, H.; Nojima, H.; Okayama, H., High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **1990,** 96, 23-28.

(192) Dower, W. J.; Miller, J. F.; Ragsdale, C. W., High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* **1988,** 16, (13), 6127-45.

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

(194) Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976,** 72, 248-254.

(195) Jones, S.; Thornton, J. M., Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** 93, (1), 13-20.

(196) Baker, E. N.; Hubbard, R. E., Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* **1984,** 44, (2), 97-179.

(197) Parker, E. J.; Coggins, J. R.; Abell, C., Derailing Dehydroquinate Synthase by Introducing a Stabilizing Stereoelectronic Effect in a Reaction Intermediate. *J. Org. Chem.* **1997,** 62, (24), 8582-8585.

(198) Stubbe, J. A.; Kenyon, G. L., Analogs of phosphoenolpyruvate. Substrate specificities of Enolase and Pyruvate Kinase from rabbit Muscle. *Biochemistry* **1972,** 11, (3), 338-345.

(199) Rhodes, G., *Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models.* Academic Press, Inc.: California, 1993; p 202.

(200) Dauter, Z., New approaches to high-throughput phasing. *Curr. Opin. Struct. Biol.* **2002,** 12, (5), 674-678.

APPENDIX ONE

Sequences included in alignment: *Arabidopsis thaliana* (GI:15236093, GI:15234217, GI:15219878), *A. mediteranni* (GI:7839575, GI:2792322), *Actinosynnema pretiosum* subsp. *auranticum* (GI: 3046984), *A. tumefaciens* (GI: 17740040), *A. nidulans* (GI:40745637), *Bartonella henselae* (GI:49475432), *Bartonella quintana* (GI), *Burkholderia cepacia* (x 2; GI:NA), *Bradyrhizobium japonicum* (GI), *Brevibacterium linens* (GI:NA), *Brucella melitensis* (GI:17982929), *Brucella suis* (GI:23501893), *Caulobacter crescentus* (GI:16126539), *C. jejuni* NCTC11168 (GI:15792065), *C. jejuni* RM1221 (GI:NA), *Clavibacter michiganensis* (GI:NA), *C. diphtheriae* (GI:38234191), *Corynebacterium efficiens* (GI:25028629), *Corynebacterium glutamicum* (GI:41326357), *Cryptococcus neoformans* (GI:NA), *C. merolae* (CMQ1206), *Desulfuromonas acetoxidans* (GI:NA), *Gibberella zeae* (GI:46109208), *Helicobacter hepaticus* (GI:32267056), *Helicobacter mustelae* (GI:NA), *Helicobacter pylori* 26695 (GI:15644764), *Hyphomonas neptunium* (GI:NA), *Kineococcus radiotolerans* (GI:46365623), *Legionella pneumophilia* (GI:NA), *Magnaporthe grisea* (GI:38102901), *Magnetospirillum magnetotacticum* (GI:NA), *Mesorhizobium loti* (GI:13470978), *Mycobacterium avium* (GI:NA), *Mycobacterium avium* subsp. paratuberculosis (GI:41408014), *Mycobacterium leprae* (GI:15827417), *M. tuberculosis* (GI:2911098), *M. bovis* (GI:31618948), *Mycobacterium microti* (GI:NA), *Mycobacterium marinus* (GI:NA), *Myxococcus xanthus* (GI:NA), *N. crassa* (GI:28925693), *N. aromaticivorans* (GI: 48851120), *Oryza sativa* (cDNA is GI:29157098, cDNA is GI:29155735), *Propionibacterium acnes* (GI:50842214), *Proteus mirabilis* (NCBI patents database GI:40080529), *P. aeruginosa* (GI:15598039), *P. putida* (GI:26988596), *Pseudomonas fluorescens* (GI:2492973, GI:48728741), *Pseudomonas syringae* (GI:46187573, GI:28868978, GI:NA), *Rhodobacter sphaeroides* (GI:NA), *Rhodopseudomonas palustris* (GI:39648904), *Rhodospirillum rubrum* (GI:NA), *Rhizobium leguminosarum* (GI:NA), *Silicibacter pomeroyi* (GI:NA), *Solanum tuberosum* (GI: 584777), *S. coelicolor* (GI:4490619, GI:21220593), *S. collinus* (GI:4884833, GI:4884848), *Streptomyces* sp. (GI:2687633), *Streptomyces venezuelae* (GI:47846869) *Streptomyces* sp. HK803 (GI:35186988), *Sinorhizobium meliloti* 1021 (GI: 15965606), *S. aurantiaca* (GI:13924499, GI:11127900), *Stenotrophomonas maltophilia* (GI:NA) *S. avermitilis* (GI:29609748), *Streptomyces hygroscopicus* (GI:2624953), *T. gondii* (GI:33356632), *Tropheryma whipplei* (GI:28493182, GI:28572704), *Ustilago maydis* (GI:46097439), *Wolinella succinogenes* (GI:34557796), *X. campestris* (GI:13172310, GI:21230391), *Xanthomonas axonopodis* (GI:21107137), *Yarrowia lipolytica* (GI:50547503).

APPENDIX TWO

Below are residues involved in the dimer interface as determined by the Protein-Protein Interaction Server (http://www.biochem.ucl.ac.uk/bsm/PP/server). Displayed are residues from subunit A and their interaction with subunit B. Contributions from residues in subunit B are not shown but are very similar.

Residue	Interface accessible surface area	% Interface accessible surface area
Gly-1	1.70	0.10
Ala0	11.94	0.67
Asn2	1.78	0.10
Trp3	91.39	5.14
Thr4	10.69	2.29
Val5	94.18	5.29
Asp6	43.58	2.45
Ille7	122.87	6.91
Pro8	68.64	3.86
Ille9	53.47	3.00
Asp10	64.03	3.60
Gln11	51.57	2.90
Leu12	40.08	2.25
Pro13	1.57	0.09
Leu15	50.56	2.84
Gln44	23.47	1.32
Ala47	31.16	1.75
Val51	12.81	0.72
Ser54	20.22	1.14
Val55	2.06	0.12
Pro56	95.32	5.36
Pro57	47.21	2.65
Val58	11.49	0.65
Val60	51.57	2.90

Ser62	23.59	1.33
Glu63	13.11	0.74
Asn94	19.40	1.09
Thr95	27.98	1.57
Glu96	33.95	1.91
Ille99	4.54	0.26
Asp165	25.01	1.41
Pro166	5.54	0.31
Ser167	43.05	2.42
Vall70	5.28	0.30
Arg171	87.15	4.90
Tyr173	9.85	0.55
Ala174	27.52	1.55
Ser177	33.91	1.91
Ala178	58.27	3.27
Asn181	105.60	5.94
Leu182	36.34	2.04
Arg184	62.09	3.49
Ala185	56.59	3.18
Leu186	9.40	0.53
Ser188	10.90	0.61
Ser189	29.10	1.64
Gly190	3.10	0.17
Thr240	5.83	0.33
Arg260	1.89	0.11
Asp263	6.94	0.39

APPENDIX THREE

Below are residues involved in the tetramer interface as determined by the Protein-Protein Interaction Server (http://www.biochem.ucl.ac.uk/bsm/PP/server). Displayed are residues from subunit B and their interaction with subunit D. Contributions from residues in subunit D are not shown. Subunit D is generated using crystallographic symmetry operations.

Residue	Interface accessible surface area	% Interface accessible surface area
Val111	8.59	0.86
Thr114	43.33	4.32
Tyr115	69.47	6.93
Gly116	1.72	0.17
Ser118	78.95	7.88
Pro120	36.71	3.66
Leu194	30.46	3.04
Glu220	36.57	3.65
Ile221	4.39	0.44
Arg223	100.86	10.06
Gly224	30.88	3.08
Leu225	19.97	1.99
Arg226	28.33	2.83
Phe227	142.94	14.26
Ser229	2.68	0.27
Ala230	82.56	8.24
Cys231	67.88	6.77
Gly232	12.52	1.25
Val233	48.26	4.82
Asn237	40.54	4.04
Leu238	34.52	3.44
Gln239	33.72	3.36
Arg461	46.35	4.63