Identification of the evolutionary divergence in DHDPS and DHDPR

A thesis submitted in partial
fulfilment of the requirements
for the degree of

Master of Science

at the

School of Biological Sciences,
University of Canterbury

by

Hamish Sean Cleland

2017

Table of contents

Abstract

DHDPS and DHDPR are the first two committed steps in the DAP pathway: a pathway responsible for the biosynthesis of lysine. It is only present in bacteria and plants making an important biological target. While DHDPS exists in a homotetrameric "dimer of dimers" formation in both bacteria and plants, the arrangement of monomers is different. In bacteria, the dimers face toward each other in a front to front arrangement. However, in plants, the orientation of the dimers is flipped into a back to back arrangement. An evolutionary difference is also observed in DHDPR. In bacteria, the protein exists in a homotetrameric conformation whereas in plants it has been shown to exist in a dimeric conformation. The exact reason for these differences in structure remain unclear but it is thought to due to evolutionary changes between the two organism types.

In this study, a lycophyte DHDPS from *Selaginella moellendorffii* was found to exist in a substrate mediated equilibrium between dimer and tetramer, with no ligands bound. When the substrate pyruvate is bound to the enzyme, the equilibrium shifted to the tetrameric species. However, in the presence of the allosteric inhibitor lysine, the equilibrium was found to shift to a dimeric species in solution. This equilibrium could exist as a "missing link" in the evolution of the plant type quaternary structure of the DHDPS enzyme.

Another subject of investigation was the characterisation of red, green and brown algal DHDPRs. The quaternary state of these species was found to be dimeric in nature. This corresponds to the proposed evolutionary lineage in which most of these species exist after the plant type species in the lineage. The exception to this is the green alga *Chlamydomonas reinhardtii* DHDPR which exists in an equilibrium between tetramer and dimer. As this organism lies in the evolutionary lineage between bacterial and plant forms, it is possible that this organisms DHDPR exists as the "divergence point" between these two species. *C. reinhardtii* DHDPR also contains a disulfide-dependent dimer interface. In the presence of reducing agent, the enzyme exists in an exclusively dimeric state. These evolutionary lineages could be applied to other enzyme evolution systems from the DAP pathway and beyond.

Acknowledgements

Firstly, I would like to thank my supervisor; Grant Pearce. Your tutelage and direction helped me become not only a better, more independent scientist but also a more rounded individual. Thank you for your support and enthusiasm.

I would also like everyone who was a part of room 630 over the year: Cameron MacDonald, William Finnis, Serena Watkin, David Coombes, Rudy Bundela and James Davies. Many a time was spent discussing subjects ranging from Cricket to Fruit Bursts and everything in between. My time as a student would certainly not be as memorable without your presence.

A special thanks also goes out to the Dobson lab for their support throughout. Along with the technical assistance, I will never forget the meetings and conversations throughout the year. A big thank you also goes to Jackie and Rayleen for keeping the lab in tip top shape and teaching me the ins and outs of the laboratory protocols. I would also like to thank the School of Biological Sciences and its residents for providing such excellent facilities and assistance.

This research would also not be possible without my friends and family. After a long day in the lab, your discussions and activities were greatly appreciated. Your support and encouragement throughout the year encouraged me to carry on and push through when things got tough.

I would also like to thank snack foods such as popcorn and pretzels for filling in the gaps between meals during a busy day in lab.

Finally, I would like to thank my Mum and Dad for supporting me throughout this endeavour. Without your unwavering confidence and reassurance during the year, this project would not be possible.

Abbreviations

(S)-ASA	(S)-aspartate β-semialdehyde
2MHI	2-methylheptyl isonicotinate
AAA	α-aminoadipate
Ala	Alanine
Å	Angstrom
Asd	Aspartate semialdehyde dehydrogenase
AU	Astronomical unit
AUC	Analytical Ultracentrifugation
BLAST	Basic Local Alignment Search Tool
Chi^2	Chi-squared test
Da	Daltons
DAP	Diaminopimelate
DapC	Succinyl-DAP aminotransferase
DapD	THPA succinylase
DapE	Succinyl-DAP desuccinylase
DapF	DAP epimerase
DapL	Diaminopimelate aminotransferase
Ddh	DAP dehydrogenase
dH_2O	Distilled H_2O
DHDP	dihydrodipicolinate
DHDPR	Dihydrodipicolinate reductase
DHDPS	Dihydrodipicolinate synthase
DSF	Differential scanning fluorimetry
Glu	Glutamic Acid
Gly	Glycine
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
His	Histidine
HTPA	(4S)-hydroxy-2,3,4,5-tetrahydro-(2S)-dipicolinic acid
LB	Luria Broth
LDS	Lithium dodecyl sulphate
Leu	Leucine

Lys	Lysine
LysA	*Meso*-diaminopimelate decarboxylase
LysC	Aspartokinase
LysE	Lysine exporter superfamily
MES	2-(N-morpholino)ethanesulfonic acid
MUSCLE	Multiple sequence alignment
MW	Molecular weight
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
ncbi	National Center for Biotechnology Information
OD	Optical Density
PCR	Polymerase chain reaction
PDC	2,6-pyridinedicarboxylate
Phe	Phenylalanine
Pro	Proline
SAXS	Small angle X-ray scattering
SDS PAGE	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SEC-MALS	Size exclusion chromatography - multi angle light scattering
Ser	Serine
SMO	*Selaginella moellendorffii*
TCEP	Tris(2-carboxyethyl)phosphine
THDP	L-2,3,4,5-tetrahydrodipicolinate
Thr	Threonine
TIM	Triosephosphate Isomerase
Tris	Tris(hydroxymethyl)aminomethane
Tyr	Tyrosine
UniProt	Universal protein resource
Val	Valine

Table of Figures

List of Tables

Chapter 1 – Introduction

1.1 Overview

The diaminopimelate (DAP) pathway is responsible for the biosynthesis of lysine in many lysine producing organisms. The first two committed steps in the pathway are catalysed by dihydrodipicolinate synthase (DHDPS) and dihydrodipicolinate reductase (DHDPR). Both proteins exhibit evolutionary differences dependent on the organism group. As lysine is an essential amino acid, the pathway is an important biological target which has many industrial and biological applications.

1.2 DAP Pathway

The diaminopimelate (DAP) pathway is responsible for biosynthesis of lysine in organisms such as bacteria, plants and lower fungi (Figure 1.1). Euglenoids and higher fungi use an alternate pathway known as the α-aminoadipate (AAA) pathway (Zabriskie and Jackson, 2000). AAA-like pathways have also been found in *Thermus thermophilus* and *Pyrococcus horikoshii*, a bacterium and an archaeon (Kobashi et al., 1999, Kosuge and Hoshino, 1998). Animals do not contain either pathway and are not able to synthesise their own lysine meaning that it must be obtained through dietary sources.

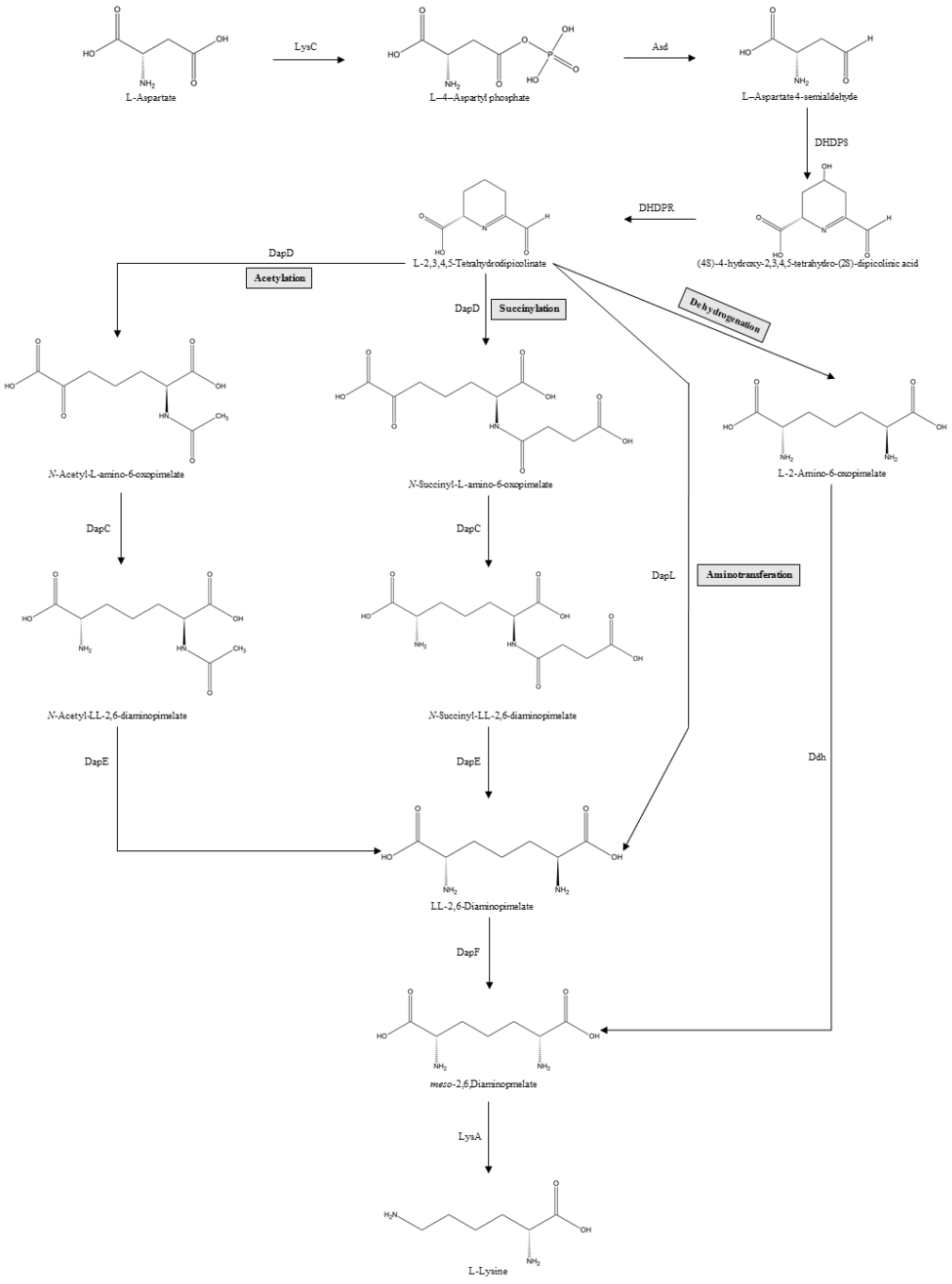


Figure 1.1: The DAP pathway with its four sub-branches (adapted from (Liu et al., 2010)

At the beginning of the pathway, aspartate is phosphorylated by aspartokinase (LysC) to form L-4-aspartyl phosphate. This is reduced to form aspartate β-semialdehyde (S)-ASA by aspartate semialdehyde dehydrogenase (Asd). These first two steps are also shared between methionine and threonine synthesis. DHDPS kicks off the first lysine specific step with an aldol condensation between (S)-ASA and pyruvate forming 2, 3 dihydrodipicolinate (HTPA). DHDPR follows catalysing HTPA reduction to 2,3,4,5-tetrahydrodipicolinate (THPA) using NADPH or NADH as a co-factor.

At this point, the pathway branches off into four sub-pathways depending on the organism type. Succinylation is one branch that is found in proteobacteria and *Mycobacterium tuberculosis* (Usha et al., 2016, Fuchs et al., 2000). This involves the succinylation of THPA in a three-step process that adds two nitrogen groups to form L,L-2,6-Diaminopimelate, a lysine precursor which functions as the convergence point for 3 sub-pathways.

Acetylation is another variation of the pathway. Using the same enzymes and steps as the succinylation pathway, this uses acetyl intermediates rather than succinyl ones. This pathway appears to be limited to certain *Bacillus* species (Sundharadas and Gilvarg, 1967). The third variation is the aminotransferase branch which directly converts THPA to the DAP precursor L,L-2,6-Diaminopimelate using diaminopimelate aminotransferase (DapL). This has been found in Cyanobacteria, Chlamydia, Methanococci, some archaea and Arabidopsis plants (Graham and Huse, 2008, Hudson et al., 2006, Liu et al., 2010, McCoy et al., 2006). In all three of these variations, DAP epimerase (DapF) catalyses the conversion of L,L-2,6-Diaminopimelate to give *meso*-2,6-Diaminopimelate, a lysine precursor.

The dehydrogenase variation directly forms *meso*-2,6-Diaminopimelate with NADPH and NH_4^+ cofactors catalysed by diaminopimelate dehydrogenase (Ddh). This is found in selected *Bacillus* species and also in *Corynebacterium glutamicum* (Misono et al., 1979). In all variations of the pathway, *meso*-2,6-Diaminopimelate is converted to lysine by *meso*-diaminopimelate decarboxylase (LysA). While most organisms only have one sub-pathway present, some contain more than one such as *C. glutamicum* (Schrumpf et al., 1991).

Production of lysine is regulated by DHDPS feedback inhibition among other factors. Lysine can bind to an allosteric site on the DHDPS enzyme which causes inhibition (Hermann et al., 1972). This effect varies between organisms. Plant enzymes are typically much more susceptible to lysine inhibition than bacterial species (Griffin et al., 2012). In terms of bacteria, most Gram-positive species are lysine inhibited whereas a large proportion of Gram-negative species are not (Soares da Costa et al., 2016). The exact physiological reason for these differing levels of regulation between species is unclear, but environmental factors and carbon availability may require a tighter regulation in these species.

Another form of regulation is feedback regulation via aspartokinase (Rodionov et al., 2003). In *E. coli*, high concentrations of lysine can repress transcription of the enzyme which functions as an alternative form of regulation in feedback insensitive DHDPS such as many Gram-negative bacteria.

1.3 Importance of Lysine

1.3.1 Industrial Production

In 2011, the global amino acid market was valued at US$3.5 billion and predicted to rise to US$5.9 billion by 2018 (Research, 2015). To meet this growing demand, industrial production is of critical importance. In a three-step process, a suitable sugar source such as molasses or sucrose is fermented with a lysine excreting bacteria for 3 to 5 days. The resulting broth is extracted and converted into a lysine powder through spray drying or crystallisation. This lysine can then be directly added to supplement animal feeds and other industrial uses.

The soil bacterium *Corynebacterium glutamicum* is the main organism used to produce industrial lysine. Its safety, handling and viability in industrial conditions make it ideal for this purpose (Eggeling and Bott, 2015). After the discovery in the 1950's that it secreted amino acids, random mutagenesis and strain selection were used to create mutant bacteria with DHDPS that exhibited reduced lysine feedback inhibition (Nakayama et al., 1978). These mutants were also found to have point

mutations in their aspartokinase gene which resulted in significantly decreased feedback inhibition (Thierbach et al., 1990). Many also had decreased threonine production which positively affected lysine production. Unfortunately, random mutagenesis of genetic information from other parts of the organism led to limited viability and decreased stress tolerance. This led to the controversial creation of genetically modified organisms with increased lysine content.

1.3.2 Importance in Nature

As an amino acid, lysine is important for the synthesis of proteins and other biomolecules. In bacteria, *meso*-2,6-Diaminopimelate is a major component of peptidoglycan, a vital component of the bacterial cell wall in most bacteria. This compound is responsible for polymeric cross links which give the cell wall both strength and shape. This therefore makes the DAP pathway an important antibiotic target.

Lysine produced by the DAP pathway is also significant in plants. While the pathway has been used as a herbicidal target, most investigation into the pathway in plants has been in search of higher lysine yields in crops (Watanabe et al., 2008, Dobson et al., 2011). As lysine and methionine are the limiting amino acids in cereal and legume crops, this limits their nutritional values to 50 - 75% of regular diets containing balanced levels of amino acids (Galili et al., 2005, Galili and Amir, 2013). The use of higher lysine producing strains of such crops or supplementation from other sources may prove beneficial in providing more balanced diets.

The recommended dietary intake in humans ranges between 30 and 64 mg/kg body weight/day for adults and infants, respectively. Typical western dietary intakes range between 40 - 180 mg/kg body weight/day with an upper limit of 300 - 400 mg/kg body weight/day (Tome and Bos, 2007). However, many developing countries utilise high cereal diets which do not contain sufficient lysine (Millward and Jackson, 2004). Deficiency of lysine and other essential amino acids have been linked to lowered disease resistance, decreased protein levels and growth development problems in children. Lysine also has more specific uses in humans such as the production of carnitine; an amino acid derivative involved in lipid metabolism (Vaz and Wanders, 2002). Lysine supplementation among other amino acids is important among

individuals that have special dietary requirements such as athletes, vegetarians and burn patients.

Lysine is also important in the agriculture industry. It is the limiting amino acid in cattle and pigs while being the second limiting amino acid in poultry after cysteine and methionine (Richardson and Hatfield, 1978). Nursery pigs are required to ingest approximately 19 g of lysine per kg of gain (Mike Tokach, 2013). White egg laying hens aged between 0 and 6 weeks require 0.85 g of lysine each per day (Manuals, 2016). This makes the supplementation of animal feeds with industrially produced lysine important to ensure health and growth of livestock.

1.4 Applications and Aims

1.4.1 Applications

With knowledge of lysine regulation mechanisms, synthesis of a genetically engineered organisms with DAP pathway modifications produced higher amounts of lysine. This is normally done through insertion of feedback insensitive DAP pathway enzymes such as DHDPS from feedback insensitive *C. glutamicum* strains have led to creation of soy, canola and maize variants with increased lysine content (Falco et al., 1995, Zhu et al., 2007). Another target is aspartokinase regulation. A point mutation at a specific position in aspartokinase was found to increase lysine production (Xu et al., 2016). The lysine exporter protein LysE is also targeted to increase lysine production (Gunji and Yasueda, 2006). This protein transports lysine out of the cell and is also thought to serve as a fail-safe against excessive lysine levels. Increased production of this protein is thought to lead to an increased efflux of lysine from the organism. The introduction of genetic modifications to commercially available strains of organisms, such as maize, remains controversial as the creation and use of genetically modification organisms is a highly debated ethical issue.

While there is ongoing investigation into industrial production in bacteria and plants as industrial sources of lysine, there are other possible organisms whose DAP pathway remain relatively unstudied. One such example is the use of algae as a possible industrial source of lysine. Industrial algae use is a recent avenue of

investigation with applications ranging from biofuel to hydrogen production, as well as environmental applications. This may prove to be a possible avenue of investigation into the industrial production of lysine as well as a target for algicides.

The DAP pathway has been proposed as an antibiotic or herbicidal target due to its absence in humans. Enzymes such as DHDPS and DapL have been investigated as possible targets in both plants and algae (Dobson et al., 2011). As these enzymes catalyse the first and last steps of the pathway, these enzymes are the most characterised in the pathway. They also catalyse steps that are present in all sub-branches of the pathway. However, there are many other enzymes in the pathway which could present themselves as attractive targets, such as DHDPR. Further characterisation of DHDPR could lead to a potential biological target.

1.4.2 Aims

While DHDPS has been characterised extensively in bacteria and plants, it remains relatively unstudied in lycophytes. One of the aims of this study is to purify and characterise DHDPS from the lycophyte *Selaginella moellendorffii*. This may give an insight into the differences in quaternary structure between bacteria and plants and may serve as a probe into the evolutionary mechanism of the protein.

DHDPR on the other hand is considerably less characterised. Only bacterial DHDPR crystal structures exist so far with only one plant DHDPR characterised. The other aim of this study is to characterise red, green and brown algal DHDPR enzymes. Characterisation of these species could lead to industrial or algicidal applications in these organisms. This study will also provide possible insight into their quaternary structure and may characterise a “divergence point” between bacterial and plant type quaternary structures, which may serve as a probe into the evolution of the protein.

Chapter 2 – Methods and Materials

2.1 Materials

2.1.1 Chemicals

All chemicals used were purchased from Lab Supply, Sigma Aldrich and Thermo Fisher. Crystal screens were purchased from molecular dimensions.

2.1.2 Equipment

An Äkta FPLC was used for all protein purifications. Vivaspin 2 10,000 Da concentrators were used to concentrate proteins prior to characterisation.

2.2 Purification

2.2.1 Plasmid Selection

All plasmids were purchased from Epoch Biolabs. Sequences were found using the UniProt protein sequence database. Algal plasmid sequences were selected through an NCBI BLAST search in comparison with either *E. coli* or *A. thaliana* DHDPR. Transformation of these plasmids into BL21 *E. coli* competent cells was performed using standard lab protocols prior to the start of the project.

2.2.2 Protein Purification

Transformed cells containing the plasmid were inoculated into 5-10 ml of sterile LB and incubated overnight. The culture (Table 2.1), antibiotics (Table 2.2) and remaining media components (Table 2.3) were added to 400 – 800 ml of autoclaved M9ZB media leaving about 1 ml remaining in case of negative result. The flasks were then incubated at 37 ºC spinning at 180 - 220 RPM until the optical density at 600 nm (OD_{600}) reached ~0.6. This required 3 - 5 hours depending on the pre-culture and media volumes. These cultures were incubated at 26 °C and induced with a 1/1000 dilution of 0.21 mol/L (50 mg/ml) IPTG. Cultures were left at 26°C spinning at 180-220 RPM overnight.

Table 2.1: The components of M9ZB media.

Component	**Amount (g/L)**
NH_4Cl	1
KH_2PO_4	3
Na_2HPO_4	6
Bacto-tryptone	10
NaCl	5
Yeast Extract	5

Table 2.2: The initial concentrations of antibiotics to be added to the M9ZB media.

Antibiotic	**Concentration (mg/ml)**
Chloramphenicol	20
Kanamycin	30
Ampicillin	100

Table 2.3: The additives to the M9ZB media after autoclaving.

Component	**Concentration (mol/L)**	**Amount (Dilution)**
Glucose	2.22	1/100
Antibiotic	See Table 2.2	1/1000
$MgCl_2$	0.4	1/1000

Following incubation, the cultures were centrifuged at 8,000 RPM for 5 minutes. The resulting cell pellet was then resuspended in 15 ml of His Tag Buffer A (Table 2.4) through vigorous shaking. Resuspended cells were lysed on ice for 20 minutes using sonication at a pulse cycle of 0.5 seconds and an amplitude of 70%.

Table 2.4: The components of the two His Tag Buffers used during His Tag chromatography.

Component	**His Tag Buffer A (mM)**	**His Tag Buffer B (mM)**
Na_2HPO_4	50	50
Imidizole	30	300
NaCl	500	500
pH	8.0	8.0

The cell lysate was centrifuged again at 16,000 RPM for 15 minutes to separate debris from the lysate. A Ge Health Care Life Sciences HisTrap FF column was prepared by flowing 15-30 ml (3-6 column volumes) of His Tag Buffer B to remove any proteins that may have remained bound from prior column use. The column was then prepared washing through 10 ml of His Tag Buffer A before the lysate was injected into the super loop and onto the column. Once the lysate had been completely injected, His Tag Buffer A was flowed through to wash unbound proteins from the column. At this point, His Tag Buffer B was flowed through the column to elute the protein of interest from the column into 1.5 ml fractions. These fractions were selected based on UV absorbance and pooled before buffer exchange. A Ge Health Care Life Sciences HiPrep 26/10 Desalting column was used to exchange the protein into a Tris buffer (Table 2.5) depending on the proteins salt preference.

Table 2.5: The Tris Buffers the proteins were stored in.

Component	**Tris Buffer (Low Salt)**	**Tris Buffer (High Salt)**
Tris-Base	20 mM	20 mM
NaCl	150 mM	500 mM

2.3 Size Exclusion Chromatography

2.3.1 Qualitative Size Exclusion Chromatography

Prior to SEC, protein samples were concentrated to a volume between 0.5 – 1.5 ml. A Ge Healthcare Life Sciences Superdex 200 column was prepared with ~30 ml (1.25 column volumes) of the desired Tris buffer. The sample was injected and protein fractions were eluted out and collected in 0.5 ml fractions. Fractions that showed higher UV absorbance were pooled on a peak by peak basis.

2.3.2 Quantitative Size Exclusion

To quantitatively run size exclusion chromatography, a Ge Healthcare Life Sciences Superdex 200 column was prepared with ~30 ml of the corresponding Tris buffer (Table 2.5). A protein sample of 1 ml was loaded into the sample loop using a syringe and injected into the column. The activity of each fraction was quantified using the kinetic assay (Chapter 2.6) to check for DHDPR activity.

2.4 SDS-PAGE Gel

Samples were prepared in accordance to Table 2.6. Samples along with one lane consisting of Novex Sharp Pre-Stained Ladder were added to a Novex Bolt 4-12% Bis Tris Plus gel submerged in a 1x dilution of MES buffer in a gel tank. The gel electrophoresis was set to run at 166V for 35 minutes.

Table 2.6: The contents of the samples for the SDS PAGE gel

Component	**Amount (µl)**
LDS Loading Dye	2.5 µl
Protein Sample	1-5 µl
Water	2.5-6.5 µl
Total Volume	10 µl

2.5 Analytical Ultracentrifugation

Analytical Ultracentrifugation (AUC) experiments were carried out using a Beckman Coulter XL1 protein characterisation system. Sample cells were made up with 400 µl of protein sample and 380 µl of buffer was used in the reference well. These were run at a speed of 50,000 RPM at 20 °C. The number of runs and wavelengths were varied depending on the protein sample.

2.6 Kinetic Assays of DHDPR

The kinetic assay of DHDPR was adapted from (Reddy et al., 1995). Assay samples were made using the components in Table 2.7.

Table 2.7: The concentrations of components used for the DHDPR kinetic assay.

Component	**Initial Concentration**	**Amount Added (µl)**	**Final Cuvette Concentration**
HEPES (pH 8.0)	100 mM	850	81.73 mM
NAD(P)H	3.24 mM	50	0.156 mM
DHDPS	~1 mg/ml	20	~0.005 mg/ml
(S)-ASA	40 mM	50	1.92 mM
Pyruvate	1 – 40 mM	10-50	0.048 – 1.92 mM
DHDPR	~1 µg/ml	20	
Total		1000-1040	

The spectrophotometer was equilibrated to read absorbance at 340 nm (NAD(P)H oxidation wavelength). To equilibrate the correct concentration of DHDPR, an undiluted sample of DHDPR was added to the assay. If the rate was not linear, the enzyme was diluted with buffer until the rate remained linear during the entire assay time. This concentration was usually around 1 - 10 µg/ml but be higher or lower depending on the effectiveness of the enzyme. Cuvettes containing HEPES, DHDPS and (S)-ASA were incubated at 25 °C for 10 - 20 minutes to equilibrate to

temperature. Once equilibrated, pyruvate was added and the cuvette was mixed by pipetting the contents up and down several times. After 3 minutes, NAD(P)H was added and the cuvette was mixed again. At this point, data collection began and the rate was checked to ensure no background activity. After 5 minutes, DHDPR was added to initiate the assay. The initial rate was measured in mAU through the linear drop in nucleotide absorbance over time. This allowed for the calculation of kinetic parameters of the substrate whilst the enzyme utilised either NADH or NADPH such as the K_m, V_{rel} and V/K_m.

2.7 Small Angle X-ray Scattering

Measurements were carried out at the Australian Synchrotron small angle scattering (SAXS) beamline with a Pilatus 1M detector (169x179 mm; pixel size 172x172 μm). The X-ray wavelength was 1.0322 Å. Protein samples were filtered and concentrated to ~4 mg/ml prior to elution through a Superdex 200 Increase 3.2/300 column immediately before data collection to remove any protein aggregates. Data was collected using a glass capillary at 20 °C in 2 second intervals. Intensity plots were radially averaged, normalised and background subtracted using the Scatterbrain program in the ATSAS suite (Petoukhov et al., 2012).

Guinier analysis was carried out using PRIMUS (Konarev et al., 2003). Porod volumes were calculated using GNOM (Svergun, 1992). GNOM was also used to generate the real space function *P(r)* using an indirect Fourier transform. Comparison of experimental and theoretical scattering curves was achieved using CRYSOL (Svergun et al., 1995). GASBOR was used to create *ab initio* models using a dummy residue model (Svergun et al., 2001).

2.8 Differential Scanning Fluorimetry

Differential scanning fluroimetry (DSF) was carried out on an Applied Biosystems QuantStudio 3 Real-Time PCR System. Each reaction condition (100 μl) was split into four 25 μl samples and placed in separate wells on a PCR plate on ice. The system was cooled to 4 °C and the temperature increased by 0.5 °C per minute until

the system reached 100 °C. The data was fitted to a Boltzmann derivative model using the Protein Thermal Shift software and the four replicate were values averaged.

Chapter 3 – DHDPS

3.1 Background

3.1.1 Quaternary Structure and Mechanism

DHDPS has been highly characterised due to its position and role in regulation of the DAP pathway. The enzyme typically exists in a homotetrameric conformation consisting of four identical subunits in most organisms (Figure 3.1). These individual subunits exist as a classic TIM barrel fold. This lyase catalyses the first committed step in the DAP pathway; a condensation reaction between (S)-ASA and pyruvate to form the product HTPA and a water molecule.

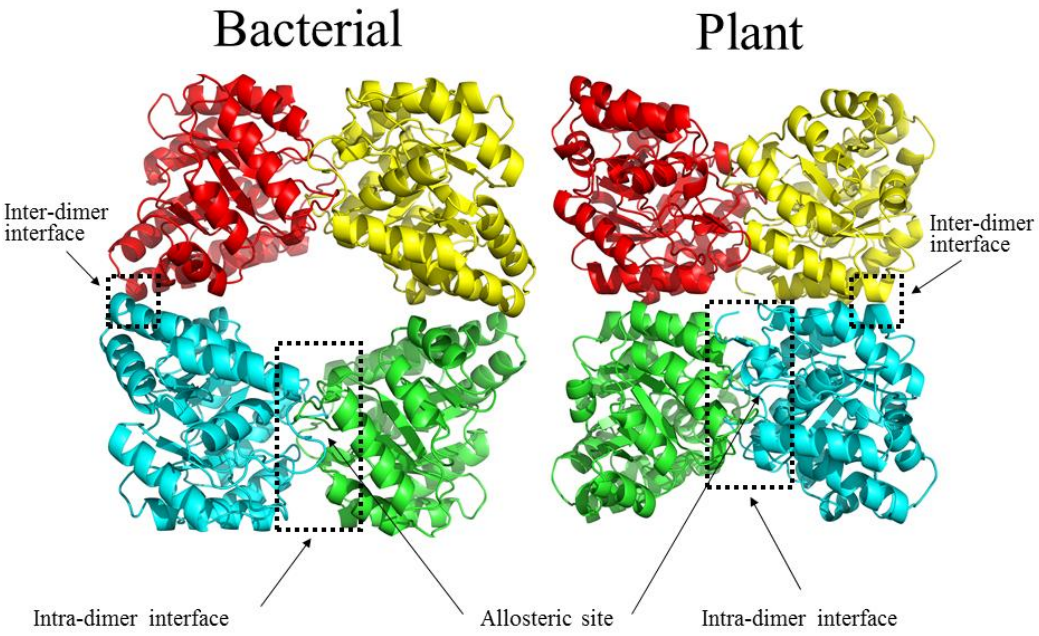


Figure 3.1: *E. coli* and *A. thaliana* crystal structures of DHDPS with key regions labelled.

Kinetically, the reaction exists as a ping pong mechanism. The proposed mechanism has been outlined in *E. coli* (Figure 3.2) (Dobson et al., 2005a). Pyruvate is attacked

by Lys 161 which forms a Schiff base. (S)-ASA subsequently binds and is dehydrated and cyclised forming the final product, HTPA. The specific importance of Lys 161 in the mechanism was tested through alanine and arginine mutants which modified the side chain position of the catalytic nitrogen. These mutants were found to be catalytically active with a slight decrease in activity (Soares da Costa et al., 2010). It was speculated that in step one, the enol form of pyruvate in equilibrium can function as an enamine in the reaction mechanism. This would allow pyruvate itself to attack the proton from Tyr 107 which may mean that Lys 161 may not have as big a role as first thought.

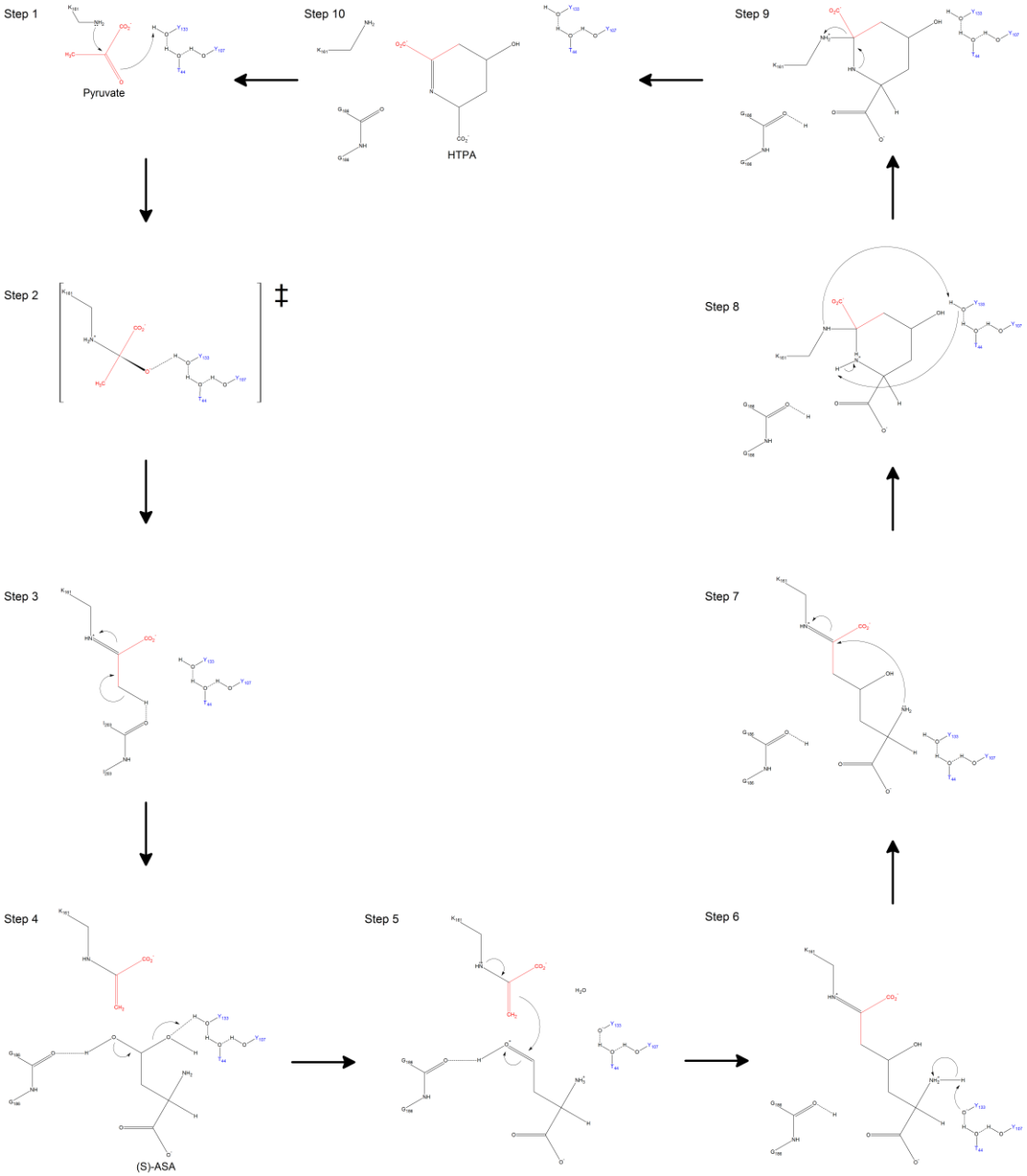


Figure 3.2: The mechanism of DHDPS (adapted from (Dobson et al., 2005a))

3.1.2 Key Residues

It is widely accepted that DHDPS contains three crucial residues located between two subunits interface. Known as the "catalytic triad", these residues form a proton shuttle from the solvent outside the protein to the active site (Figure 3.3). The first such residue is Tyr 133 (*E. coli* numbering). This residue sits directly above Lys 161 and is proposed to have a role in step 1 of the mechanism (Dobson et al., 2004). Mutation to a glutamic acid resulted in a 99.7% drop in activity compared to wild type (Dobson et al., 2004). The specific placement of the residue allows binding to the hydrate of (S)-ASA and proton donation to occur.

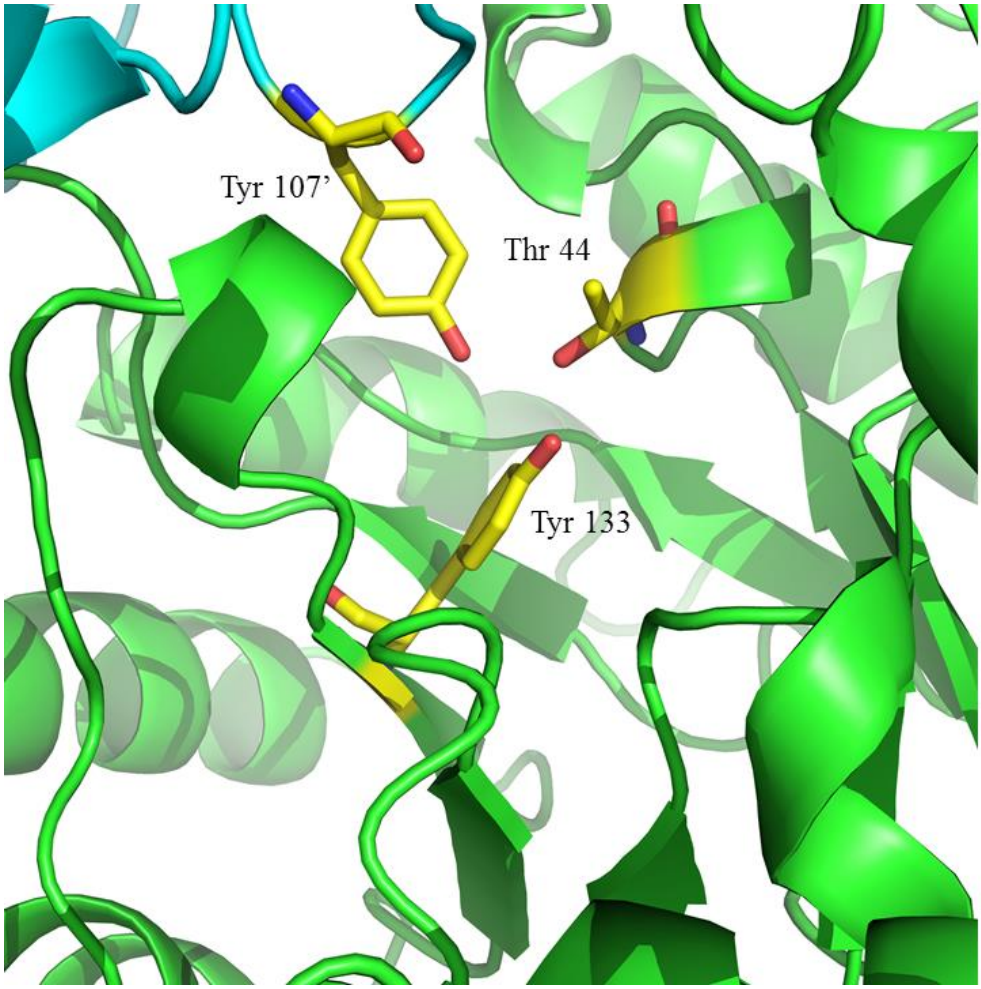


Figure 3.3: The catalytic triad of *E. coli* DHDPS active site residues: Thr 44, Tyr 133 and Tyr 107' from the adjacent monomer.

Another critical residue in the catalytic triad is Thr 44. Located adjacent to Lys 161, it is hydrogen bonded to the two other catalytic triad residues forming a proton shuttle (Dobson et al., 2004). While the residue is not suggested to be directly involved in substrate binding, it appears to play a part in Schiff base formation mechanistically stabilising Tyr 133 through hydrogen bonding (Blickling et al., 1997). This was tested with a mutation to a valine which resulted in a 99.9% drop in maximum rate but no change in affinity for either substrate (Dobson et al., 2004). A slight change in Tyr 107 location was also observed which rendered the enzyme much less catalytically active. While it was speculated that a mutation to a serine would establish activity due to its retained hydroxyl group , it was only restablished to 8% of the wild type and the mechanism was modified to a ternary complex mechanism with a trapped α-ketoglutarate adduct in the active site (Dobson et al., 2009) (Figure 3.4). This adduct has also been observed in various dimeric and catalytic triad mutants causing speculation that flexibility of active site residues can reduce specificity and cause other similar molecules to bind the enzyme instead (Reboul et al., 2012).

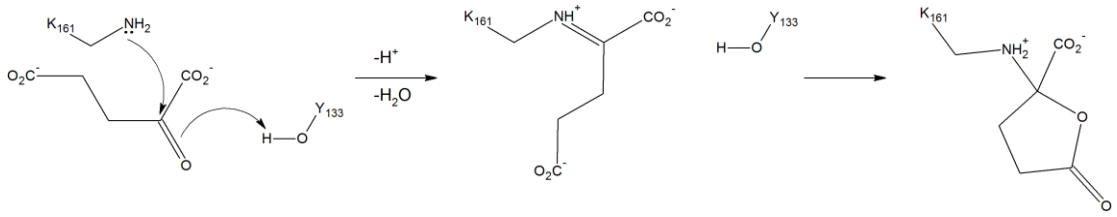


Figure 3.4: The mechanism of formation of the α-ketoglutarate adduct formed in certain dimeric DHDPS mutants (adapted from (Dobson et al., 2009))

The final residue in the catalytic triad is Tyr 107. This residue appears to contribute to the tight binding interface of the dimers while also playing a role in the proton shuttle between the active site and solvent similar to Thr 44 (Blickling et al., 1997, Dobson et al., 2004). Phenylalanine mutation of the residue altered the positioning of the relay-stabilising hydroxyl group causing a tenfold loss of activity which was not as pronounced as the other triad mutations (Dobson et al., 2004). This may be due to the positioning of the residue in a channel leading to the dimer interface. A water molecule from this channel could move into a position near where the hydroxyl of Tyr

107 would have been to partially fulfil its role (Dobson et al., 2004, Griffin et al., 2008).

While the triad residues are most important in its mechanism, there are other important residues key to enzyme activity. Arg 138 is one such residue sitting at the active site entrance (Dobson et al., 2005a). Highly conserved among DHDPS, its role appears to be in binding (S)-ASA and stabilising the proton relay. Its importance was tested through the creation of two mutants; R138H and R138A. The histidine mutant retains the steric bulk and basic activity of the arginine but removes direct interactions with the substrate, and the alanine mutant removes all side chain functionality previously present. Both mutants display greatly decreased substrate affinity and catalytic function showing that specificity for arginine at this position is critically important to catalysis. The R138A mutant specifically causes the three catalytic triad residues to exist in several different conformations, increasing flexibility and distances between residues in the hydrogen bonding network.

Ile 203 is also thought to play a role in catalysis in *E. coli* DHDPS (Dobson et al., 2008). This residue binds to a hydroxyl group of the substrate analogue β-hydroxylpyruvate. The residue also has a highly strained torsion angle of 201° which is greater than four standard deviations from the average torsion angle. It is speculated that the residue could play a role in the mechanism (Dobson et al., 2008).

3.1.3 Dimeric Interfaces

The interdimer interface is the smaller dimer interface responsible for holding the tetrameric structure together along with reducing protein dynamics. In *E. coli*, the total contact area of the solvent accessible surface area of the inter dimer interface is 4.8% compared to the 11.2% in the intra dimer interface (Dobson et al., 2009). This has given rise to the terms “tight” and “loose” dimer interfaces for the intra-dimer and inter-dimer interfaces, respectively.

Several residues present at the inter-dimer interface are critical in its interactions. Leu 196 is a residue that contacts the equivalent leucine from the neighbouring subunit through hydrogen bonding (Dobson et al., 2009). Mutation to a negatively charged aspartic acid removed these hydrogen bonding interactions and introduced

electrostatic repulsion between monomers forming a dimer-tetramer equilibrium in solution. The adjacent residue, Leu 197, is also important in the stability of the interdimer interface (Griffin et al., 2008). Mutation of the residue to an aspartic acid also led to dimeric variants and greatly reduced activities compared to the wild type (1.4% and 2.5% respectively).

Asp 193 is another important residue in the interface. It hydrogen bonds with water molecules that occupy cavities in the interface (Dobson et al., 2009). Substitutions of the residue with alanine and tyrosine removed hydrogen bonding interactions and in the case of tyrosine, introduced steric hindrance. These mutations both led to a dimer-tetramer equilibrium in solution. Glu 234 and Glu 175 also form cross interface contacts with each other. When these residues were replaced with a negatively charged aspartic acid, electrostatic repulsions similar to the Leu 196 mutant formed a dimer-tetramer equilibrium. This was not to as great of an extent as the Leu 196 and Asp 193 mutations, indicating that this interaction is less important than these residues.

The larger intradimer interface is of critical importance to the lysine inhibition mechanism due to its proximity to the binding site. Crystallised *E. coli* DHDPS with lysine bound showed the interactions and residues in the interface (Figure 3.5).

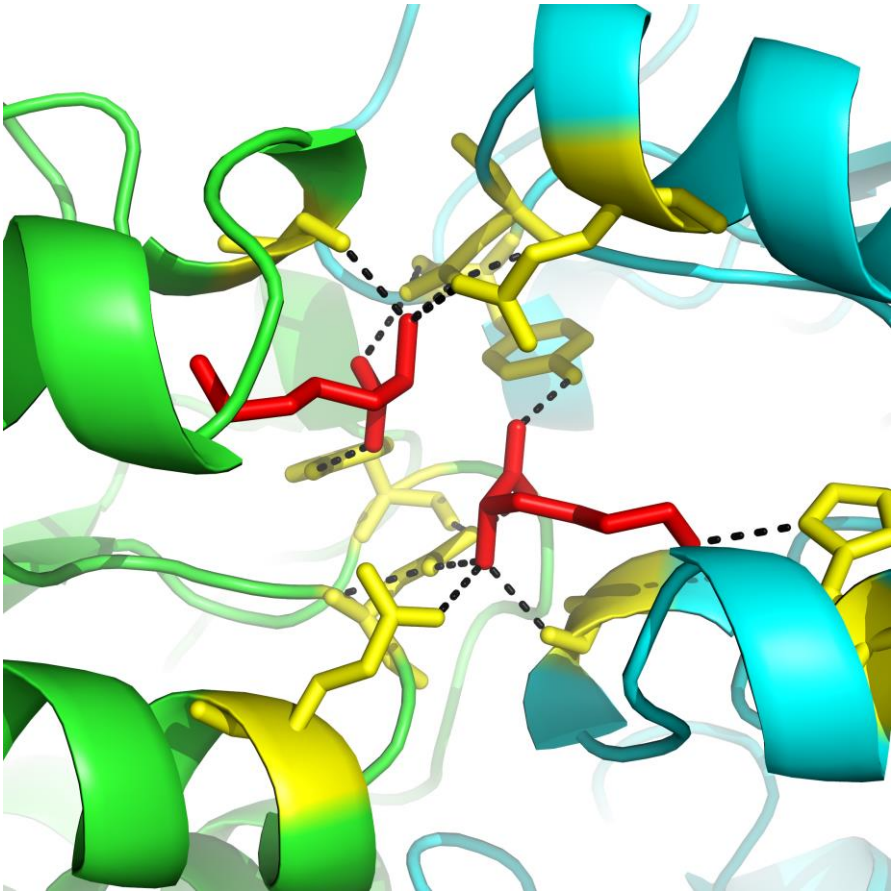


Figure 3.5: Binding region for lysine at the intradimeric interface in *E. coli*. The five critical residues (yellow) shown from each monomer that bind lysine (shown in red) are: Ala 49, His 56, Asp 80, Glu 84 and Tyr 106.

Comparison of a lysine insensitive DHDPS from *C. glutamicum* with *E. coli* DHDPS showed several differences in the residues involved in lysine binding (Rice et al., 2008). In *E. coli*, seven amino acid residues are important in stabilisation of lysine at the binding site. However, in *C. glutamicum*, key differences prevent these interactions. His 56 (Lys 71 in *C. glutamicum*) hydrogen bonds to lysine in *E. coli* but the equivalent residue in *C. glutamicum* DHDPS does not. It is thought that this is the key residue in determining whether the enzyme undergoes lysine inhibition or not (Soares da Costa et al., 2016). If a glutamic acid or histidine is present, lysine binding can occur but a lysine or arginine in this position stops lysine binding due to a lack of side chain interactions. Mutation of His 56 in *E. coli* to its *C. glutamicum* counterpart along with Glu 84 was found to remove lysine inhibition (Geng et al., 2013). His 53 appears to be another important residue in *E. coli* which acts as a "lid" which stabilises the lysine in place whereas the equivalent Ala 68 in *C. glutamicum* DHDPS

was too small to serve the same function. Glu 48 in was also found to be important in *E. coli*. This residue hydrogen bonds with the substrate lysine but the equivalent Thr 99 in *C. glutamicum* DHDPS does not have a long enough side chain to provide the same function.

Monomeric mutant DHDPS have also been created using a combination of both inter and intra dimeric dissociation (Muscroft-Taylor et al., 2010). A dual Y107F/L197D mutant was found to exist as a monomer in size exclusion chromatography and analytical ultracentrifugation. This variant also had greatly reduced activity (8% of the wild type) with the ternary complex mechanism due to triad disruption and did not undergo lysine inhibition. This shows the importance of the quaternary structure in lysine inhibition.

3.1.4 Lysine Inhibition Mechanism

It has been speculated that lysine mediated DHDPS inhibition is a result of disruption of the catalytic triad proton relay. This is thought to be caused by Tyr 107 displacement instigated by movement of its neighbouring Tyr 106 in *E. coli* (Dobson et al., 2005b). This is also observed in the plant DHDPS *Vitis vinifera* where a rearrangement of the neighbouring hydroxyl group in Tyr 131 (Tyr 106 in *E. coli*) causes displacement in the neighbouring catalytic triad residue Tyr 132 (Tyr 107 in *E. coli*) disrupting the proton relay and inhibiting the enzyme (Atkinson et al., 2013).

Another example of lysine inhibition is in the Gram-negative bacteria *Campylobacter jejuni* DHDPS. Binding of lysine shows domain movement of residues 104-184 (Conly et al., 2014). Tyrosine residues 110 and 111 (106 and 107 in *E. coli*) were found to be critical in this domain movement. A Y110F mutant removed the domain movement and substantially reduced inhibition. Active site size also appears to play a role in lysine regulation (Conly et al., 2014). When pyruvate bound to the apo form of DHDPS, the active site volume decreased by 44% and the allosteric site volume by 16% compared to the lysine bound form which increased the active site volume by 30%. The Y110F mutant did not have these changes in volume further showing the importance of this tyrosine residue in the lysine binding mechanism.

3.1.5 DHDPS as a Drug Target

As the first committed step in the DAP pathway, inhibition of DHDPS through small molecules could be a viable drug target. However, this would only be the case if high levels of both specificity and inhibition could be assured (Hutton et al., 2007). Most attempts to achieve this have been through the construction of DHDPS substrate (S)-ASA analogues. While many heterocyclic compounds that covalently bind to the enzyme were found to be effective in inhibiting activity, only a few show any form of specificity (Mitsakos et al., 2008). Chelidamic acid and its diester were found to potently inhibit *E. coli, M. tuberculosis* and *S. aureus* DHDPS but did not inhibit *B. anthracis* DHDPS. Other compounds such as piperidine ester and thiazanes inhibited *E. coli* and *B. anthracis* DHDPS much more strongly than *M. tuberculosis* and *S. auerus* DHDPS (Mitsakos et al., 2008). Antibacterial activity of each compound exhibited various levels of success showing the potential for future development of inhibitors. It has also been noted that partial enzyme activity has been observed with some DHDPS substrate analogues (Devenish et al., 2010). Substrates with an electronegative substituent in the 3' position were found to be compatible with the enzyme but at a significantly lower rate than (S)-ASA. This information could be used to create new inhibitors which could covalently bind or block further pathway enzymes such as DHDPR.

There have been other recent developments in targeted inhibitors for DHDPS. A molecule known as 2-methylheptyl isonicotinate (2MHI) was isolated from *Streptomyces sp.* 201 (Singh et al., 2012). This molecule was found to form favourable interactions with binding sites in *M. tuberculosis* and exhibit anti-bacterial activity. Another recent development has led to creation of a novel DHDPS inhibitor known as "bislysine" (Skovpen et al., 2016). This molecule effectively exists as two lysine molecules linked together by an ethylene bridge between its α-carbon binding in two binding sites at the same time. This inhibitor was found to bind *Campylobacter jejuni* DHDPS 300x tighter than lysine, making it a potential antibiotic molecule.

3.2 Evolutionary Differences

DHDPS has a significant evolutionary structural difference between bacteria and plants. While it is known that the plant version exists in a different arrangement of monomers to the bacterial protein, it is uncertain why this change occurs. These differences can be explained by the proposed evolution of its quaternary structure (Griffin et al., 2008). It is hypothesised that the enzyme first existed as a monomer containing a catalytic dyad made up of the two residues from the individual monomer; Tyr 133 and Thr 44. To fill the role of Tyr 107, an ordered water molecule is thought to complete the reaction. However, this reaction could be optimised through the addition of Tyr 107 from another monomer to form a dimer and complete the triad. However, this setup is still not optimal as this dimer would suffer from small "breathing" movements of residues which reduces substrate specificity and catalytic efficiency. Specifically, movement of Tyr 107 causes the hydrogen bonding network to become more strained, negatively affecting catalysis. Formation of a tetramer from two dimers would prevent excessive dynamics increasing rate and substrate specificity. However, this tetramer could be formed in either front or back facing formations leading to the differing monomeric orientations between plants and bacteria.

There are special cases of DHDPS that do not adhere to the typical tetrameric structures. These include the dimeric DHDPS of *S. auerus*, *P. aeruginosa* and *L. pneumophila* (Burgess et al., 2008, Kaur et al., 2011, Soares da Costa et al., 2016). While dimeric mutants of tetrameric DHDPS are not catalytically optimal due to "breathing" motions, these native dimeric enzymes exhibit similar activity and specificity to their tetrameric counterparts. A much more extensive intradimeric interface is observed which has a greater solvent inaccessible surface area than other DHDPS. While *E. coli* was found to have 7 residues involved in hydrogen bonding and 3 involved in hydrophobic contacts at the intradimeric interface, dimeric *S. auerus* DHDPS has 17 residues involved in hydrogen bonding and 18 hydrophobic contacts (Burgess et al., 2008). This makes *S. auerus* much more effective as a dimeric protein compared to mutant *E. coli* dimers.

There are also other interesting differences between *S. aureus and E. coli* DHDPS. One of the catalytic triad residues Tyr 109 (Tyr 107 in *E. coli*) has a significantly different orientation in which the aromatic ring of the tyrosine is rotated by 60°. This gives a more tightly packed aromatic stacking interaction, eliminating the need for a tetrameric structure. A similar movement of a catalytic triad residue was observed in *P. aeruginosa* DHDPS (Kaur et al., 2011). Tyr 133 was found to have rotated 46° compared with *E. coli*. This may have a similar effect to *S. auerus* DHDPS in keeping the aromatic stacking interactions more tightly packed than in typical tetrameric DHDPS.

3.3 SMO DHDPS Sequence Analysis

Selaginella moellendorffii is an ancient lycophyte that is commonly used in comparative genomics as a model organism. Due to its location on the evolutionary lineage as the oldest extant division of vascular plants, it is an ideal target organism to characterise the evolution of both bacterial and plant type DHDPS enzymes.

To help ascertain whether SMO DHDPS exhibits properties similar to other DHDPS enzymes, its sequence was compared to a bacterial type *E. coli* and a plant type *A. thaliana* DHDPS (Figure 3.6).

Figure 3.6: Sequence alignment of *E. coli, A. thaliana* and *S. moellendorffii* DHDPS. Catalytic triad residues are indicated in blue, other catalytic residues in orange and *E. coli* lysine binding residues are shown in green.

The sequence alignment shows many conserved residues from both *E. coli* and *A. thaliana* DHDPS. All three catalytic triad residues (Thr 44, Tyr 107 and Tyr 133) are conserved along with other residues thought to be important such as Lys 161, Gly 186 and Ile 203. The enzyme was found to exhibit typical DHDPS activity with tight lysine inhibition similar to other plant DHDPS enzymes (unpublished data). One minor difference is observed in the major lysine binding residues. In *E. coli,* Ala 49 is thought to be involved in lysine binding (Soares da Costa et al., 2016). However, in both *A. thaliana* and *S. moellendorffii* the residue is altered to a glutamine and histidine respectively. This may partial account for the tighter lysine inhibition in plant type DHDPS enzymes. All three sequences were found to contain the critical His 56 residue which indicates that they all experience some form of lysine inhibition.

3.3 Transformation, expression and purification of SMO DHDPS

The plasmid containing the gene for SMO DHDPS was transformed into *E. coli* BL21 DE3 cells using standard lab protocols prior to purification. The enzyme was purified using the techniques described in the methods. This produced 25.8 mg of protein from 1.8 L of media. The protein was stored in "low salt" 20 mM Tris and 150 mM NaCl buffer.

3.4 Analytical Ultracentrifugation

Analytical ultracentrifugation (AUC) is a technique which measures the real-time sedimentation of a protein solution under a centrifugal force. This gives information about the protein such as the mass and shape. In this context, it can be used to test the molecular mass of SMO DHDPS while bound to lysine and pyruvate, as well as no ligand which allows the quaternary state of the protein to be determined (Figure 3.7).

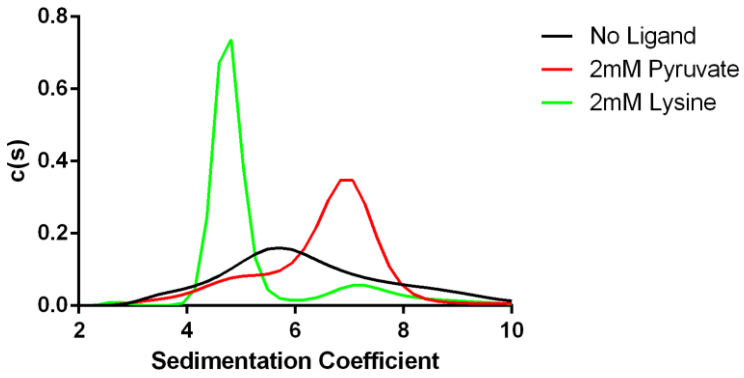


Figure 3.7: AUC data from SMO DHDPS in solution under 3 different conditions.

With no ligand present, it could be hypothesised that the protein exists not in a dimer nor a tetramer but in an equilibrium of the two forms. The single peak at ~5.7 S is observed. This value can be converted to an approximate molecular weight using the Svedberg equation. This equates to an estimated molecular weight of ~99,000 Da. With the theoretical monomer weight calculated to be 40,713 Da, this indicates a dimer-tetramer equilibrium or a trimeric species. However, the long broad peak observed is more indicative of an association-disassociation equilibrium in which the

protein is constantly switching between dimeric and tetrameric conformations. As this is not two separate species in solution, this only appears as one peak in the graph.

The lysine bound form exhibits a major peak at ~4.8 S which equates to an estimated molecular weight of ~77,000 Da. This information indicates that the protein forms a dimeric species in the presence of lysine.

With a sedimentation coefficient of ~7 S, the pyruvate bound species has an estimated molecular weight of ~135,000 Da which is consistent with a tetrameric protein in the presence of pyruvate.

3.5 Small Angle X-ray Scattering

In order to confirm the shape of the protein in solution, small angle X-ray scattering (SAXS) was used. The scatter patterns of SMO DHDPS can be compared with model bacterial and plant DHDPS and Chi^2 fits compared.

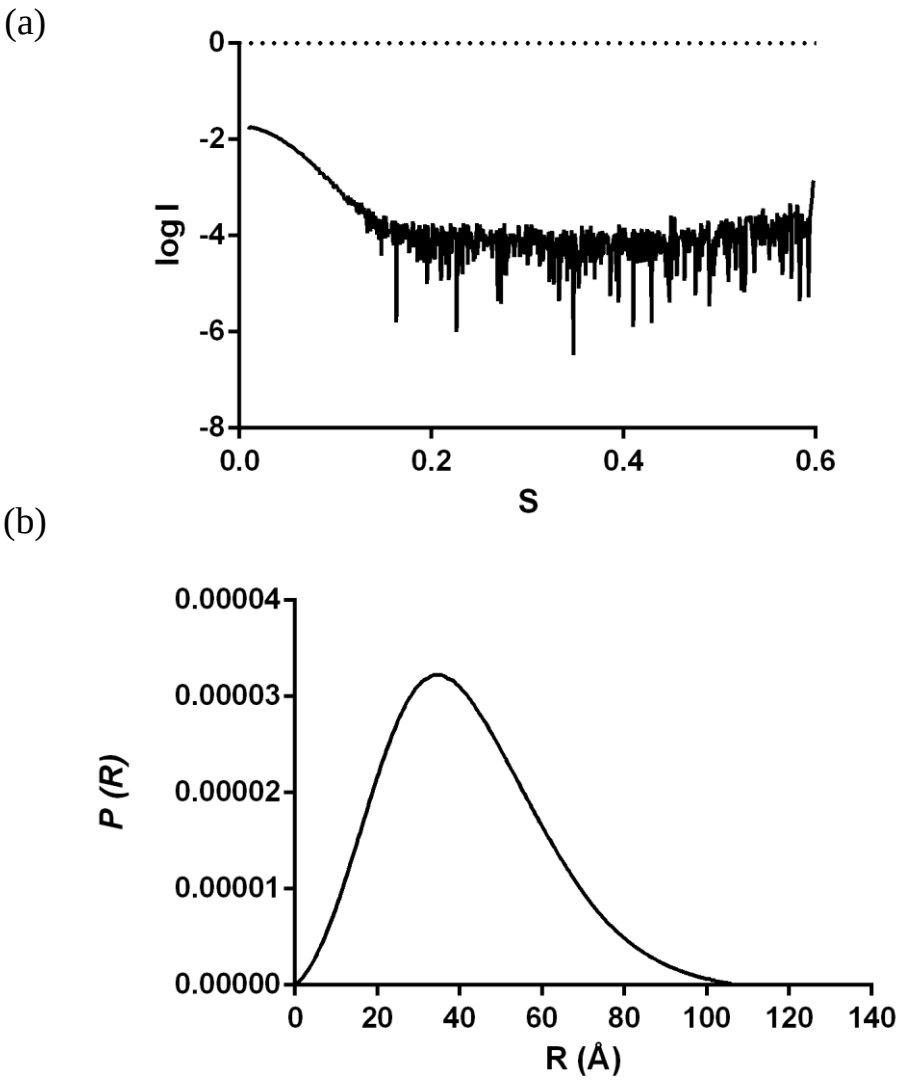


Figure 3.8: Intensity plot (log I vs S) (a) and distance distribution plot (P (R)) (b) for SMO DHDPS in the presence of 5 mM lysine.

Table 3.1: Chi^2 fits for SMO DHDPS in the presence of 5 mM lysine

Comparison Structure	**Chi^2**
E. coli DHDPS	1.60
A. thaliana DHDPS	1.35
Intradimer	0.58
Interdimer	0.59

The SAXS data shows interesting insights into the nature of both lysine and pyruvate bound SMO DHDPS. The *P(R)* plot can show information about the proteins quaternary structure. This plot measures the paired set of distances between points. The *P(R)* plot for lysine bound SMO DHDPS indicates a non-globular rod-like structure (Figure 3.8). In a globular structure, the distribution would follow a standard bell curve as the distribution of mass is spread evenly around the centre of the protein. However, this data shows a noticeable skew off to the left with an uneven distribution. This indicates that the protein is not globular in solution and may represent a rod-like dimeric structure.

The proteins scatter data can also be compared with the *A. thaliana* and *E. coli* native tetrameric crystal forms and two hypothetical dimeric crystal forms; each representing a different possible dimer (Figure 3.9 and Table 3.1). The lysine bound form much more strongly fits the two dimeric forms (0.58,0.59) rather than the bacterial or plant type tetrameric form (1.60, 1.35). This also indicates that the lysine bound form is likely to be a dimeric species.

However, this does not indicate which dimer is being formed. It could be speculated that the interdimeric species is the one that forms. Lysine binding may cause dissociation of the lysine binding site interface. However, the greater surface area of this interface compared to the interdimer interface could mean that the interface is more resilient to movement in residues. The interdimeric interface has much less surface area and may be more susceptible to residue movement which could cause dimeric dissociation. The Chi^2 fits do not show a statistically significant difference between the two dimers. The native dimeric DHDPS enzymes such as those from *S.*

auerus exist in the intradimeric formation. Further investigation as to the dimeric nature of the enzyme is of interest to see if it is in an identical dimeric configuration to the native DHDPS enzymes or in a novel interdimeric configuration. If the dimeric orientations were identical, this would have implications in the proposed evolutionary mechanism.

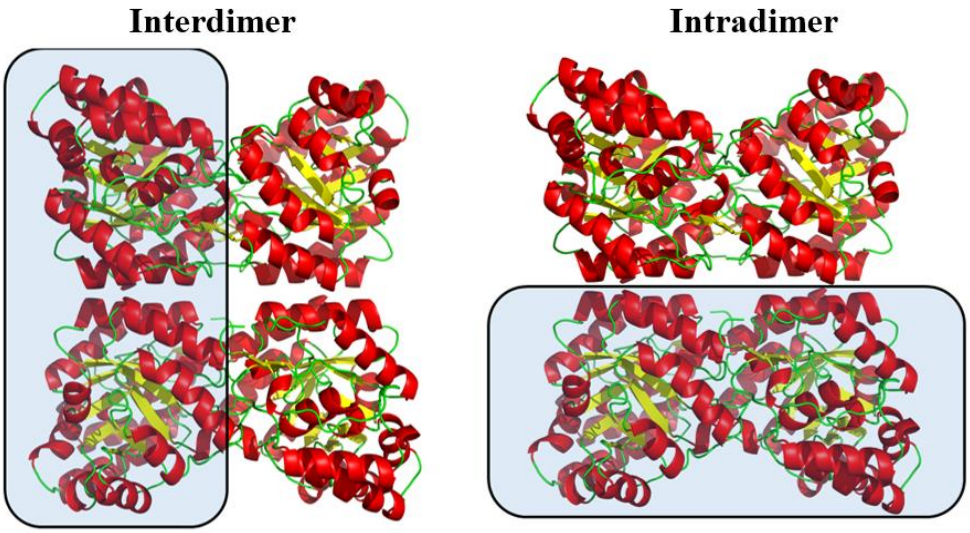
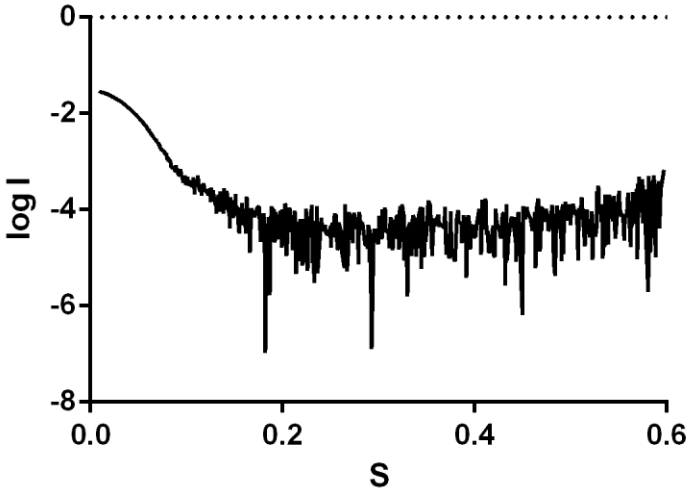


Figure 3.9: The two possible dimeric forms of SMO DHDPS based on the structure of *A. thaliana* DHDPS.

(a)



(b)

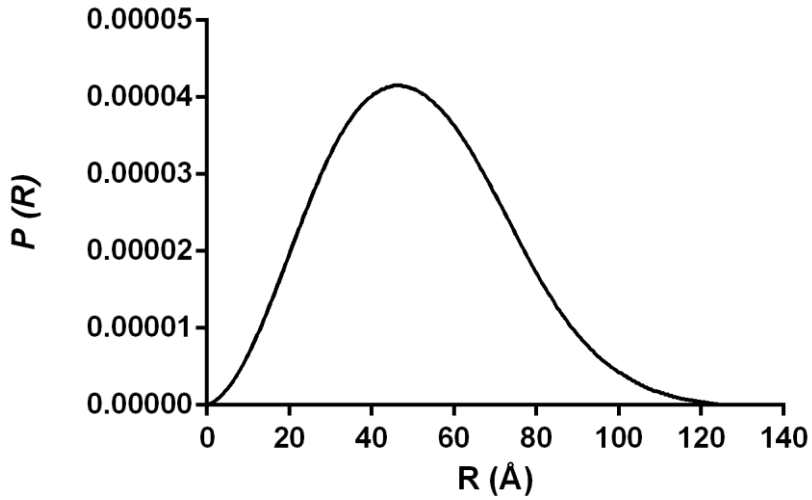


Figure 3.10: Intensity plot (log I vs S) (a) and distance distribution plot (P (R)) (b) for SMO DHDPS in presence of 5 mM pyruvate.

Table 3.2: Chi^2 fits for SMO DHDPS in the presence of 5 mM pyruvate

Comparison Structure	**Chi^2**
E. coli DHDPS	23.30
A. thaliana DHDPS	1.54
Intradimer	17.22
Interdimer	17.23

The *P(R)* plot of pyruvate bound DHDPR indicates a globular structure (Figure 3.10). A relatively uniform standard distribution is observed indicating that unlike the lysine

bound curve, the protein exists in a much more globular tetrameric shape than its lysine bound equivalent. The Chi^2 fits also indicate a tetrameric structure (Table 3.2). The fit for the tetrameric plant type species from *A. thaliana* is much stronger (1.54) then either the bacterial form tetramer from *E. coli* (23.30) or the two dimeric species (17.22, 17.23). Both of these factors are strong evidence for the pyruvate bound form of SMO DHDPS existing in a plant type tetrameric arrangement.

Table 3.3: SAXS parameters for SMO DHDPS in two different buffers along with *A. thaliana*.

Organism	**Ligand**	**Rg (Å)**	**Io**	**Porod Volume ($Å^3$)**
A. thaliana	No Ligand	36.4±0.2	0.055	271,021
S. moellendorffii	5 mM Lysine	31.5±0.3	0.019	127,080
S. moellendorffii	5 mM Pyruvate	38.6±1.9	0.03	285,373

These dimeric and tetrameric observations are backed up by the radius of gyration and zero angle intensity (Io) (Table 3.3). The radius of gyration is an indicator of the distribution of mass around the centre of a protein. This gives an indication of the diameter and size of the protein. The zero-angle intensity gives an indication of the intensity of radiation scattered through zero-angle. The radius of gyration for the pyruvate bound form is similar to the *A. thaliana* radius. However, when both *A. thaliana* and the pyruvate bound scatter patterns are compared to the lysine bound SMO DHDPS, the radius of gyration for this protein is markedly lower than the tetrameric forms. Porod volumes express an estimated molecular weight of 1.6x that of the actual weight. Taking this into account, it also shows that the lysine bound form (~79,000 Da) has a volume approximately half that of the pyruvate bound structure (~178,000 Da). These factors reinforce the idea that the quaternary structure of SMO DHDPS depends on substrate binding.

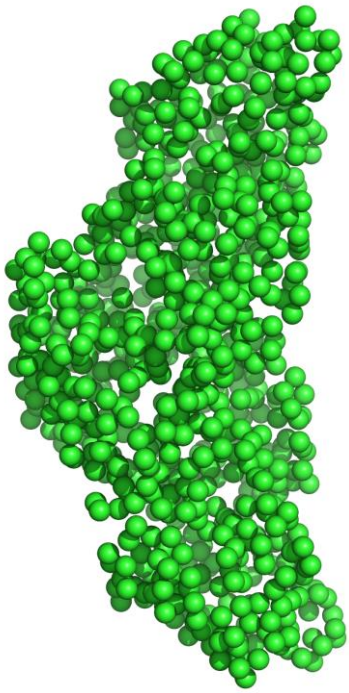


Figure 3.11: GASBOR model created from SMO DHDPS in the presence of 2 mM lysine (Chi^2 fit against raw data = 0.20, P2 symmetry applied)

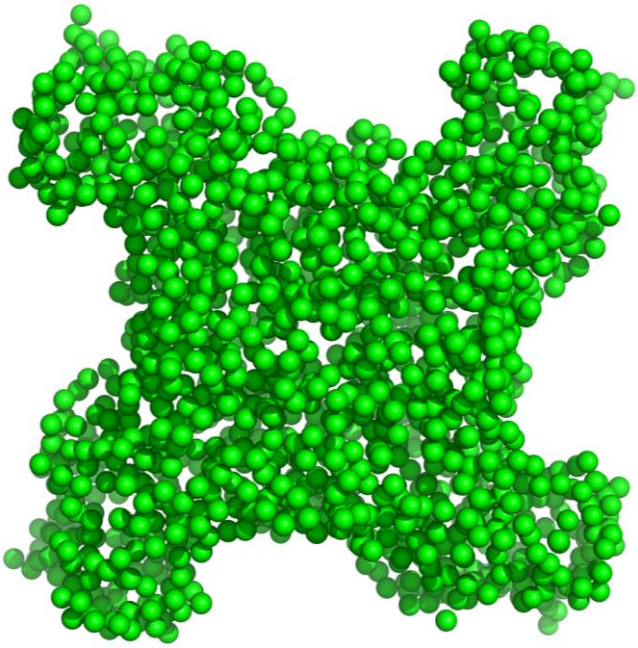


Figure 3.12: GASBOR model created from *Selaginella moellendorffii* in the presence of 2 mM pyruvate (Chi^2 fit against raw data = 0.17, P4 symmetry applied)

In order to gain an insight into the 3D structure; a low resolution *ab initio* model was created using the program GASBOR (Svergun et al., 2001). The model created using the lysine bound data more closely resembled the dimer than the tetramer (Figure 3.11). This was in contrast to the pyruvate bound model which much more closely resembled a plant type tetramer (Figure 3.12).

While SAXS is a powerful technique in the determination and comparison of single species solutions, it has limitations when more than one species is present in solution. This makes it unsuitable to find information about the species with no ligand present. It also does not have a high enough resolution to determine exactly which dimer is formed.

3.6 X-ray Crystallography

To generate a 3D high resolution structure, the protein was screened using JCSG+ and PACT crystal screens due to their high success rate in DHDPS-like proteins in the past. If successful, a high-resolution 3D structure of the protein could be formed. Initially, crystallisation was attempted using a method that led to crystal formation in other DHDPS and DHDPS-like proteins. Two separate trials were done, one with 5 mM lysine and the other 5 mM pyruvate. Both ligands force the equilibrium into a single species which should allow for a higher chance of crystallisation. However, these attempts were unsuccessful. The next attempt increased the ligand concentrations from 5 to 10 mM to increase to stability and homogeneity of the protein. The protein concentration was also greatly lowered to ~2.5 mg/ml in an attempt to replicate a previous successful attempt to crystallise *Vitis vinifera* (grapevine) DHDPS (Atkinson et al., 2011). However, this also proved unsuccessful. In a final attempt to generate crystals, a tray using the same crystallisation conditions but a vastly increased protein concentration (~28 mg/ml) was created. This attempt was also unsuccessful. Successful crystallisation of the protein is an avenue of further investigation into the exact position of its dimer interface residues and may indicate exactly how dimer dissociation occurs.

There are many other parameters that could have been changed to allow for protein crystallisation. For example, alternate screens such as Morpheus and Midas would include conditions that were not present on JCSG+ and PACT. However, JCSG+ and PACT contain a large variety of conditions that have previously been used to crystallise other DHDPS so it is likely that these screens had the highest likelihood of crystallisation.

The sitting drop method was used in these experiments however the vapour diffusion is an alternate method that could be initiated. This technique involves vapour from a precipitant well diffusing into a protein well in a sealed environment. Another form of crystallisation is the hanging drop method. This technique involves a drop of protein solution hanging slightly above a precipitant solution in a sealed environment. This is the technique used in many examples of DHDPS crystallisation and may be more conducive to protein crystallisation (Dobson et al., 2005a, Griffin et al., 2008, Pearce et al., 2008). It is also possible that variation in temperature could increase the probability of crystallisation. Other examples of DHDPS crystallisation use a variation of temperatures ranging from 4-12 °C. Therefore, there remains a number of parameters that remain to be tested in order to optimise protein crystallisation.

3.7 Differential Scanning Fluorimetry

In order to determine protein stability of dimeric and tetrameric species and the influence of ligand binding, melting temperatures were determined by Differential scanning fluorimetry (DSF). This is a technique that measures the unfolding of a protein with increasing temperature using a hydrophobic fluorescent dye that binds the protein as it unfolds. Using this data, the melting point of SMO DHDPS under various ligand conditions can be calculated.

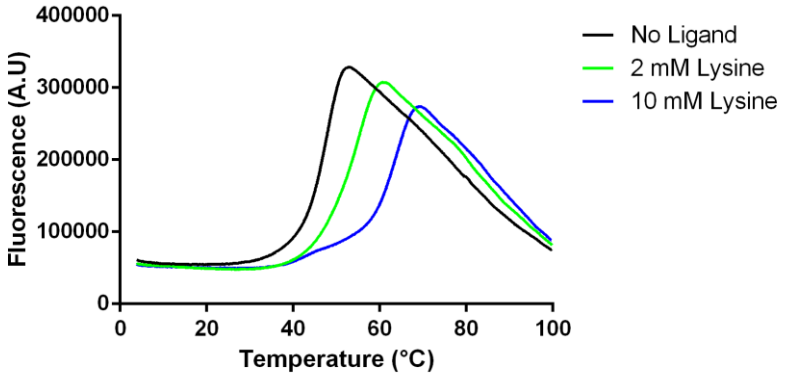


Figure 3.13: DSF data for conditions of no ligand and two concentrations of lysine.

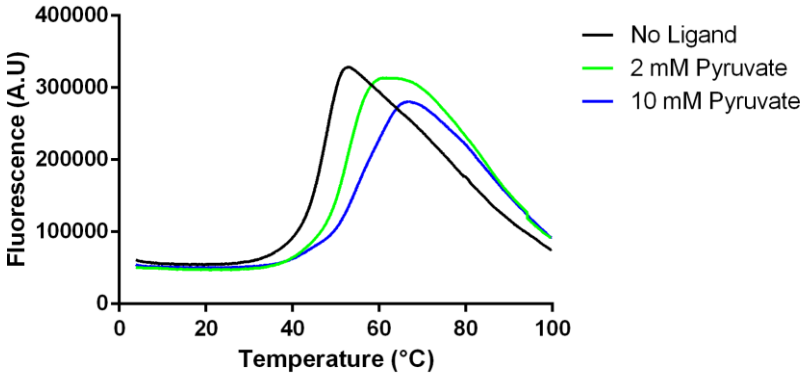


Figure 3.14: DSF data for conditions of no ligand and two concentrations of pyruvate.

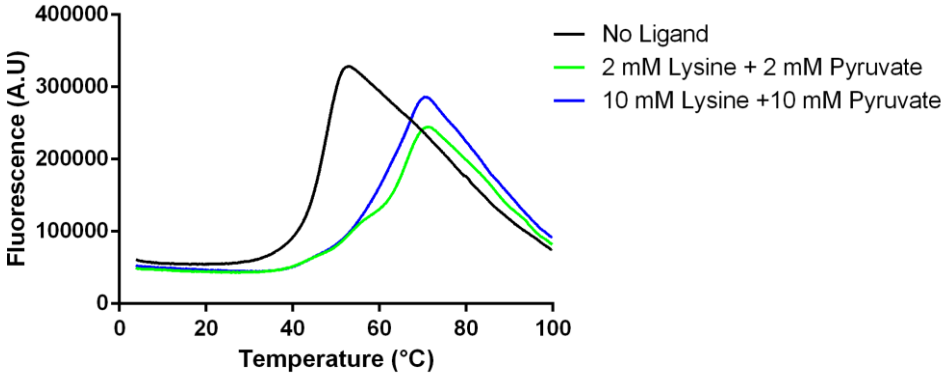


Figure 3.15: DSF data for conditions of no ligand and two concentrations of both lysine and pyruvate.

Table 3.4: Melting temperatures for SMO DHDPS under various ligands

Ligand Condition	**Tm_{ave} (°C)**
No Ligand	47.9±0.1
2 mM Lysine	55.5±0.1
10 mM Lysine	63.8±0.2
2 mM Pyruvate	52.9±0.3
10 mM Pyruvate	55.4±0.1
2 mM Lysine + 2 mM Pyruvate	66.0±0.1
10 mM Lysine + 10 mM Pyruvate	66.7±0.2

From these melt curves, the temperature stability of the DHDPR enzymes can be found (Table 3.4). In absence of ligand, the enzyme has a Tm of 47.9 ºC (Figure 3.13). This is considerably lower than other measured Tms from other DHDPS enzymes such as *B. anthracis* and *E. coli* (59.8 ºC and 61.3 ºC) (Burgess et al., 2008). In other DHDPS, pyruvate has been found to stabilise the enzyme (Kefala et al., 2008, Domigan et al., 2009, Burgess et al., 2008). SMO DHDPS is no different with a 5.1 ºC increase upon addition of 2 mM pyruvate and a 7.6 ºC increase when 10 mM pyruvate is added (Figure 3.13). Interestingly, a greater stability increase is also observed in the presence of lysine (Figure 3.14). A 7.6 ºC increase in stability is observed in 2 mM lysine and a 15.9 ºC increase with 10 mM lysine present. An increase in stability is also observed in the presence of both ligands with 18.6 °C and 18.8 °C increases at 2 and 10 mM concentrations of both lysine and pyruvate (Figure 3.15). As the equivalent amount of lysine stabilises the protein more than the equivalent amount of pyruvate, it seems that binding of ligand at the lysine binding site increases thermal stability to a higher extend than the substrate binding site.

3.8 Conclusions

3.8.1 Proposed Evolutionary Pathway

With discovery of the dimer-tetramer equilibrium of SMO DHDPS, an expanded evolutionary pathway can be proposed. A dimeric DHDPS from *S. auerus* or *P. aeruginosa* could be considered as the "divergence point" for the evolution of the

organism. A substrate mediated equilibrium similar to SMO DHDPS where pyruvate shifts the equilibrium towards the tetrameric form has been observed in *B. anthracis* (Voss et al., 2010). This equilibrium could be considered an intermediate between the bacterial tetrameric arrangement and dimeric DHDPS. With the discovery of the SMO DHDPS equilibrium, this allows an equivalent pathway for the formation of the plant type DHDPS from the ancestral dimer. This makes the discovery of SMO DHDPS an important "missing link" in the evolution of the enzyme (Figure 3.16).

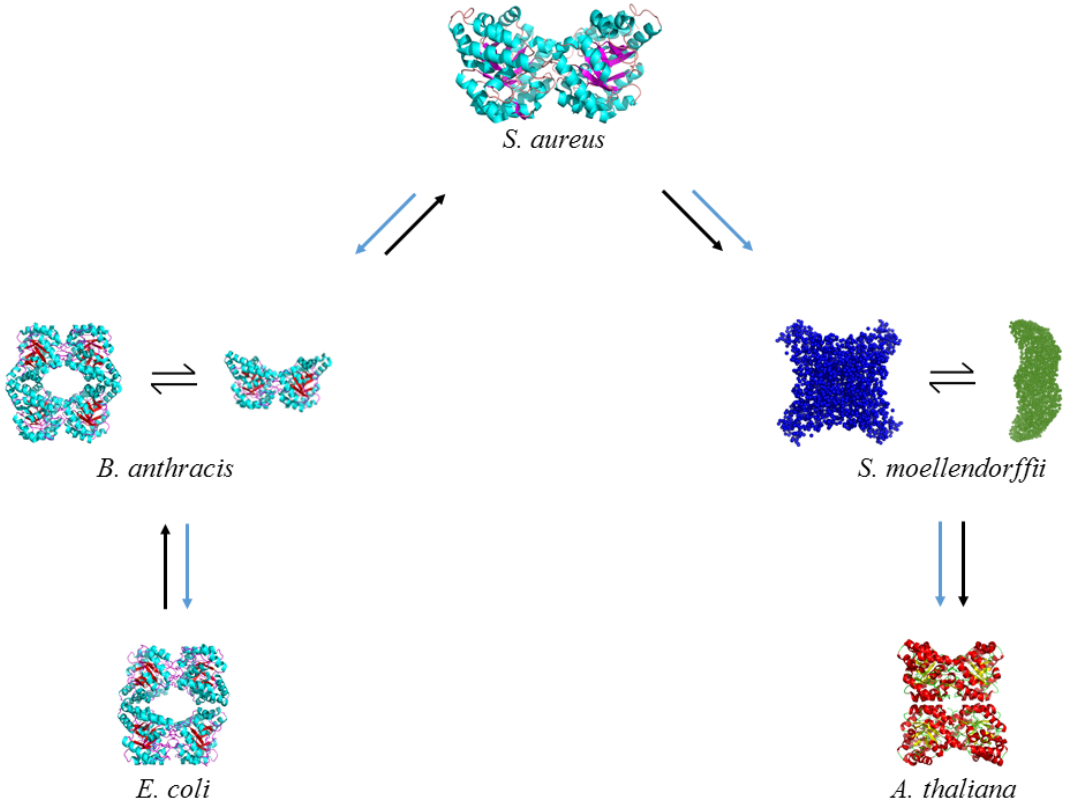


Figure 3.16: The two possible evolutionary pathways of DHDPS. The bacterial DHDPS could possibly have evolved into the dimeric form and towards the plant form (black). It is also possible that both the bacterial and plant tetramers both evolved from the dimeric ancestor (blue). Shown on the diagram: *S aureus, B. anthracis, E. coli, S. moellendorffii* (GASBOR) model and *A. thaliana.*

While each stage of the proposed evolutionary pathway has been found, the exact method that the pathway has formed is unknown. It could occur in two different ways. One possible pathway involves a split from the dimeric form to generate the bacterial and plant tetramers. The other possible mechanism of evolution proposes that the

bacterial tetrameric form evolved back into the dimeric from which then evolved into the tetrameric form.

This proposed evolutionary mechanism could have possible implications in the evolution of other DAP pathway enzymes and other enzymes. For example, it is possible that a similar evolutionary mechanism exists in DHDPR which will be explored later in the DHDPR sections. The model may also be able to be applied to the evolution of other enzymes in the future.

3.8.2 Summary

SMO DHDPS exists in a substrate-dependent equilibrium between dimer and tetramer. AUC and SAXS show that in the presence of the allosteric inhibitor lysine, the protein exists in a dimeric conformation. However, the presence of pyruvate shifts the quaternary structure into a tetrameric conformation similar to the native plant type tetrameric structure. With no ligand present, AUC shows that the protein exists in a dimer-tetramer equilibrium. DSF shows that while both pyruvate and lysine binding increase the stability of the protein, lysine stabilises it by a higher amount. The observation of this equilibrium can be seen as an evolutionary “divergence point” in the evolution of DHDPS for the formation of the plant type tetrameric formation.

Chapter 4 – Algal DHDPR Enzymes

4.1 Background

4.1.1 Quaternary Structure

DHDPR catalyses the second committed step in the DAP pathway. As the second committed step in the DAP pathway, the enzyme remains relatively unstudied compared the first committed step catalysed by DHDPS. The (4S)-hydroxy-2,3,4,5-tetrahydro-(2S)-dipicolinic acid (HPTA) formed from the DHDPS step is converted to L-2,3,4,5-tetrahydrodipicolinate (THDP) using an NADH or NADPH cofactor. The enzyme was first isolated and purified in 1973 by Hadassah Tamir and Charles Gilvarg (Tamir and Gilvarg, 1974). Its monomeric arrangement, like DHDPS, exists as a repeating subunit in a dimeric or a tetrameric arrangement, depending on the organism type (Pearce et al., 2008, Reddy et al., 1995). In all bacterial DHDPR characterised so far, the protein exists as a homo-tetramer. However, plant type DHDPR enzymes characterised so far have been found to exist as homo-dimers (Griffin et al., 2012, Watkin, 2014).

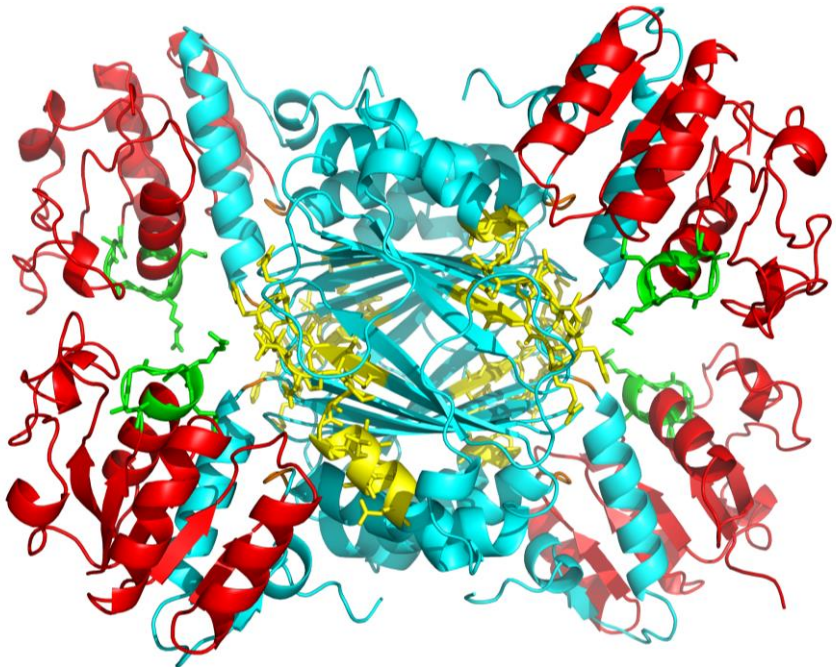


Figure 4.1: *E. coli* DHDPS in its tetrameric structure with the nucleotide binding domain shown in red and the binding loop shown in green. The substrate binding domain is shown in cyan and the binding loop shown in yellow.

The general structure of a DHDPR subunit in *E. coli* consists of two distinct domains; the C- and N-terminus domains (Figure 4.1). Each domain has a different structure and function (Scapin et al., 1995). The C-terminus domain consists of four β-sheets and two α-helices which are arranged on the bottom of the β-sheets. This domain is involved in substrate binding as well as maintaining quaternary structure of the protein. The N-terminus domain consists of seven β-sheets and four α-helices arranged in a Rossmann fold; a fold common among other nucleotide binding proteins such as Lactate Dehydrogenase and Ferrodoxin.

4.1.2 Mechanism

The reaction proceeds via a ordered sequential ping pong reaction order in which the nucleotide cofactor binds first, followed by HTPA, with the product THPA leaving first after the reaction followed by the reduced nucleotide cofactor (Figure 4.2) (Reddy et al., 1995) (Scapin et al., 1997). It is characterised by rapid hydride transfer followed by a slower rate-determining step to stabilise the nitrogen ring and complete the reaction.

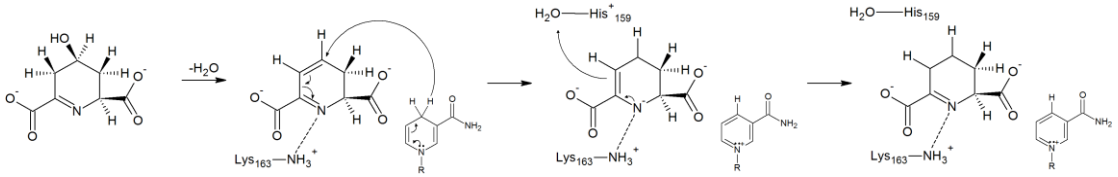


Figure 4.2: The proposed mechanism for DHDPR (adapted from (Scapin et al., 1997))

In the proposed reaction mechanism, Lys 163 and His 159 are thought to be critical to the reaction chemistry. Mutation of these residues to alanine and glutamic acid – residues that did not contain any charged side chains markedly reduced substrate affinity and almost eliminated the reaction rate to zero (Scapin et al., 1997). As an unexpected feature of the mechanism, the protein exhibits dehydratase activity to convert HTPA to DHDP (Devenish et al., 2010). It was initially thought that DHDPR used DHDP as its substrate and that the dehydration of HTPA occurred spontaneously in solution. The discovery that an increased rate of DHDP formation is linked with DHDPR concentration shows that DHDPR directly accepts HTPA as a substrate. This is hypothesised to occur with His 159 acting as a base, deprotonating HTPA to form DHDP (Devenish et al., 2010). However, the exact mechanism of this step in the reaction remains uncertain.

Specificity is another important aspect of DHDPR substrate binding. Several (S)-ASA substrate analogues were found to be turned over in DHDPS (Devenish et al., 2010). However, the resulting product was found to not be compatible with DHDPR. This shows that DHDPR has a different element of substrate specificity than DHDPS.

4.1.3 Domain Dynamics

Domain dynamics are also critically important in the mechanism. Binding of the substrate is thought to rely on a conformational change between the two domains (Figure 4.3) (Janowski et al., 2010). One proposed mechanism states that the protein exists in the "open" conformer with no ligands bound. Nucleotide binding is thought to shift the protein into an equilibrium between open and closed forms with HTPA binding forcing the monomer into the "closed" conformation.

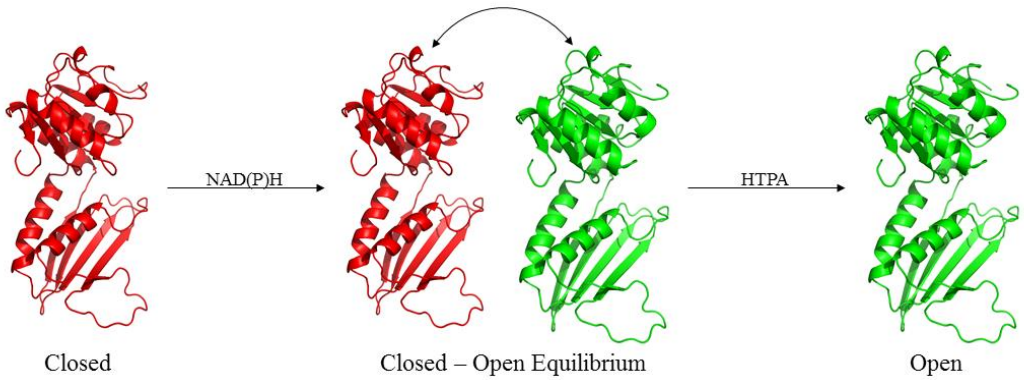


Figure 4.3: The proposed binding mechanism of DHDPR substrate binding.

The conformational change is centred around a "hinge region" consisting of two loops containing residues Phe 129 – Ser 130 and Ser 239 – Arg 240 in *E. coli* (Figure 4.4). Upon overlay of no ligand bound and PDC/NADH bound crystal forms, a 16° rotation around the hinge region is observed (Scapin et al., 1997). Similar domain movement has also been observed in *S. aureus* DHDPR with a rotation angle of 11° between two conformers (Girish et al., 2011).

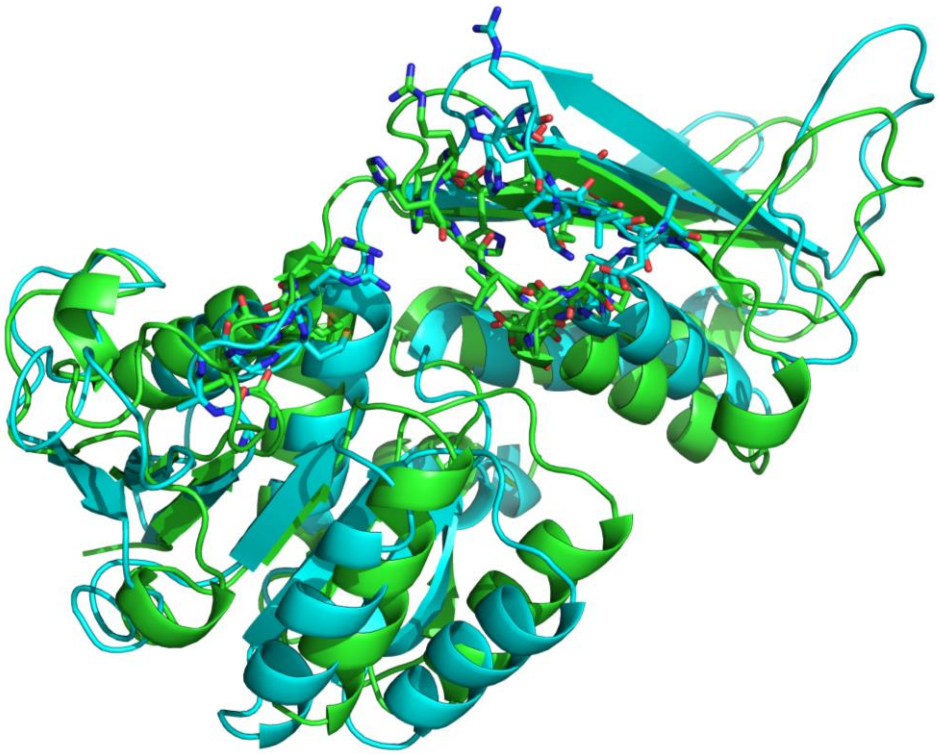


Figure 4.4: This is an overlay of the ligand unbound (cyan) and PDC-NADH bound (green) *E. coli* DHDPR.

This conformational change is thought to occur in three parts. HTPA binds first to the enzyme. This was found induce a greater conformational change than the other steps (Ge et al., 2008). The next two steps are both NADH binding events. One monomer was found to bind NADH before the others resulting in negative cooperativity. This was followed by a 2nd binding event of either two or three NADH molecules binding at once. While crystal structures containing NADH typically contain three of their active sites filled with nucleotide and one that remains unbound, it is unclear whether this is representative of the actual binding mechanism or an artefact of crystal packing. Exactly how this cooperativity occurs is also uncertain. It is speculated that a change in the position of Arg 16 is initiated by Glu 51 movement upon NADH binding, which causes a small rearrangement leading to increased inter monomer repulsion with more than one NADH bound (Ge et al., 2008).

4.1.4 Nucleotide Binding Loop

In binding of the nucleotide cofactor, a conserved region consisting of GXXGXXG in most DHDPR enzymes has been observed in *E. coli*, *M. tuberculosis*, *T. maritima* and *A. thaliana* DHDPR, among others (Figure 4.5) (Griffin et al., 2012, Scapin et al.,

1995, Janowski et al., 2010, Pearce et al., 2008). As a protein that can utilise both NADH and NADPH, *M. tuberculosis* DHDPR is an important model in nucleotide binding. In this case, lysine residues 9 and 11 have been found to anchor NADPH to the protein (Cirilli et al., 2003). Mutations of these residues greatly decreased NADPH affinity.

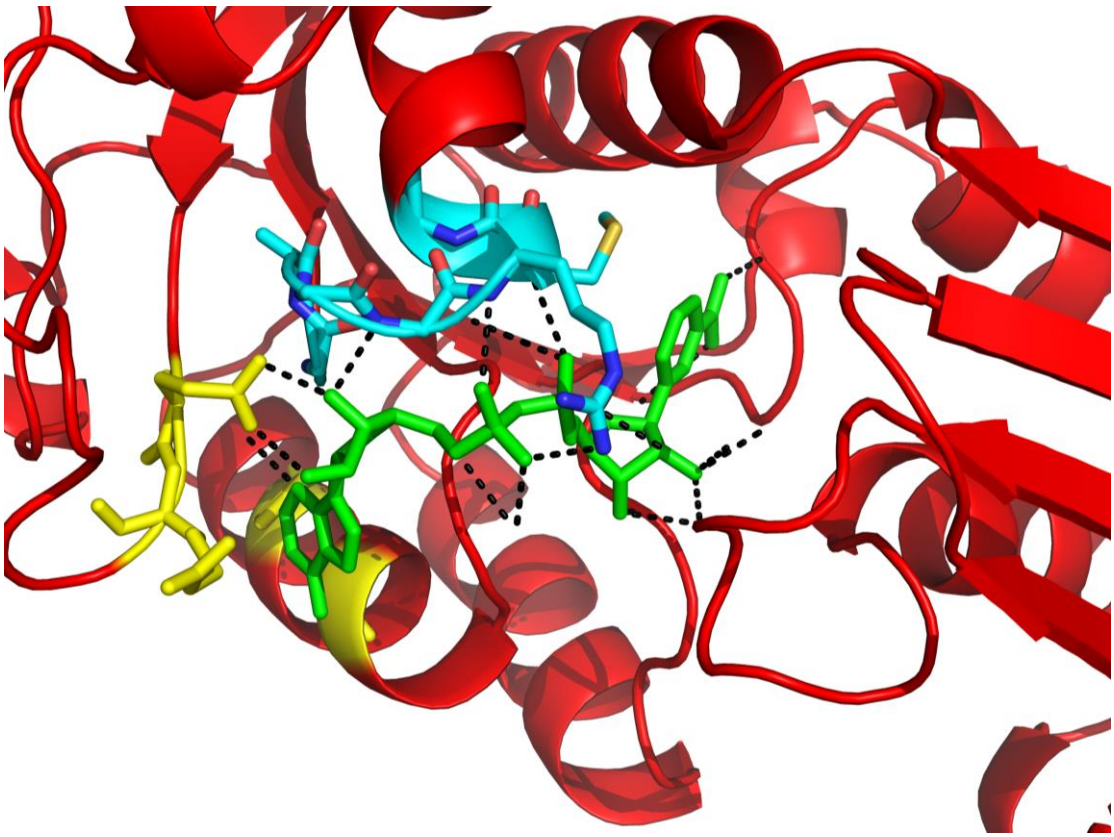


Figure 4.5: The nucleotide binding loop consisting of residues Gly 12 to Gly 18 (shown in cyan) binding to NADH (green) along with other residues deemed important (Glu 38, His 39, Gly 84 and His 88) (yellow) in nucleotide binding.

However, these residues are not conserved in other DHDPR enzymes, even in those that bind NADPH, indicating that the binding residues may differ slightly between individual DHDPR enzymes. While most characterised and uncharacterised DHDPR enzymes discovered so far contain the critical glycine residues in the nucleotide binding motif, some organisms such as *S. auerus* have small variations such as an asparagine in place of the third glycine with similar nucleotide affinity (Dommaraju et al., 2011). This may mean that the exact nature of the residues nucleotide binding motif is not as important as initially thought.

The nucleotide binding motif along with two specific upstream residues have been speculated to dictate whether the enzyme expresses a preference towards NADH, NADPH or both nucleotides (Dommaraju et al., 2011). In *E. coli,* an aspartic or glutamic acid located 19-20 residues upstream from the GXXGXXG at position 38 was found to hydrogen bond with two hydroxyls in the ribose moiety of NADH (Scapin et al., 1995). The phosphate group of NADPH was also found to interact with the basic Arg 39. These residues were hypothesised to decide if the protein prefers to utilise NADH or NADPH. However, there are exceptions to this. For example, *S. auerus* DHDPR exhibits substrate affinity and rate while utilising NADPH but lacks a basic residue equivalent to Arg 39 (Dommaraju et al., 2011). It is also thought that N terminus rearrangement allows Lys 35 interaction with the phosphate of NADPH, accounting for its nucleotide specificity (Girish et al., 2011). NADPH binding was reduced 20-fold in a K35A mutant which reduced side chain functionality. It is therefore uncertain exactly which residues are responsible for nucleotide specificity due to inconsistencies in proposed binding hypotheses.

4.1.5 Substrate Binding Loop

The substrate binding loop is a critically important region of the protein in terms of its function. This is marked by a short binding loop region between Glu 157 and Ala 171 (in *E. coli*) (Scapin et al., 1995). These residues along with residues Gly 102 – Phe 106 and Ala 126 – Ser 130 from the other domain form a substrate binding pocket. Residues in this pocket hydrogen bond with the substrate allowing the residues involved in the mechanism (Lys 163 and His 159) to be positioned correctly (Figure 4.6).

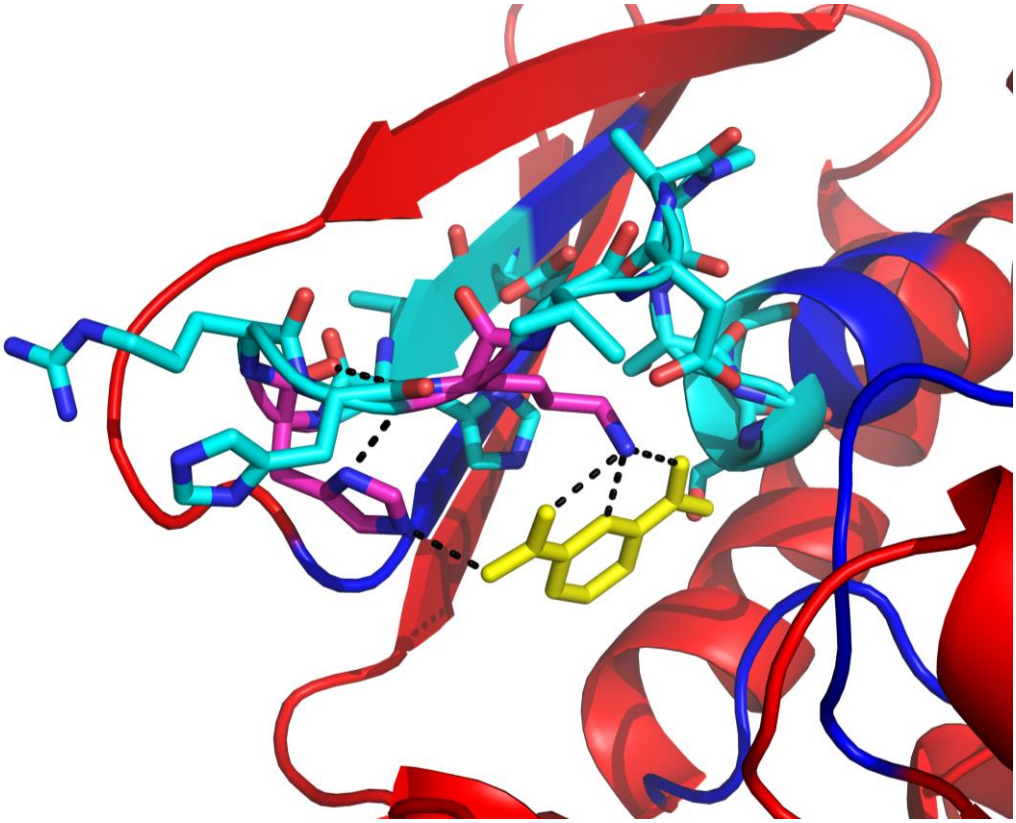


Figure 4.6: This image shows the substrate binding region (shown in cyan) in addition to the two key catalytic residues (purple) hydrogen bonded to the substrate analogue PDC.

4.1.6 Dimer Interfaces

The dimer interfaces of the protein are also important in stabilising the dynamics of the protein. One dimer interface is characterised by four-stranded β-sheets from each monomer, hydrogen bonding to form an eight-stranded mixed β-sheet (Figure 4.7). This pairs face to face with the other dimer interface in tetrameric species, forming a 16-strand mixed β-barrel. This structure is anchored by the other dimer interface loops and the outermost β-strand.

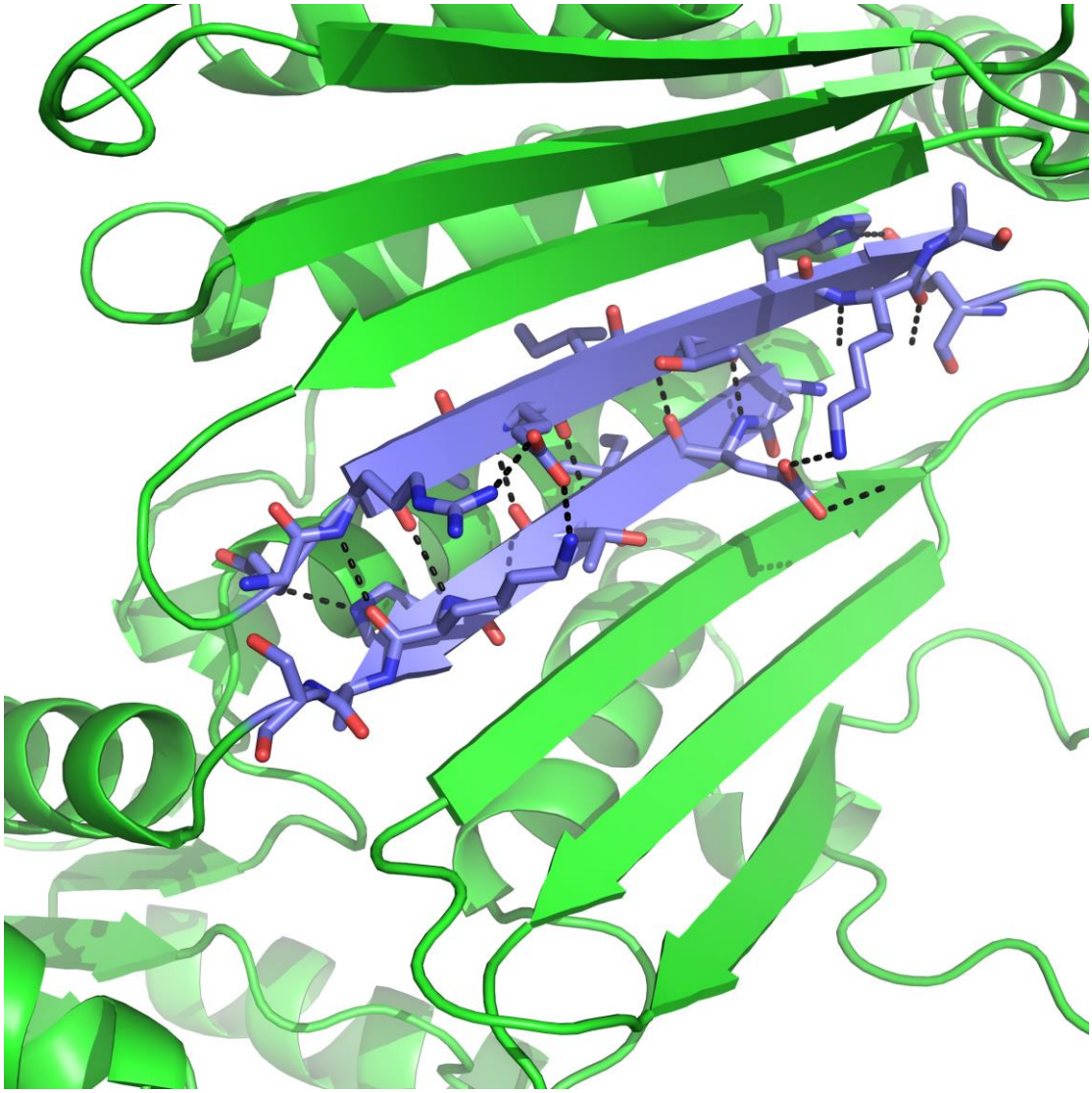


Figure 4.7: The β-sheet dimer interface. Residues 229-238 from each dimer are shown in blue.

The other dimer interface is created by two loops, Glu 195 – Pro 203 from one monomer and Val 164’ and Pro 167’ from another (Figure 4.8). These two loops form hydrogen bonding interactions which anchors the two subunits together. These two dimer interfaces combine to form a tetrameric structure.

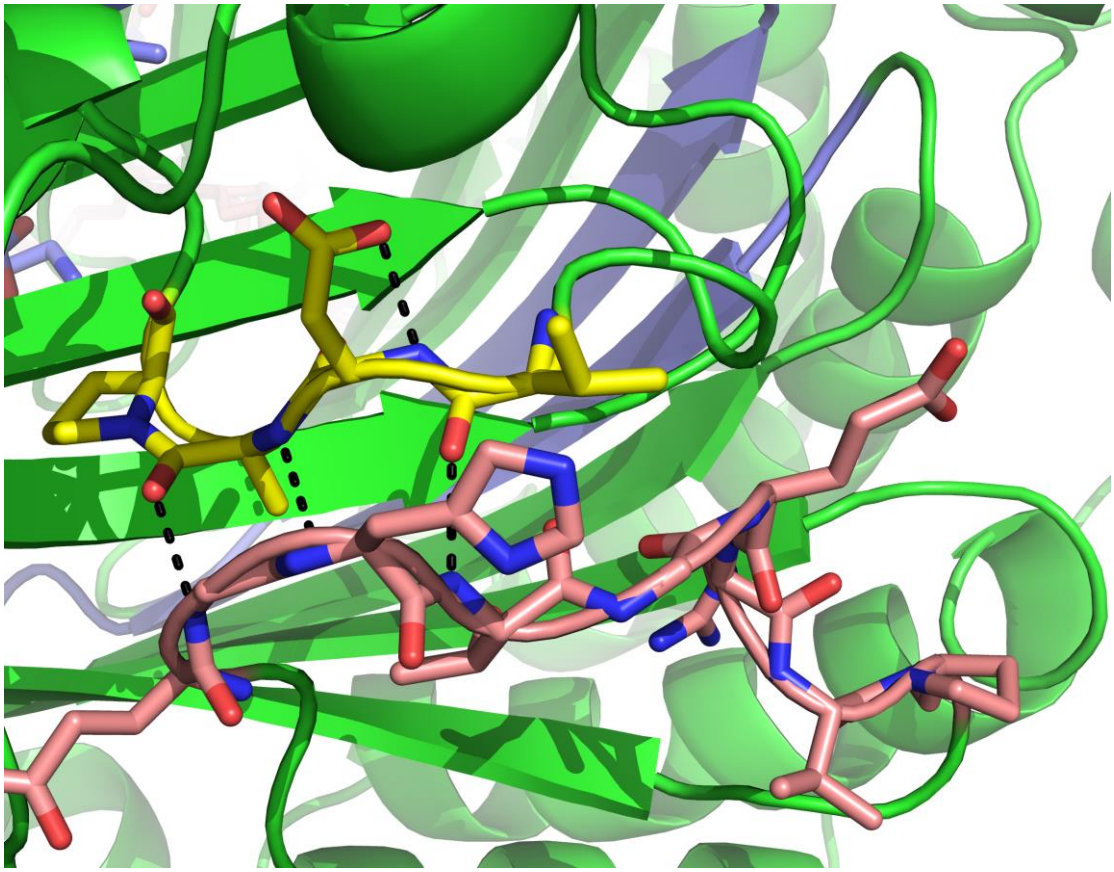


Figure 4.8: The loop binding interface and the two loops that form this interface: Glu 195 – Pro 203 (pink) and Val 164 and Pro 167 (yellow).

4.2 Evolutionary Divergence

4.2.1 Quaternary Structural Differences

DHDPR, like DHDPS, also exhibits quaternary structural differences between plant and bacterial proteins. All bacterial species characterised so far have been found to exist as tetramers (Cirilli et al., 2003, Girish et al., 2011, Janowski et al., 2010, Scapin et al., 1995). While characterisation of plant DHDPRs so far remains limited, *A. thaliana* and *Vitis vinifera* DHDPR have both been found to exist as dimers (Griffin et al., 2012, Watkin, 2014). This difference in quaternary structure is due to an evolutionary divergence at some point that made it favourable to switch to a dimeric structure but it is unknown exactly how or where this occurred.

4.2.2 Nucleotide Utilisation

DHDPR enzymes have varying affinities for nucleotide co-factors. Some bacterial DHDPR enzymes such as those from *E. coli* and *M. tuberculosis* have dual specificity for both NADH and NADPH (Cirilli et al., 2003, Reddy et al., 1996). Other bacterial DHDPR express a strong preference towards NADPH, such as that from *S. aureus* and *T. maritima* (Girish et al., 2011, Pearce et al., 2008). *A. thaliana* DHDPR has been shown to utilise NADPH, possibly due its association with the light reactions in plants (Griffin et al., 2012). These nucleotide associations may be linked to evolutionary advantages depending on the organism's environment.

Substrate inhibition occurs in many DHDPR enzymes. While most express a higher affinity to a specific nucleotide, high concentrations of substrate can lead to substrate inhibition whilst utilising that nucleotide. This effect has been observed in *S. auerus* and *T. maritima* among others (Dommaraju et al., 2011, Pearce et al., 2008). In both cases, high concentrations of substrate led to greatly increased levels of inhibition compared to the other nucleotide. However, the concentration required for this inhibition was not thought to be physiologically relevant. Substrate inhibition is thought to result from HTPA binding to the enzyme before the oxidised nucleotide has left its binding site, resulting in a dead-end complex (Dommaraju et al., 2011).

It has also been speculated that some forms of the enzyme can be inhibited by lysine (Tsujimoto et al., 2006). A DHDPR from *Methylophilus methylotrophus* was found to be inhibited by lysine with inhibition increasing along with lysine concentration. However, this effect has not been observed in any other DHDPR characterised so far making it uncertain whether this effect is observed in other DHDPR enzymes or is exclusive to this group of organisms.

DHDPR substrate inhibition is an aspect of the DAP pathway that is commonly ignored in the creation of higher lysine DAP pathways. With the many attempts that have been made at higher lysine organisms in bacteria and plants, most of these have neglected to take account of feedback inhibition when utilising a specific nucleotide. In bacteria that can utilise both nucleotides without a loss in rate or substrate affinity, this would not be a great deal. However, in bacteria, plants and algae that exhibit nucleotide induced substrate inhibition, increased levels of substrate may not necessarily lead to increased product. It is therefore important to ensure that the correct nucleotide is being used when the pathway is being manipulated to increase lysine production.

4.3 Phylogenetic Analysis

As DHDPR exists in two separate quaternary conformations in bacteria and plants, its evolution is of great interest in algal species. Phylogenetic analysis of DHDPR sequences would allow the proposed grouping of each algal structure to be estimated.

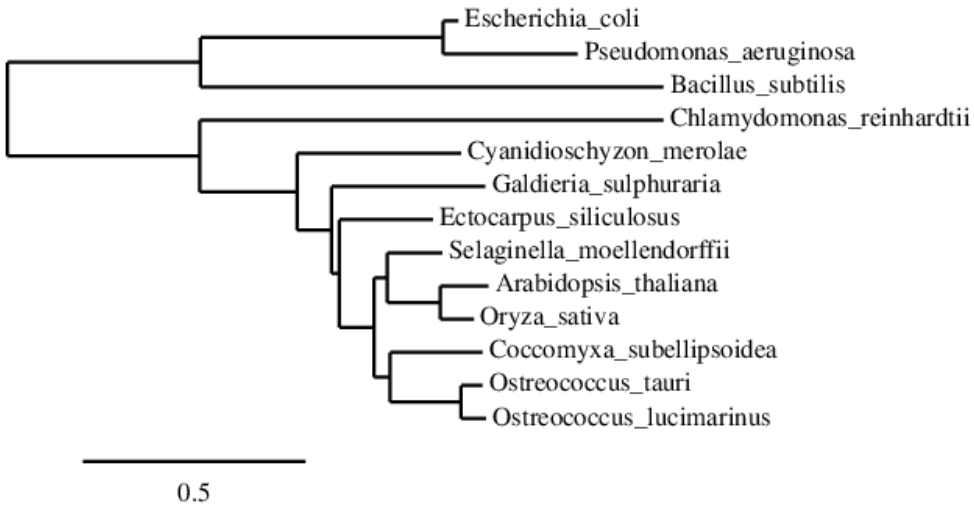


Figure 4.9: This is a phylogenetic tree created using www.phylogeny.fr

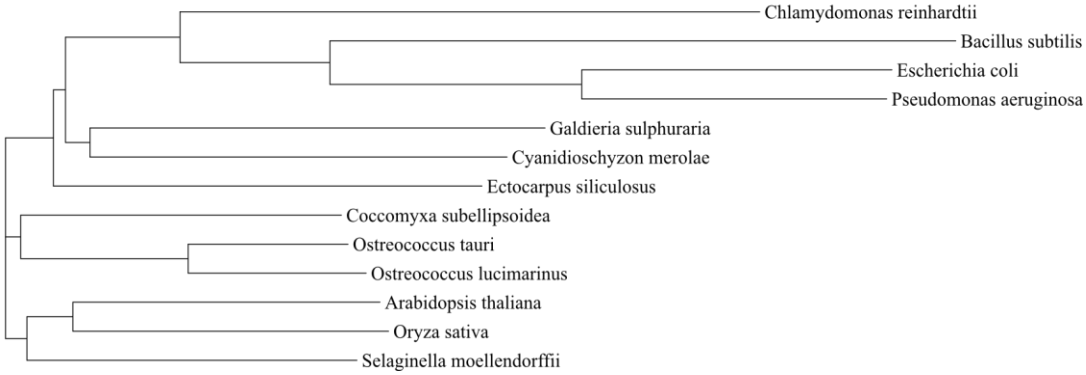


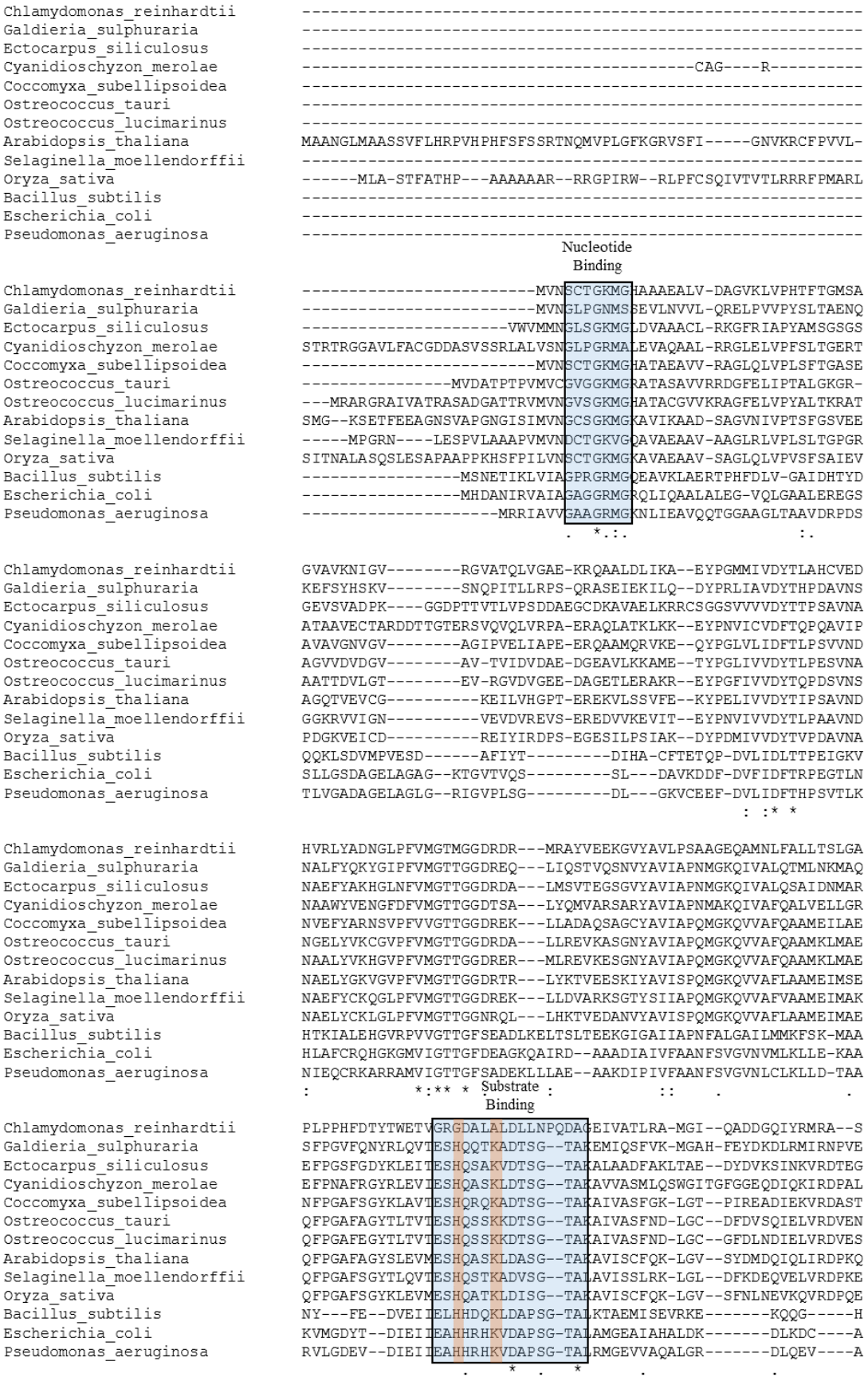
Figure 4.10: Another phylogenetic tree; created using www.ebi.ac.uk

Interestingly the two trees have significant differences (Figures 4.9 and 4.10). These two phylogenetic trees use slightly different sorting methods. This explains the slight differences between trees. In both instances, the bacterial and plant species are grouped together as expected. *B. subtillis*, *P. aeruginosa* and *E. coli* remain grouped together in both trees. Lycophytic and plant enzymes from *A. thaliana, O. sativa* and *S. moellendorffii* remain bunched together in a separate portion of the tree, showing that they are likely to have similar properties. Red algal species *G. sulphuraria* and *C. merolae* along with the brown algal species *E. siliculosus* remain bunched very close to each other in both trees. All of these species remain above plants in the tree making their quaternary structure uncertain between plant and bacterial formations. Characterisation of these proteins will allow the evolutionary nature of each of these DHDPR to be found.

Green algal species placement varies depending on the tree. In the phylogeny.fr tree, *O. tauri* and C. *subellipsoidea* exist after plant species but in the ebi.ac.uk tree, they exist before plants. Due to their close grouping to plants in either case, it can be hypothesised that they will retain similar characteristics to plant type DHDPR enzymes. What's more interesting, however, is the placement of the green algal *C. reinhardtii* DHDPR. In the phylogeny.fr tree, it is placed on its own; right after the bacterial species and before the algal species. In contrary to this, the ebi.ac.uk tree places the protein squarely amongst the bacterial species. The exact location of its placement is unknown but this discovery suggests that the *C. reinhardtii* DHDPR requires further investigation.

4.4 Sequence Alignment

To gain more information about specific residues of each protein, a multiple sequence alignment (MUSCLE) containing each amino acid sequence was created using Clustal Omega. This allows comparison of residues in equivalent positions between various DHDPR proteins.


Nucleotide Binding
Substrate Binding

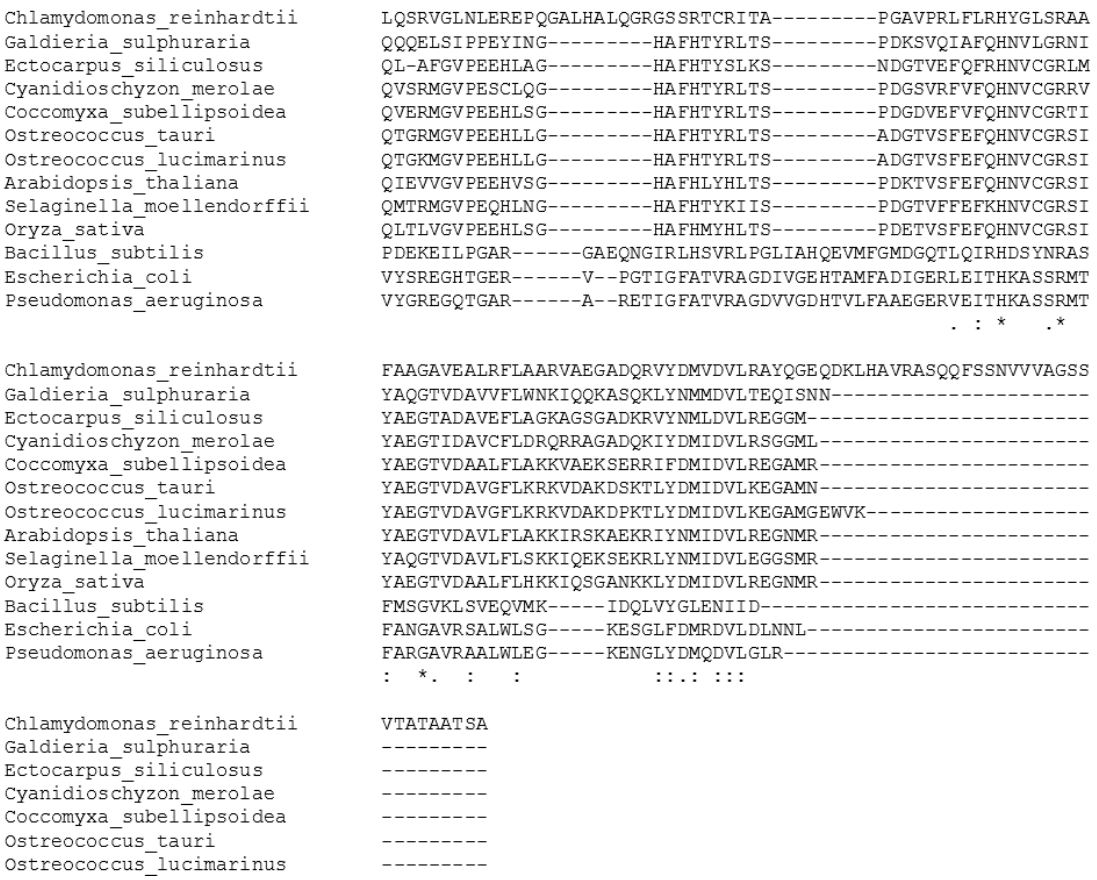


Figure 4.11: Sequence alignment for *Chlamydomonas reinhardtii, Galdieria sulphuraria, Ectocarpus siliculosus, Cyanidioschyzon merolae, Coccomyxa subellipsoidea, Ostreococcus tauri, Ostreococcus lucimarinus, Arabidopsis thaliana, Selaginella moellendorffii, Oryza sativa, Bacillus subtilis, Escherichia coli* and *Pseudomonas aeruginosa.*

Two regions in the protein are of particular importance. The GXXGXXG nucleotide binding motif is critical to NAD(P)H binding and is conserved among other nucleotide binding proteins such as Lactate Dehydrogenase and Ferrodoxin. While the motif is conserved in *A. thaliana, E. coli E. siliculosus* and *O. tauri*, in other algal and lycophyte protein sequences the motif is disrupted (Figure 4.11).

The first glycine residue is replaced by an aspartic acid in *S. moellendorffii.* This first glycine is highly conserved among DHDPR enzymes making it important in nucleotide binding. However, in *E. coli,* this residue does not make as many direct contacts with the NADH molecule as the critical second residue. This may mean that

it can be replaced by another residue such as an aspartic acid which may play a similar role.

The second glycine in the motif is completely conserved in every DHDPR analysed. This indicates that the residue is critically important in nucleotide binding and its positioning is highly specific; it hydrogen bonds to a hydroxyl on the ribose ring of NADH as well as an oxygen on the phosphate backbone of the nucleotide. It is likely that this residue plays a significant role in binding nucleotide to the protein in a position that NADH can donate its proton.

In the two red algal species characterised, the third glycine residue is replaced with another residue; in *G. sulphuraria* it is swapped out for a serine, whereas in *C. merolae*, it is replaced with an alanine. Like the first glycine residue, this does not make as much contact with the NADH molecule as the critical second glycine residue. This may mean that the serine and alanine replacements can fulfil a similar role to the glycine in nucleotide binding.

The substrate binding loop containing residues Glu 157 to Ala 171 is another sequence of interest. In all DHDPR enzymes characterised thus far, two residues (His 159 and Lys 163) are thought to be critical to the mechanism. While they are present in all other sequences analysed, they are not present in *C. reinhardtii.* This contrasts with the proposed mechanism which uses a Schiff base formed by lysine that stabilises the substrate along with a histidine electron acceptor. However, in *C. reinhardtii* these residues are replaced with an alanine and glycine respectively. These residues do not contain the side chain functionality that their counterparts do so *C. reinhardtii* must utilise alternative residues with a similar function, or a different mechanism. It is also interesting to note that the rest of the binding region only retains 2 of the 15 other residues from the *E. coli* substrate loop and contains an additional residue. These two retained residues are not thought to directly contribute to substrate binding but may function in stabilisation of the substrate.

Another thing that is interesting to note about the *C. reinhardtii* DHDPR sequence is the presence of an extra cysteine residue. It is possible that this extra residue may play a role in a disulfide-dependent dimer interface. However, the exact residues that form

this interface are unknown. This extra residue matches with *E. coli* DHDPR in a region that looks unlikely to form a disulfide bond. This may mean that in *C. reinhardtii* DHDPR, the residue is orientated in a position that could allow a disulfide interface to occur. It is also possible that the other conserved cysteine residues are orientated in positions that allow a disulfide bond to occur. The properties of this enzyme are explored further in the *C. reinhardtii* DHDPR chapter.

4.5 Enzyme Kinetics

To test the kinetics of each DHDPR enzyme, enzyme activity assays were performed. This gave an insight into the nucleotide utilisation of each enzyme as well as substrate affinity and inhibition (Table 4.1).

Table 4.1: The kinetic parameters for each organism while utilising each nucleotide.

Organism	**Nucleotide**	**K_m (mM)**	**V_{rel} (µmol/mg/s)**	**V_{rel}/K_m**	**K_i (mM)**
E. coli	NADH	0.32±0.055	0.0047±0.00036	0.015	
E. coli	NADPH	0.29±0.00010	0.0079±0.00010	0.027	
A. thaliana	NADH	0.73±0.16	0.0046±0.00067	0.0063	2.20±0.60
A. thaliana	NADPH	0.62±0.088	$0.0011\pm5.414\times10^{-5}$	0.0018	
T. maritima	NADH	$8.67\times10^{-14}\pm0.014$	0.00023±0.00090	2.62	0.93±0.34
T. maritima	NADPH	0.051±0.0055	$0.028\pm2.743\times10^{-5}$	0.56	24.60±14.30
E. siliculosus	NADH	0.099±0.021	$8.433\times10^{-5}\pm8.317\times10^{-6}$	0.00085	1.23±0.28
E. siliculosus	NADPH	0.11±0.0087	$6.932\times10^{-5}\pm2.253\times10^{-6}$	0.00061	3.67±0.44
C. subellipsoidea	NADH	0.81±0.15	0.017±0.0014	0.022	
C. subellipsoidea	NADPH	0.87±0.15	0.017±0.0011	0.017	
O.tauri	NADH	0.14±0.024	0.001869±0.00014	0.013	7.34±3.02
O.tauri	NADPH	0.39±0.053	$0.0013\pm4.943\times10^{-5}$	0.0032	
O. lucimarinus	NADH	0.00050±0.00014	$0.0001877\pm3.101\times10^{-6}$	0.38	

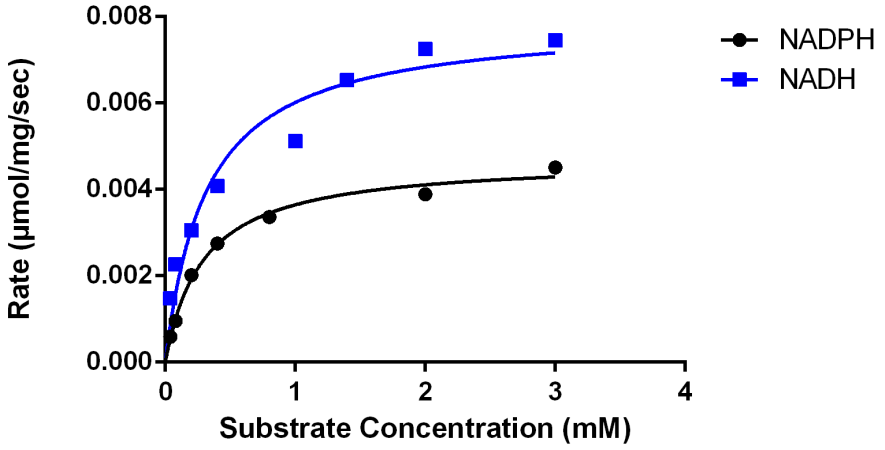


Figure 4.12: Kinetics for *E. coli* DHDPR. The data was fitted to a standard Michalis-Menten kinetic profile.

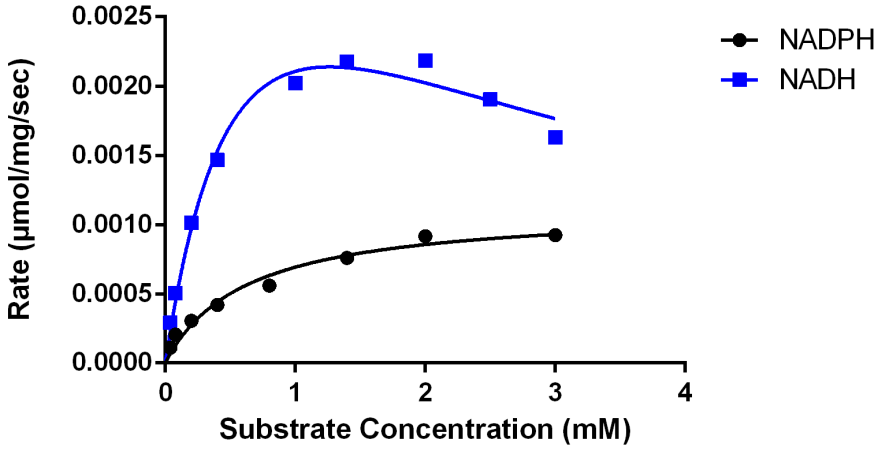


Figure 4.13: The kinetics for *A. thaliana* DHDPR. The NADPH data was fitted to a standard Michalis-Menten kinetic profile whereas the NADH data was fitted to a substrate inhibition model.

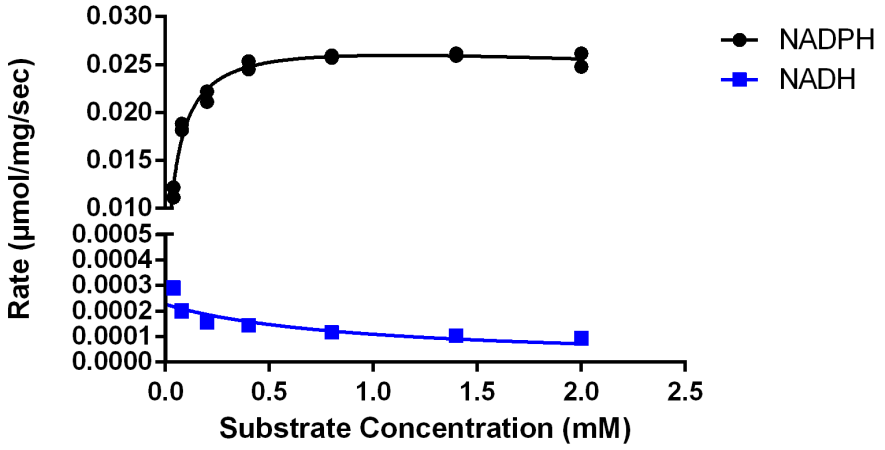


Figure 4.14: The kinetics for *T. maritima*. Both curves were fitted using a substrate inhibition model.

Three enzymes were used as controls, the first of which was *E. coli* DHDPR (Figure 4.12). This typically has a greater affinity for NADPH but can utilise both nucleotides (Coulter et al., 1999). In this case, the K_m values are similar but the V_{rel} is higher when NADPH is utilised. This kinetic profile will be used as one of the models for kinetics for a bacterial DHDPR.

The other bacterial model that will be used is *T. maritima* DHDPR (Figure 4.13). This uses a different model of nucleotide utilisation which expresses a much stronger substrate affinity and rate whilst utilising NADPH (Pearce et al., 2008). With NADH, it exhibits extremely strong substrate inhibition and a rate many times lower than that with its phosphorylated counterpart.

A. thaliana DHDPR has a noticeably different kinetic profile than the two bacterial models (Figure 4.14). It expresses a similar affinity for both NADH and NADPH (Griffin et al., 2012). However, it expresses a higher rate along with substrate inhibition while utilising NADH. As the reaction takes place in the chloroplast, it is likely that NADPH is the biologically relevant co-factor (Griffin et al., 2012).

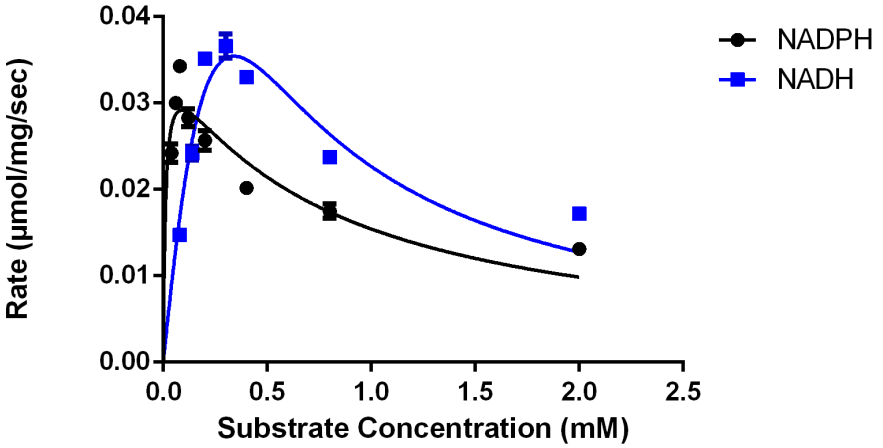


Figure 4.15: The kinetics for *C. merolae* DHDPR. Both curves were fitted using a substrate inhibition model.

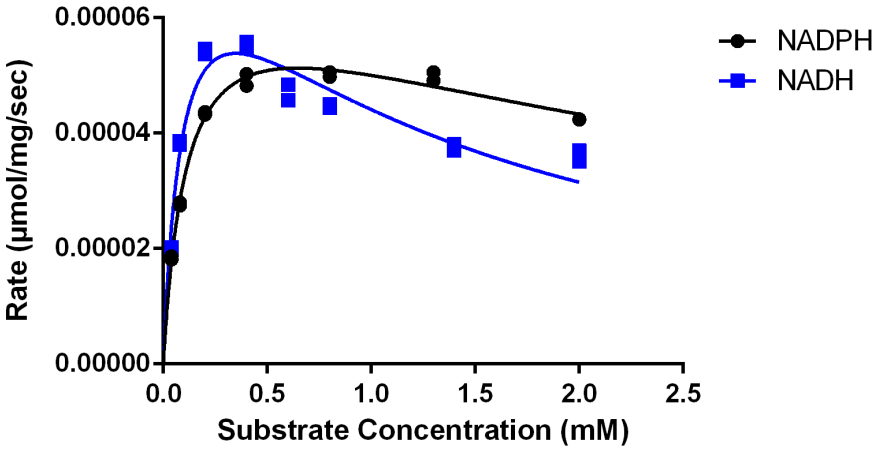


Figure 4.16: The kinetics for *E. siliculosus* DHDPR. Both curves were fitted using a substrate inhibition model.

Both red and brown algal species characterised have similar kinetic profiles. The red algal DHDPR from *C. merolae* may be able to utilise both substrates *in vivo*. It has a much higher affinity for NADPH; however, it has a greater maximum rate whilst utilising NADH (Figure 4.15). In each case, a novel substrate utilisation profile is observed, with severe substrate inhibition observed while utilising both nucleotide substrates. However, this may not usually be the case due to the stability of the enzyme. The enzyme was unstable over long periods of time and the substrate inhibition may have been caused by protein degradation causing formation of the dead-end enzyme complex responsible for substrate inhibition. It could be speculated

that as red algae contain chloroplasts, they utilise NADPH due to the enzymes location in the chloroplast where NADPH is ubiquitous due to its role in the light reactions. The brown algal DHDPR from *E. siliculosus* exhibits a similar kinetic profile to that of *C. merolae* (Figure 4.16). While its substrate affinity appears similar between both nucleotides, it exhibits a slightly higher rate whilst using NADH rather than NADPH. Substrate inhibition is observed in both cases but is more pronounced in NADH. This reinforces the idea proposed by the evolutionary lineage that the brown algal species are closely related to their red algal counterparts.

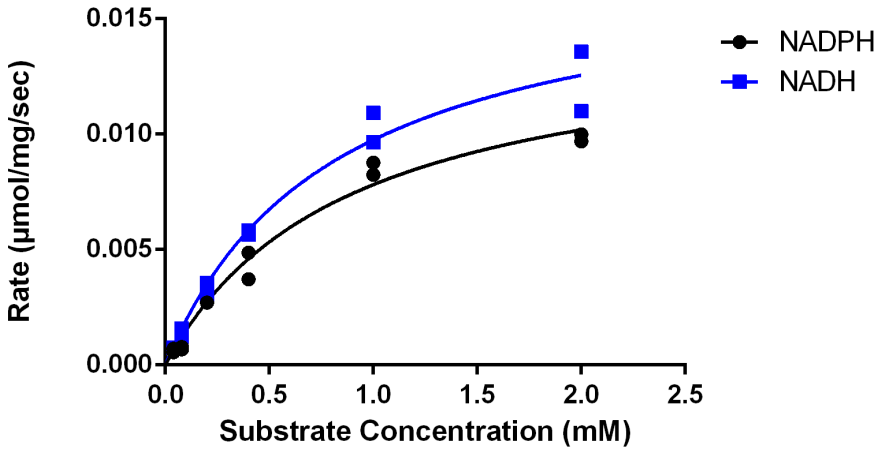


Figure 4.17: The kinetics for *C. subellipsoidea* DHDPR. Both data sets were fitted to a standard Michalis-Menten kinetic profile.

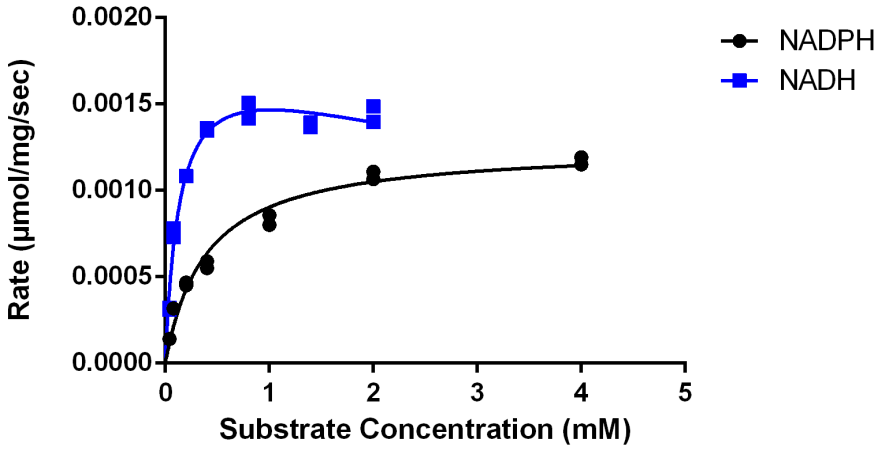


Figure 4.18: The kinetics for *O. tauri* DHDPR. The NADPH data was fitted to a standard Michalis-Menten kinetic profile whereas the NADH data was fitted to a substrate inhibition model.

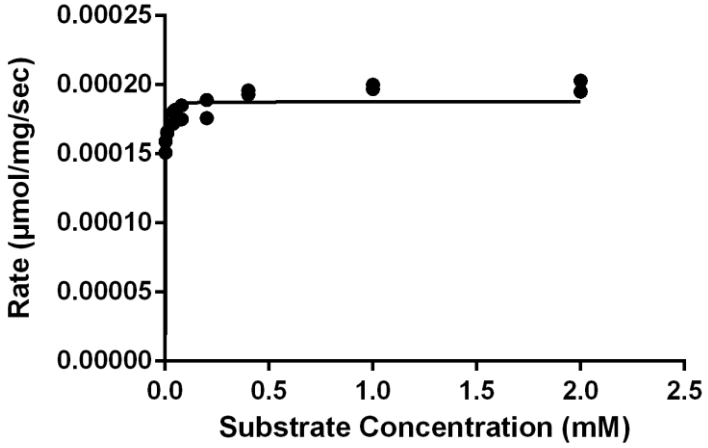


Figure 4.19: The kinetics for *O. lucimarinus* DHDPR. The NADPH data was fitted to a standard Michalis-Menten kinetic profile whereas the NADH data was fitted to a substrate inhibition model. Data for NADPH is not shown as the protein was unreactive when this nucleotide was used.

As for the green algal species, *C. subellipsoidea* shows a similar kinetic profile to *E. coli* DHDPR (Figure 4.17). This is characterised by a complete lack of substrate inhibition and an extremely similar affinity towards both nucleotides. This contrasts with *O. tauri* DHDPR which exhibits a much stronger affinity and higher rate whilst utilising NADH rather than NADPH (Figure 4.18). However, substrate inhibition is much more prevalent whilst utilising NADH. This kinetic profile is similar to *A. thaliana* DHDPR.

O. lucimarinus has another unique kinetic profile. This utilises NADH with an extremely tight affinity with minimal substrate inhibition (Figure 4.19). However, when NADPH was utilised, no increase from a minimal level of activity was observed. This profile appears to be similar to *T. maritima* except with a reversed nucleotide utilisation with NADH being the main nucleotide. The *O. lucimarinus* DHDPR protein sequence contains all of the catalytic residues responsible for catalytic activity so it is uncertain exactly why it does not utilise NADPH. However, it is possible that the sample was contaminated with Lactate Dehydrogenase. Lactate Dehydrogenase catalyses a reaction between pyruvate to lactate and back again using specifically NADH as a co-factor. It is possible that the small amount of this protein had entered

the sample which would explain the lack of observed activity when NADPH was used as the nucleotide co-factor. It is also possible that the protein concentration was not sufficient to observe enzyme activity. High levels of substrate inhibition may have meant that an observable rate was simply not possible at the enzyme assay concentration.

The exact kinetic parameters for these enzymes is difficult to characterise for many of the algal DHDPR proteins. For example, many of them are extremely unstable and lose measurable activity very quickly. Due to this fact, it can be challenging to find an accurate maximum rate of the enzyme. This means that the rates between enzymes cannot readily be compared.

Another reason the rates cannot be trusted between enzymes is an issue with calculation of the concentration of the protein. For all calculations of protein concentration, a Nanodrop was used. This is a device that measures light absorbance at 280 nm; the wavelength at which the peptide bond absorbs. However, a much larger peak at 260 nm was observed which renders the peak at 280 nm distorted. Variables that can modify absorption at 260 nm include DNA, nucleotides and changes in acidity. As this protein binds nucleotides as part of its mechanism, it is likely that this contamination is the caused by binding of nucleotide during purification.

The contamination meant that while the V_{rel}'s within the same protein sample can be compared, interspecies comparison of rates may not necessarily be possible. This would not affect the K_m however, which is independent of protein concentration. This certainly can be compared between species and is an important tool in characterisation of the algal DHDPR's.

4.6 Quantitative Size Exclusion Chromatography

To analyse the molecular weights of each protein, quantitative size exclusion chromatography was used. This involved loading a set volume of protein onto a size exclusion column. As higher volume objects elute before lower volume ones, the

protein weights can be estimated to determine whether they are in a tetrameric or dimeric state (Table 4.2).

Table 4.2: DHDPR elution time on a size exclusion column.

Organism Name	**Organism Type**	**Elution Volume (ml)**	**Theoretical Mass (Da)**	**Proposed Quaternary State**
T. maritima	Bacteria	10.12	94,424	Tetramer
A. thaliana	Plant	14.03	75,740	Dimer
C. subellipsoidea	Green Algae	14.07	58,750	Dimer
O. lucimarinus	Green Algae	17.79	64,602	Monomer
O. tauri	Green Algae	13.25	59,884	Dimer
C. reinhardtii	Green Algae	10.91/12.50	133,716/66,858	Tetramer/Dimer
C. merolae	Red Algae	13.51	68,094	Dimer
G. sulphuraria	Red Algae	13.49	62,552	Dimer
E. siliculosus	Brown Algae	13.24	59,574	Dimer

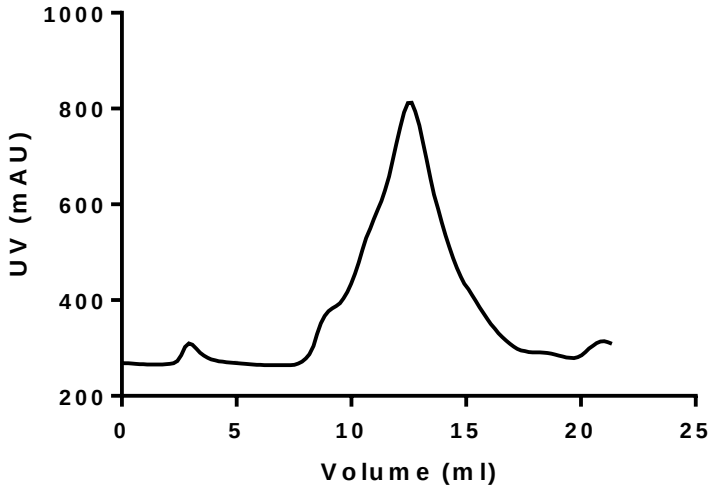


Figure 4.20: A sample SEC UV trace from the experiment. The two major peaks that exhibited activity are shown; at 10.97 ml peak with a height of 550 mAU residing on the side of the main peak and the main peak at 12.50 ml which is 800 mAU high.

To be used as controls, tetrameric *T. maritima* and dimeric *A. thaliana* DHDPR enzymes were the first to be characterised using the column. This allows the remaining DHDPR enzymes to be compared to these benchmarks. Two of the green algal DHDPR enzymes were found to elute at a similar volume to the plant DHDPR enzymes. This is evidence that they may exist in a dimeric conformation. The other green algal DHDPR from *C. reinhardtii* had two peaks that exhibited activity corresponding to dimer and tetramer (Figure 4.20). The red algal species were found to exist at very similar elution volumes to the plant type dimer indicating that they may exist in a dimeric conformation. The brown algal species *E. siliculosus* also elutes at a similar volume to other dimers, indicating that it also exists as a dimer.

The results indicate that almost all the algal species characterised appear to exist in a dimeric formation. This corresponds to the proposed evolutionary lineage in which the algal species tested apart from *C. reinhardtii* are hypothesised to exist in a dimeric conformation. The exception was the green algal DHDPR from *O. lucimarinus.* As this enzyme was unstable during purification and kinetic assays, it is possible that the enzyme lost its quaternary structure and exists in a monomeric form. This form could still be catalytically active albeit at a much lower rate which was observed in the kinetic assays.

The reason that this technique was used instead of more accurate techniques such as AUC or size exclusion chromatography – multi angle light scattering (SEC-MALS) is that many of these proteins do not have sufficient stability or purity to be analysed using these techniques. AUC requires a sample purity of 95% which is not necessarily possible to achieve with some of these algal proteins despite a two-step His Tag and size exclusion chromatography column purification. This technique was used over SEC-MALS due to the sensitivity of the technique to aggregates and contaminants. As both are prevalent in these algal species, an upscaled version of the technique was used to characterise the proteins. This technique required more labour than a traditional SEC-MALS experiment but due to the lower resolution of the column, allowed specific fractions to be characterised for activity. This led to the formation of a technique that was much more applicable to the situation.

4.7 Small Angle X-ray scattering

Conformation of the quaternary structure of DHDPR enzymes can be found using small angle X-ray scattering. This allows determination of the size and shape of DHDPR enzymes. To be used as controls, both tetrameric *E. coli* and dimeric *A. thaliana* DHDPR enzymes were used as references for *C. merolae* quaternary structure determination (Figures 4.21 and 4.22).

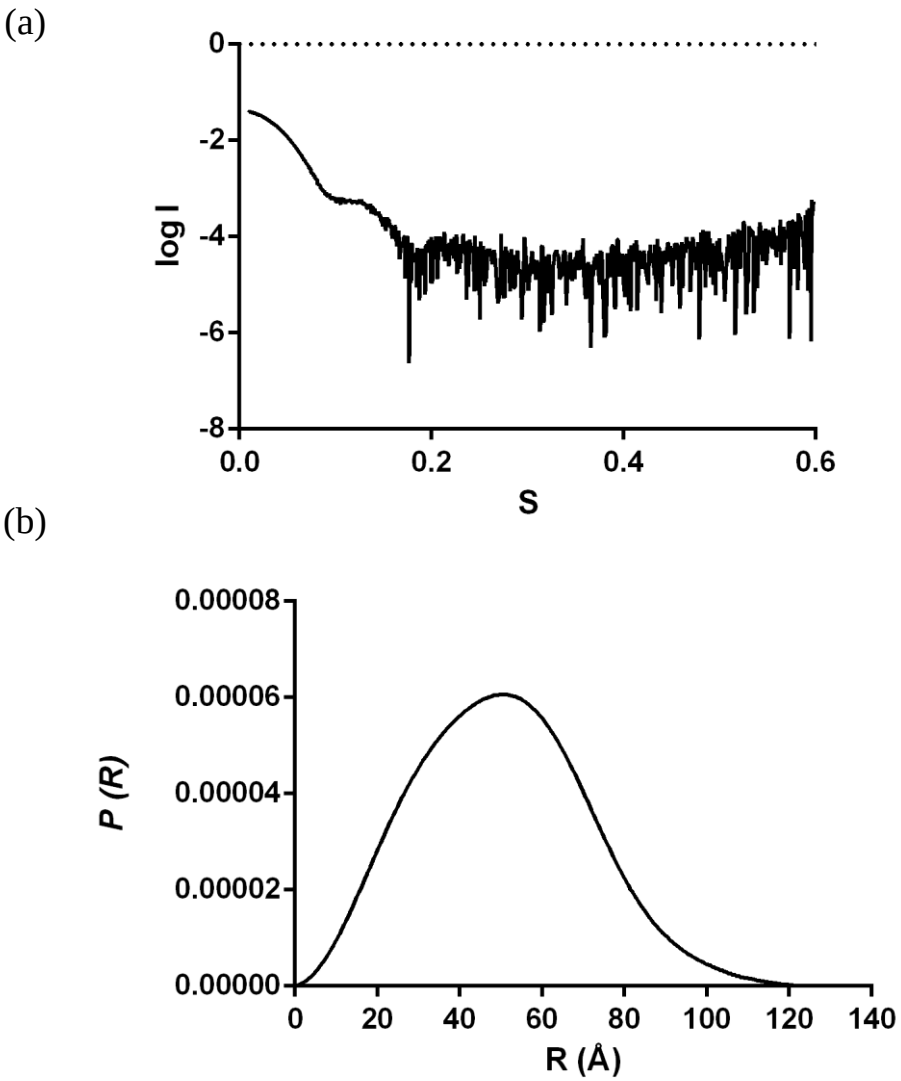
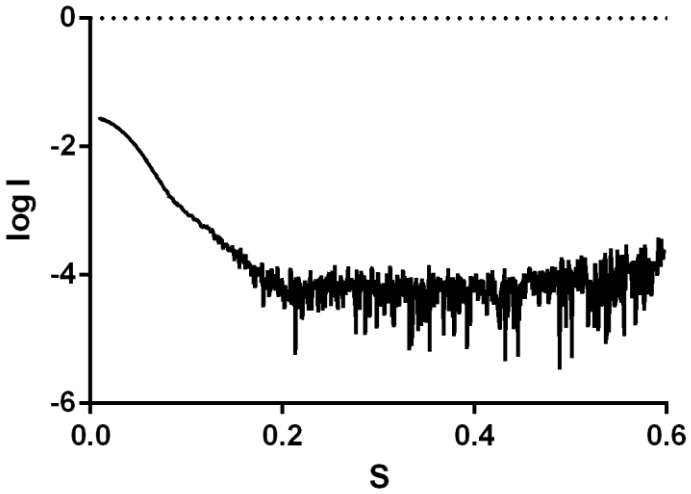


Figure 4.21: Intensity plot (log I vs S) (a) and distance distribution plot (P (R)) (b) for *E. coli* DHDPR.

(a)



(b)

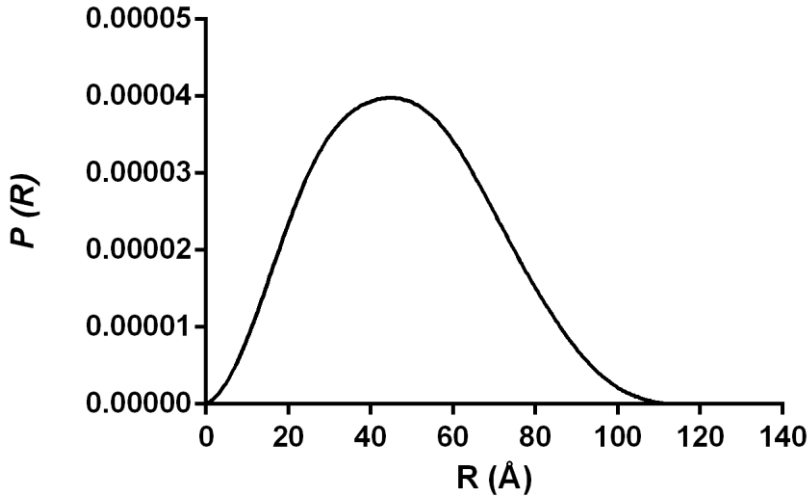
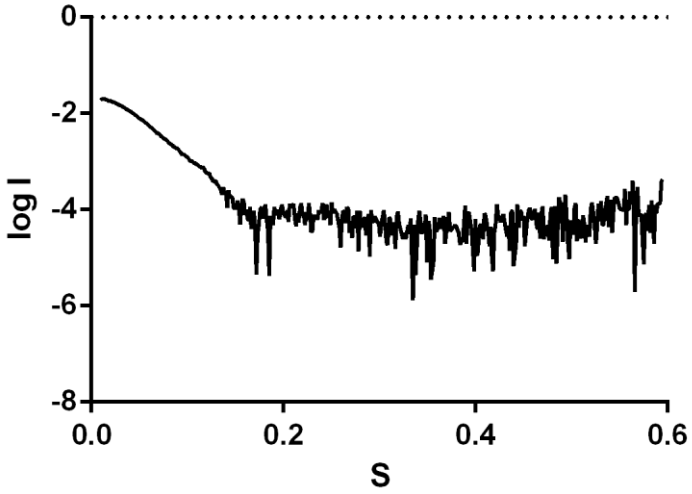


Figure 4.22: Intensity plot (log I vs S) (a) and distance distribution plot (P (R)) (b) for *T. maritima* DHDPR.

(a)



(b)

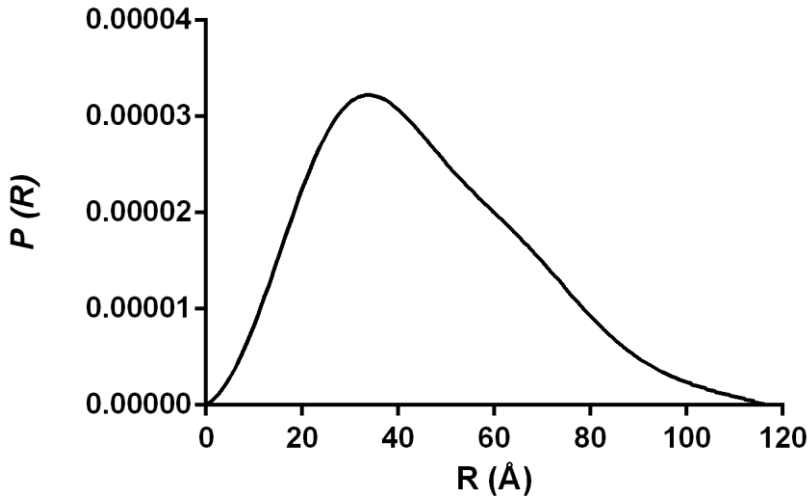


Figure 4.23: Intensity plot (log I vs S) (a) and distance distribution plot (P (R)) (b) for *C. merolae* DHDPR.

Table 4.3: The parameters for the SAXS analysis.

Organism Name	**Rg (Å)**	**Io**	**Porod Volume (Å^3)**
A. thaliana	35.06±0.47	0.048	128,525
T. maritima	37.16±0.24	0.049	199,135
E. coli	39.88±3.59	0.046	241,859
C. merolae	35.14±0.66	0.021	120,226

C. merolae DHDPR *P(R)* plot indicates a dimeric structure in solution (Figure 4.23). Like the *A. thaliana* scatter pattern, it exists in a non-standard distribution with a

significant skew towards the left side. This indicates that the protein does not have an even distribution of mass around the centre of the protein meaning that it likely exists in a dimeric formation. Scatter patterns can also can be compared to *E. coli* and *A. thaliana* controls to deduce which quaternary state the protein conforms to (Table 4.3). It's radius of gyration compared favourably to the *A. thaliana* dimeric value as opposed to the much higher *E. coli* and *T. maritima* tetrameric values. This is also backed up the Porod volume which is similar to the *A. thaliana* value. These factors indicate that *C. merolae* DHDPR exists in a dimeric formation.

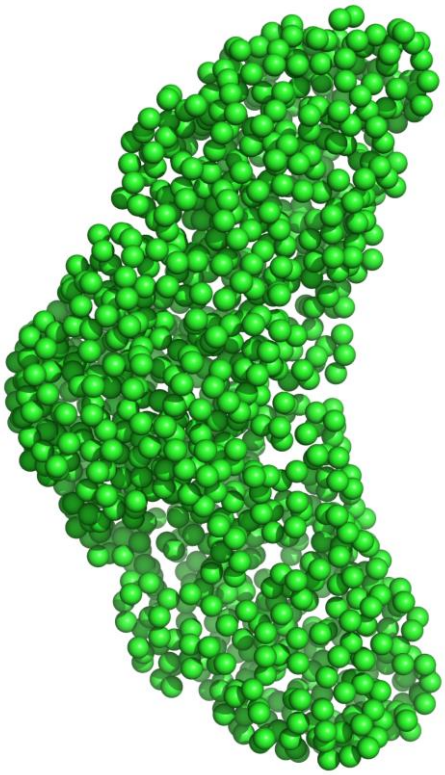


Figure 4.24: GASBOR model for *C. merolae* (Chi^2 fit = 0.25)

These observations are also observed in the *ab initio* modelling of the scatter data (Figure 4.24. A dimeric structure reminiscent of a half an *E. coli* DHDPR is observed adding to the evidence that *C. merolae* DHDPR exists in a dimeric formation.

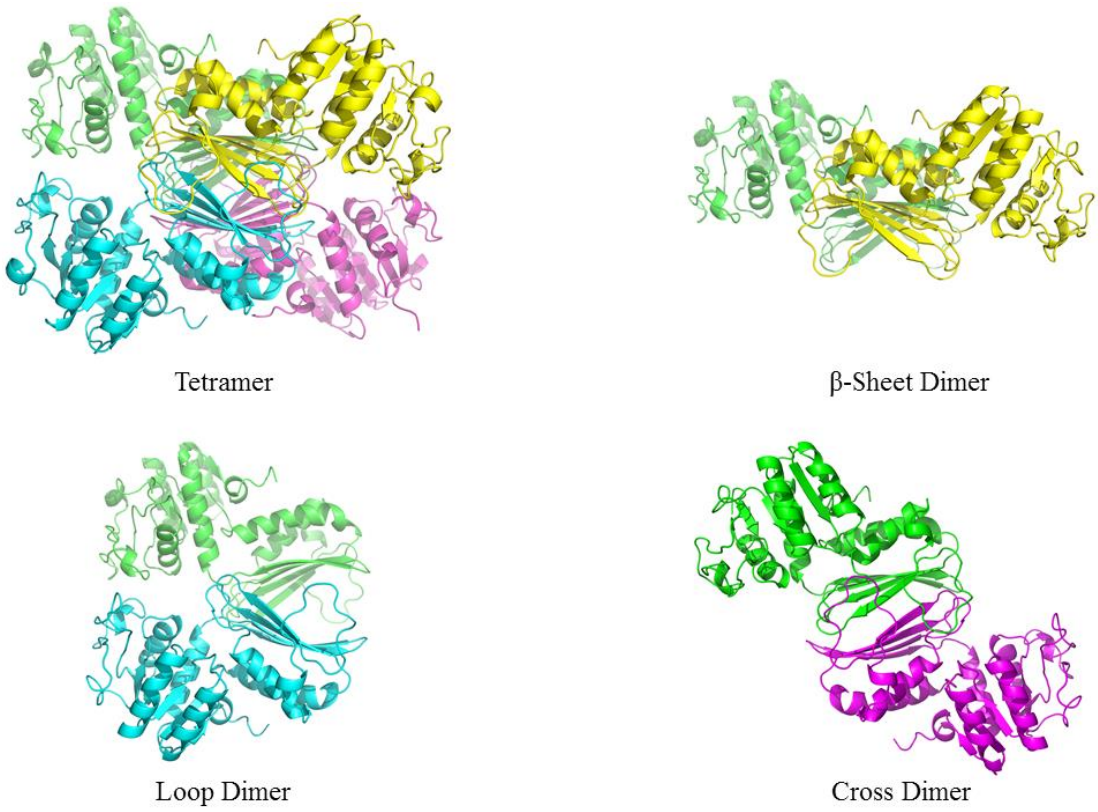


Figure 4.25: Possible formations for a dimeric DHDPR.

If *C. merolae* DHDPR does exist in a dimeric formation, it is uncertain exactly which orientation the dimer forms. It could form one of 3 different dimers (Figure 4.25). The most likely configuration is the "β-sheet dimer". In this conformation, two β-sheets pair together to form an eight stranded β-sheet. Another possible conformation is known as the "loop dimer". This is formed by two loops that hydrogen bond with each other. The other possible dimer is known as the "cross dimer". This dimer forms a diagonal conformation across the former tetramer with hydrogen bonds forming a β-sheet "sandwich" which has a length half of the full tetrameric equivalent version. The last two do not show the features of the SAXS model with a large gap in between the two monomers. This is much more indicative of a β-sheet dimer than the other two forms however these forms cannot be ruled out.

4.8 Summary

Phylogenetic tree analysis showed that both red and green algal DHDPR enzymes cluster around plant organisms indicating that they may be more closely related to the plant type protein than the bacterial. Green alga *Chlamydomonas reinhardtii* DHDPR was found to exist in varying positions on the evolutionary lineage but may exist between bacterial and plant type species. Sequence analysis showed that while almost all other DHDPR sequences retained residues thought to be critical to catalysis, *C. reinhardtii* did not contain these residues making its mechanism unknown. In the quantitative size exclusion experiment, almost all algal species characterised were found to exist as dimers. This fits with the phylogenetic tree which indicates that these species exist after *A. thaliana* in the evolutionary lineage. This could indicate an evolutionary divergence point between bacterial and plant type DHDPR enzymes.

Kinetic assays showed that was no clear patterns within green algal species with varying levels of substrate inhibition along with nucleotide and substrate affinities observed. Red and brown algal species were found to act similarly to each other with each species observing a similar affinity for each nucleotide with substrate inhibition. The exception to this was *O. lucimarinus* which exhibited an extremely strong affinity for NADH but no measurable activity whilst utilising NADPH. In *C. merolae* DHDPR, small angle X-ray scattering showed that the dimeric arrangement is thought to exist in a "β-sheet" dimeric conformation similar to the hypothesised *A. thaliana* conformation.

Chapter 5 - *C. reinhardtii* DHDPR

5.1 Overview

C. reinhardtii is green alga existing worldwide among soil and fresh water. Due to the dimer-tetramer equilibrium of the *C. reinhardtii* DHDPR and possible position between bacteria and plant enzymes in the evolutionary lineage, it may serve as a divergent DHDPR enzyme. To establish more information about the properties of *C. reinhardtii* DHDPR, several characterisation techniques were used.

5.2 Purification and Expression

The plasmid was purified transformed into *E. coli* BL21 DE3 cells using standard lab techniques and the enzyme was purified using the techniques described in the methods with a 19.7 mg yield from 1.6 L. The protein was stored in “high salt” 20 mM Tris and 500 mM NaCl buffer and frozen until use. On an SDS PAGE gel, the protein shows two separate bands (Figure 5.1. As this process should remove any quaternary interactions, the fact that two bands exist shows that a disulfide bond is present in solution.

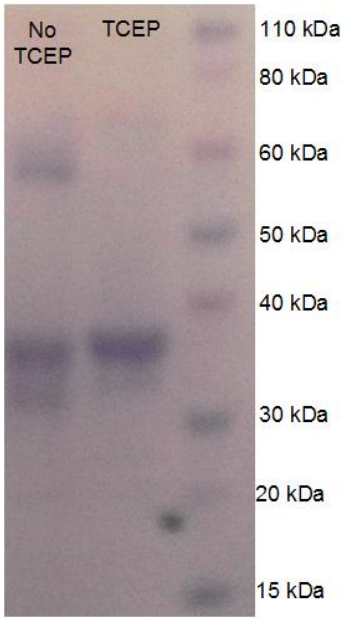


Figure 5.1: SDS-PAGE gel of *C. reinhardtii* DHDPR in its native form and in the presence of 5 mM of tris(2-carboxyethyl)phosphine (TCEP).

5.3 Analytical Ultracentrifugation

AUC was used to more effectively characterise the quaternary state of *C. reinhardtii* DHDPR. This technique allowed the molecular weights of the species in solution to be found.

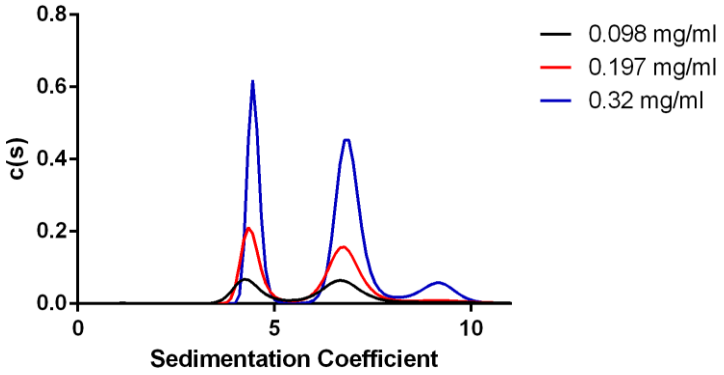


Figure 5.2: Sedimentation of *C. reinhardtii* DHDPR at 3 different concentrations.

C. reinhardtii DHDPR exists in two quaternary states (Figure 5.2. One peak exists at ~4.5 S with the other residing at ~6.8 S. These two peaks exist at theoretical molecular weights of ~60,000 Da and ~110,000 Da which approximately correspond to dimeric and tetrameric species. Interestingly, in the highest concentration of protein, a peak is observed at ~9 S which equates to a molecular weight of ~185,000 Da. This may indicate the presence of a small amount of hexameric protein in solution, which may indicate three dimeric species joining together. This, along with quantitative size exclusion, reinforces the idea that *C. reinhardtii* DHDPR exists in a dimer-tetramer equilibrium, with a small amount of possible hexamer at higher concentrations.

The *C. reinhardtii* DHDPR equilibrium is modified in the presence of reducing agent (Figure 5.3). This is also expressed in AUC data in the presence of TCEP. With no reducing agent present, a typical dimer-tetramer equilibrium is observed. However, when 2 mM TCEP was added, the equilibrium shifts almost exclusively to the dimeric form. This has not been observed in any other DHDPR so far making it the first disulfide-dependent dimeric interface observed in DHDPR enzymes.

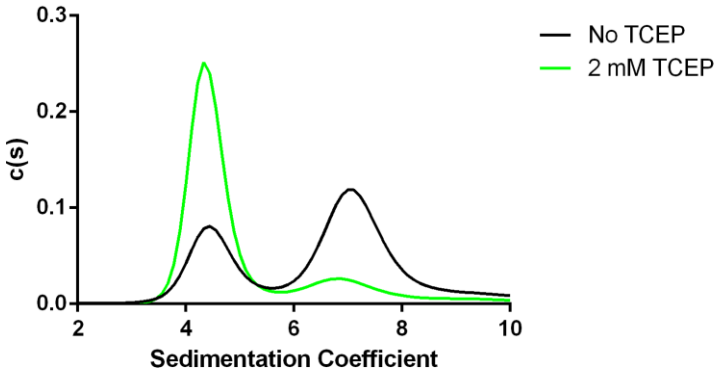


Figure 5.3: 0.5 mg/ml of *C. reinhardtii* DHDPR with and without the presence of reducing agent.

It is uncertain how this mechanism occurs. The sequence of *C. reinhardtii* DHDPR contains three cysteine residues instead of the regular two that are conserved in all other DHDPR enzymes so far. However, these residues do not appear to be in positions that would allow a disulfide dimer interface to occur. It is possible that the disulfide interface exists on the outside of the protein and is not naturally involved in the protein. Rather, it may exist as a "daisy chain" dimer where disulfide bonds are formed on the outside of the protein which may appear in solution as a tetramer or even a hexamer depending on protein concertation.

While DHDPS in plants has been known to exist in the chloroplast, the location of DHDPR in non-bacterial organisms has not been determined (Ghislain et al., 1990). However, the presence of a transit peptide; a key sequence important for protein entry into the chloroplast indicates that *C. reinhardtii* DHDPR exists in the chloroplast. This would likely mean that the enzyme would exist in a non-reduced form *in vivo* as opposed to an exclusively dimeric form in the reducing environment of the cytoplasm where the bacterial DAP pathway takes place.

5.4 Enzyme Kinetics

To test the kinetic parameters of *C. reinhardtii* DHDPR, kinetic assays were performed. This was carried out not only in no presence of reducing agent but also in

the presence of 2 mM TCEP. This ensures that any changes in kinetics for the reduced form of the enzyme are observed (Table 5.1).

Table 5.1: Kinetic parameters for *C. reinhardtii* DHDPR.

Buffer	**Nucleotide**	**K_m (mM)**	**V_{rel} (µmol/mg/s)**	**V_{rel}/K_m**	**K_i (mM)**
No TCEP	NADH	0.45±0.26	0.079±0.033	0.17	0.33±0.33
No TCEP	NADPH	0.035±0.012	0.050±0.0041	1.4	23.54±33.00
2 mM TCEP	NADH	0.45±0.26	0.079±0.033	0.21	0.55±0.33
2 mM TCEP	NADPH	0.015±0.0088	0.072±0.0073	4.74	3.37±1.54

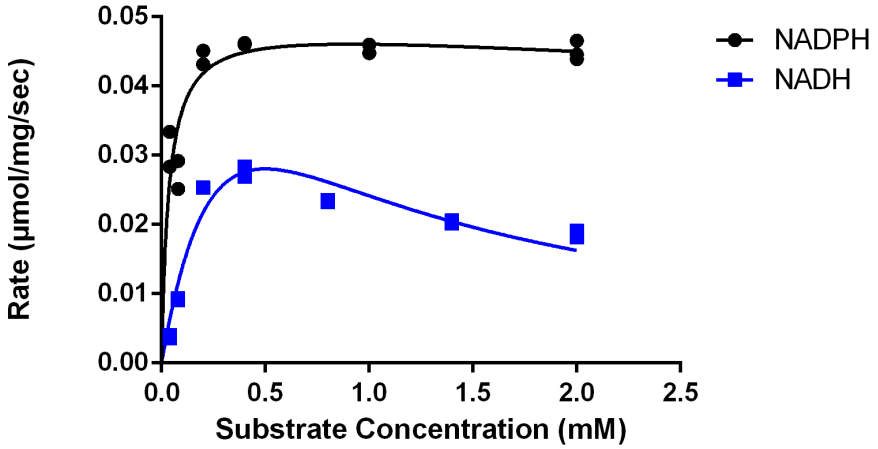


Figure 5.4: Kinetics for *C. reinhardtii* utilising both NADPH and NADH. Both data sets were fitted to a substrate inhibition model.

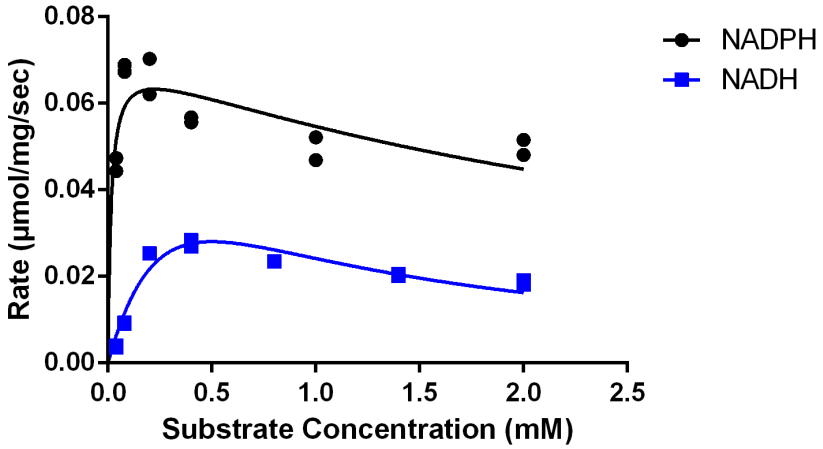


Figure 5.5: Kinetics for *C. reinhardtii* in the presence of reducing agent TCEP. Both curves were fitted using a substrate inhibition model.

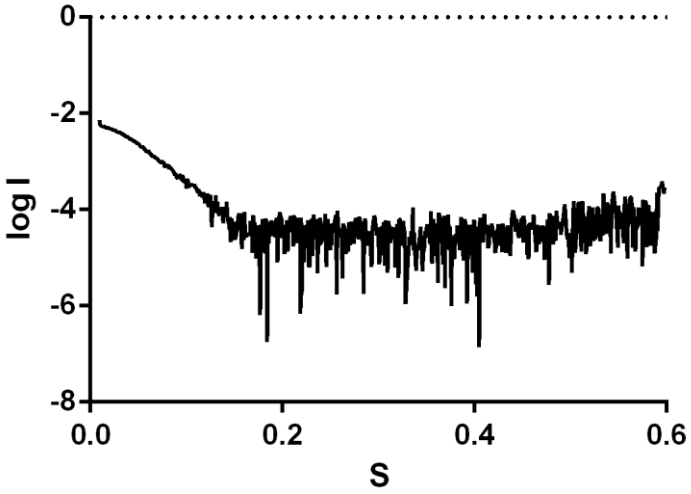
When *C. reinhardtii* activity is compared between nucleotides, it exhibits a much greater rate and substrate affinity whilst utilising NADPH (Figure 5.4). This exhibits a nucleotide specificity somewhat like *T. maritima* DHDPR, however with a much higher rate in the presence of NADH. This may indicate that it is more closely related to bacteria. However, NADPH may be the nucleotide preference of this algal enzyme due to the association with the light reactions in this chloroplast containing organism. This may indicate plant-like nucleotide utilisation tendencies.

The kinetic behaviour of the reduced form of the enzyme was tested in the presence of the reducing agent TCEP (Figure 5.5). When NADH was used as the nucleotide, both buffer conditions gave extremely similar profiles. Affinities were similar in both cases however the maximum rate was slightly higher in the presence of no reducing agent. However, when NADPH is utilised as the nucleotide, the reverse is observed. In the presence of a reducing agent, the observed rate is noticeably higher with a marked increase in substrate inhibition compared to the unreduced counterpart. Overall however, there is not a great deal of difference between the reduced and unreduced forms of the enzyme indicating that the reduced form of the enzyme does not exhibit noticeably different enzyme activity.

5.5 Small Angle X-ray Scattering

To further characterise the quaternary structure of the protein in the presence of reducing agent, small angle X-ray scattering was performed. This allowed further confirmation of the quaternary structure of the protein. To ensure that the protein was completely reduced without reducing agents negatively affecting the data, the protein was treated with reducing agent a few days prior to data collection. Following treatment, the reducing agent was removed using SEC 24 hours prior to data collection with the sample ran on an SDS PAGE gel to check for reformation of the disulfide bond. This allowed for both a clean sample as well as no contaminating buffer components.

(a)



(b)

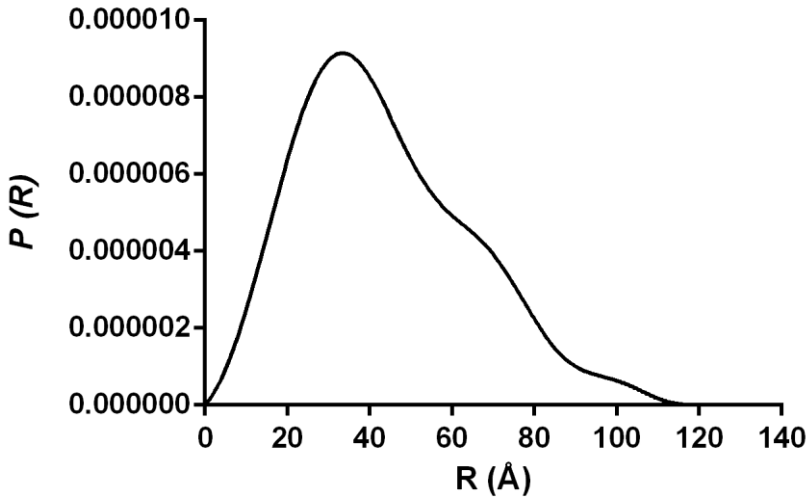


Figure 5.6: Intensity plot (log I vs S) and distance distribution plot (P (R)) for *C. reinhardtii* DHDPR.

Table 5.2: Parameters for SAXS analysis in reducing agent

Organism Name	**Rg (Å)**	**Io**	**Porod Volume (Å^3)**
A. thaliana	35.06±0.47	0.048	128,525
E. coli	39.88±3.59	0.046	241,859
C. reinhardtii	33.53±0.51	0.0097	120,825

The *C. reinhardtii P(R)* plot indicates that the mass of the object is not uniformly distributed around the centre of the protein (Figure 5.6). Normally, this would be indicative of a dimer in solution. However, it retains a strange curve shape in which

there are multiple peaks and valleys on the right side of the graph. This may indicate that the dimer has formed an interesting non-uniform structure.

Like *C. merolae* DHDPR, *C. reinhardtii* DHDPR X-ray scatter patterns can be compared between bacterial and plant type DHDPR enzymes (Table 5.2). The radius of gyration of the reduced form is slightly less than *A. thaliana* which equilibrates to a dimeric quaternary structure. This is backed up by the Porod volumes which show a slightly smaller volume than *A. thaliana* DHDPR which also calculates to half of the tetrameric volume.

Table 5.3: Chi^2 fit for *C. reinhardtii* scattering data against the crystal structures of *E. coli* and *T. maritima* DHDPR and various modifications.

Organism Name	**Comparison Structure**	**Chi^2**
C. reinhardtii DHDPR	*E. coli* DHDPR Tetramer	0.33
C. reinhardtii DHDPR	*E. coli* DHDPR β-sheet Dimer	0.26
C. reinhardtii DHDPR	*E. coli* DHDPR Loop Dimer	0.25
C. reinhardtii DHDPR	*E. coli* DHDPR Cross Dimer	0.41

To find out which dimer is formed, the scatter patterns were compared to the crystal structures for each theoretical dimer (Table 5.3). The best fits are for the β-sheet dimer and the Loop dimer, which have similar Chi^2 fits to the data. The cross dimer appears unlikely due to the improbable configuration of the dimer. The GASBOR model appears to resemble the β-sheet dimer due to the perceived monomeric arrangement. In any case, it appears that both the *C. merolae* and *C. reinhardtii* have a similar monomeric arrangement.

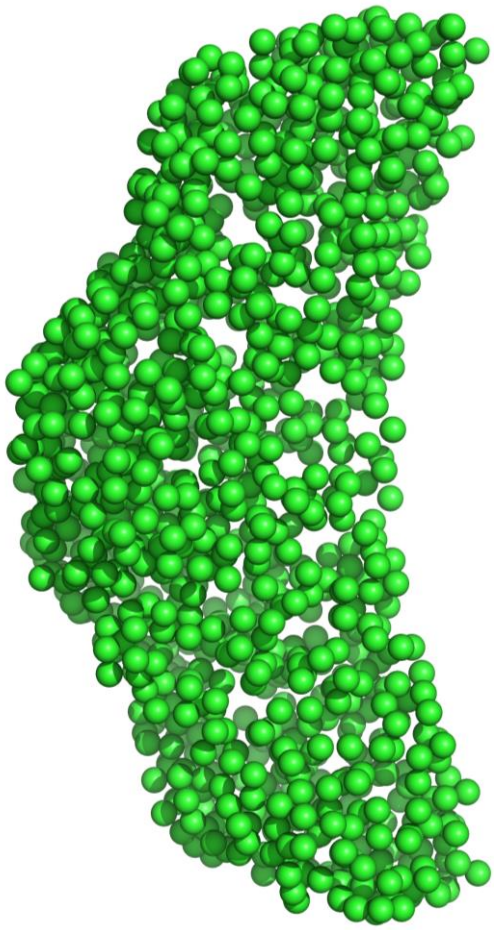


Figure 5.7: GASBOR model created from *C. reinhardtii* DHDPR (Chi^2 fit against raw data = 0.13)

GASBOR *ab initio* modelling suggests a very similar dimeric structure to *C. merolae* DHDPR (Figure 5.7). This indicates that the protein forms a dimer in reducing conditions.

5.6 Differential Scanning Fluorimetry

To test the thermal stability of *C. reinhardtii* DHDPR in comparison to other DHDPR enzymes, DSF was used. This also allowed comparison when bound to nucleotides as well as the substrate analogue, 2,6-pyridinedicarboxylate (PDC).

Table 5.4: DSF data for *E. coli* and *A. thaliana* DHDPR

Organism Name	**Ligand**	**Tm_{Ave} (°C)**
E. coli	No Ligand	80.28±0.00
E. coli	1.5 mM NADH	80.96±0.39
E. coli	1.5 mM NADPH	81.48±0.05
A. thaliana	No Ligand	43.21±0.24
A. thaliana	1.5 mM NADH	46.43±0.00
A. thaliana	1.5 mM NADPH	44.22±0.42

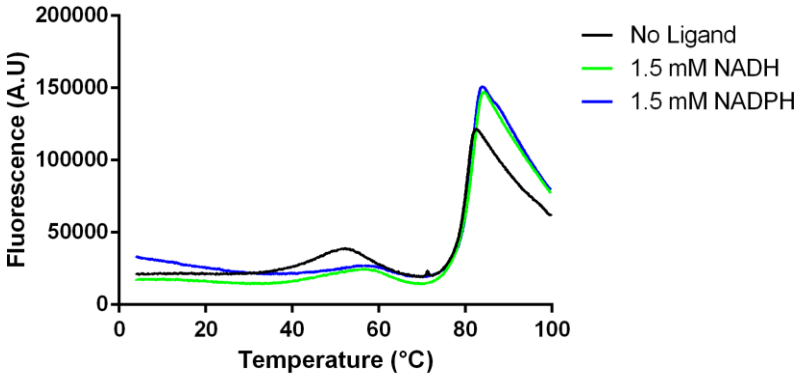


Figure 5.8: Melt curves for *E. coli* with no ligand and in the presence of nucleotide.

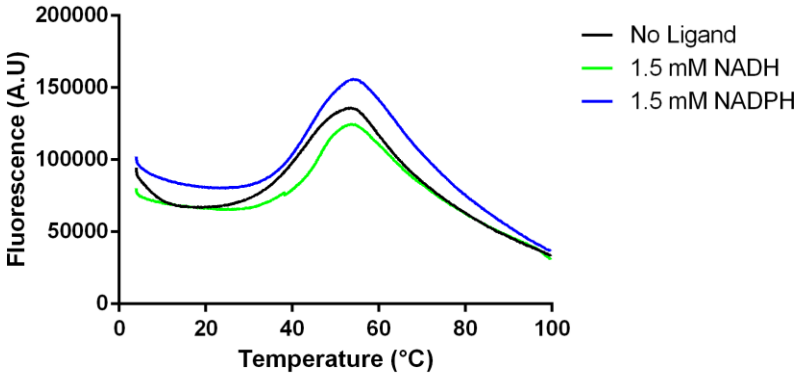


Figure 5.9: Melt curves for *A. thaliana* with no ligand and in the presence of nucleotide.

To compare with *C. reinhardtii* DHDPR, the melting points of *E. coli* and *A. thaliana* DHDPR were taken (Figures 5.8 and 5.9, Table 5.4). These were found to be 80.3°C and 43.2°C respectively with no ligand present. With the addition of nucleotide (the first ligand that binds to the enzyme), the thermal stability of the protein marginally increased by a small amount in both cases. This had been previously observed for *T. maritima* DHDPR (Pearce et al., 2008). What was interesting to note was the presence of a small secondary melt curve in *E. coli*. With no ligand present, a small peak at ~47°C was observed. The peak was also observed in the nucleotide bound species but to a much lesser extent.

Table 5.5: DSF data for *C. reinhardtii* DHDPR in the presence of no reducing agent

Ligand	**Tm_{Ave} (°C)**	**Secondary Tm_{Ave} (°C)**
No Ligand	45.9±0.3	72.5±0.0
1.5 mM NADH	47.1±0.2	76.2±0.2
1.5 mM NADPH	47.7±1.4	74.7±0.1
PDC	48.3±2.1	
PDC, 1.5 mM NADH	55.4±2.1	
PDC, 1.5 mM NADPH	46.1±0.1	66.7±0.2

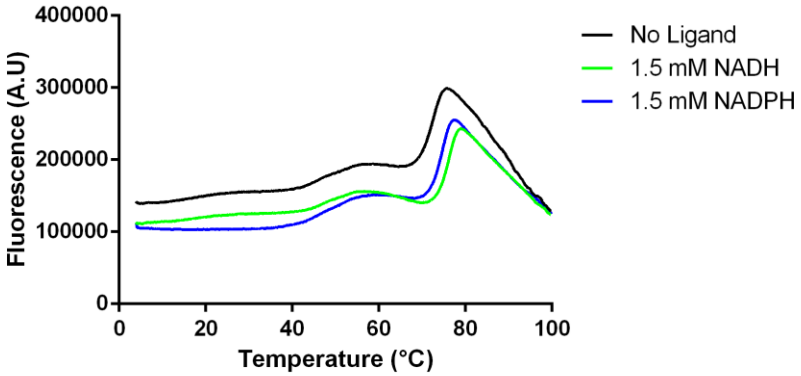


Figure 5.10: Melt curves for *C. reinhardtii* DHDPR in the presence of no ligand along with two different nucleotides.

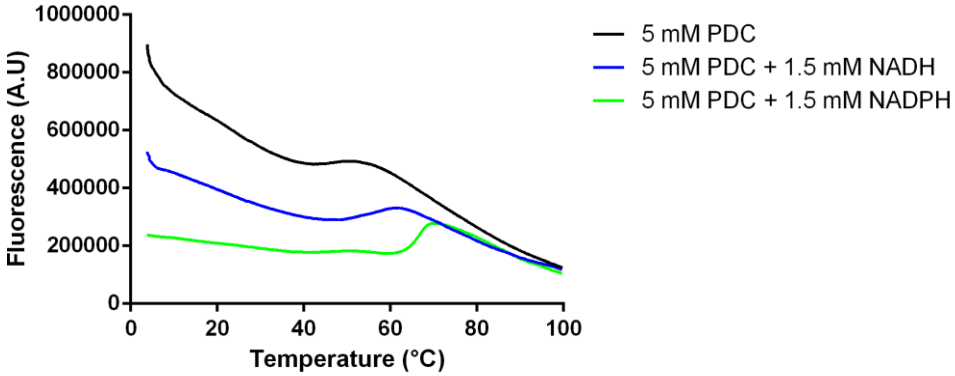


Figure 5.11: Melt curves for *C. reinhardtii* DHDPR in the presence of substrate analogue PDC in addition to two different nucleotides.

This data can be compared with *C. reinhardtii* DHDPR to understand its stability (Table 5.5). Like in *E. coli*, it contains two melt peaks (Figure 5.10). The smaller first peak was found to have a Tm of 45.9 °C. This was followed by a much larger second peak with a Tm of 72.5°C. This observation indicates that the enzyme undergoes a two-step melting process. However, addition of various ligands changes the melting profile. Addition of nucleotide increased the thermal stability of the initial peak by 1.2 °C and 1.8 °C degrees for NADH and NADPH respectively. A 3.8 °C and a 2.2 °C increase in stability was also observed in the second larger peak.

The addition of substrate analogue PDC greatly increases the stability of the enzyme (Figure 5.11). With PDC present, a 2.4 °C increase from the initial ligand unbound stability is observed. However, there is only one peak observed instead of the usual

two. This is also observed with addition of both NADH and PDC with a 9.5 °C increase from the ligand unbound value but no secondary peak. When NADPH is used as the substrate, the melt curve reverts to a typical two-step curve with a small peak at 46.1°C followed by a normal sized peak at 66.7°C. This second peak value exhibits a significantly increased melting temperature, indicating that the protein may have a higher affinity towards NADPH as its substrate. The two-step binding mechanism also indicates that the binding mechanisms may not be the same for each nucleotide.

Table 5.6: DSF data for *C. reinhardtii* DHDPR in the presence of 2 mM TCEP

Ligand	**Tm_{Ave} (°C)**	**Secondary Tm_{Ave} (°C)**
No Ligand	62.9±0.0	
1.5 mM NADH	48.1±0.9	70.6±0.7
1.5 mM NADPH	49.9±1.1	71.2±0.1
PDC	49.3±0.6	
PDC, 1.5 mM NADH	55.5±1.0	
PDC, 1.5 mM NADPH	46.5±1.2	65.7±0.3

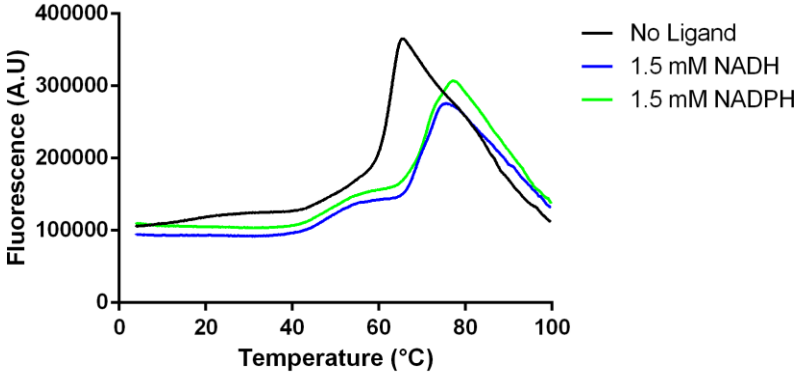


Figure 5.12: Melt curves for *C. reinhardtii* DHDPR in presence of reducing agent with no ligand along with two concentrations of nucleotide.

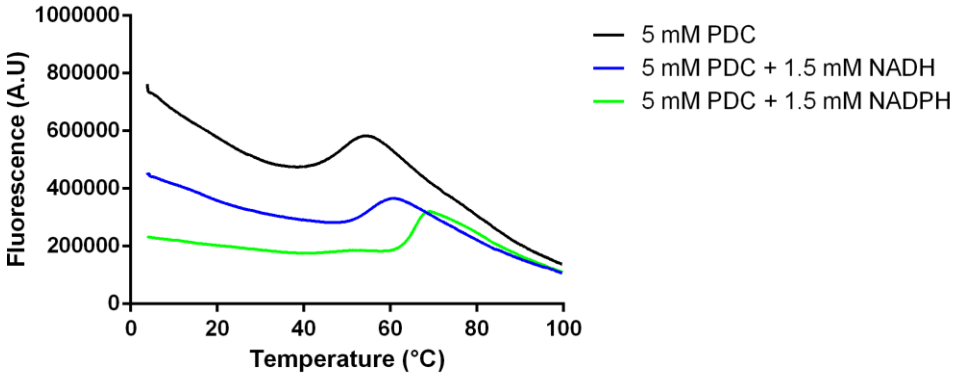


Figure 5.13: Melt curves for *C. reinhardtii* DHDPR in the presence of reducing agent with substrate analogue PDC in addition to two different nucleotides.

The addition of TCEP influences the melting profile of the protein (Table 5.6). With the presence of 2mM TCEP forming a reduced species, the first peak remains in approximately the same position, albeit with a smaller peak height that is not quantified in analysis (Figure 5.12). The second peak is decreased by 9.6°C. NADH and NADPH bound species also lost stability by 5.6°C and 8.0°C respectively. These values indicate that the reduced dimeric species has a lower stability than the native *C. reinhardtii* DHDPR species. In the presence of PDC alone and PDC in solution with NADH, a one-step melt curve is observed, similar to the unreduced species. Thermal stabilities under these conditions were within 1 °C of their unreduced counterparts. The two-step PDC and NADPH utilisation is also present in the reduced form of the enzyme with only ~1°C difference between the native enzyme (Figure 5.13). These values indicate that there is little difference in the binding mechanism between the reduced and unreduced forms of the enzyme.

5.7 Conclusions

5.7.1 Proposed DHDPR Evolutionary Lineage

With the discovery that *C. reinhardtii* DHDPR exists as both a dimer and a tetramer in solution, a possible evolutionary mechanism can be inferred (Figure 5.14). With tetrameric and dimeric DHDPR enzymes existing, it is likely that at some point in the evolutionary lineage, a divergence similar to DHDPS occurred from the bacterial to the plant form. If *C. reinhardtii* DHDPR exists in both the bacterial tetramer and plant dimer forms, it is possible that it is the divergence point between these species. This is corroborated by the fact that *C. reinhardtii* DHDPR exists on the evolutionary lineage exactly between bacterial and plant DHDPR species in one of the proposed evolutionary lineages. This makes *C. reinhardtii* DHDPR an important species in the evolution of the protein.

This discovery is interesting when compared to the proposed DHDPS evolution mechanism. The split between the dimeric and plant type tetramer in DHDPS was found to occur in the lycophytic *S. moellendorffii* DHDPS. It appears to occur at a different positon in DHDPR with SMO DHDPR thought to exist as a dimer (unpublished data). As they are both members of the same pathway, the fact that the individual members of the pathway appear to have different evolutionary paths is interesting.

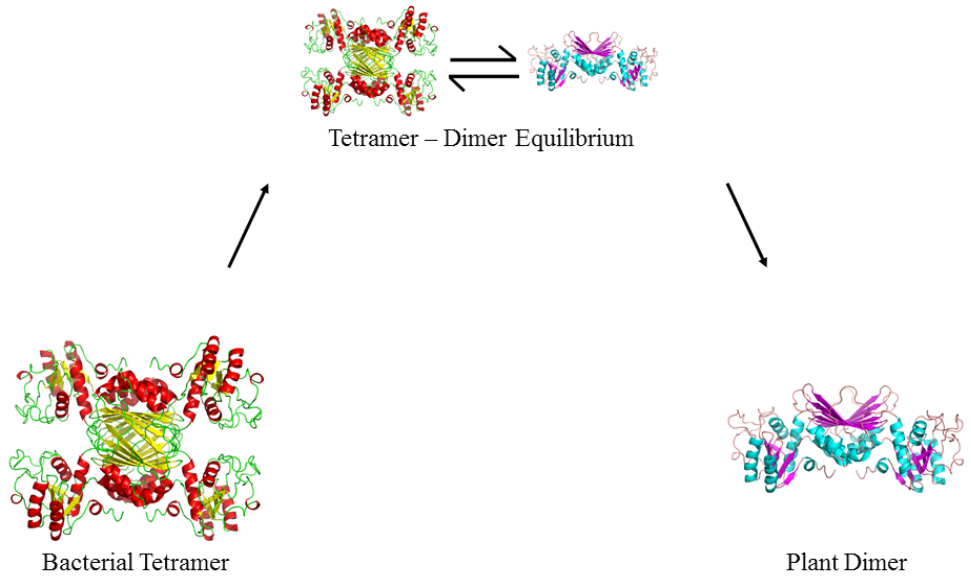


Figure 5.14: This image shows the proposed evolutionary pathway for DHDPR. On the left is an *E. coli* DHDPR which represents the tetrameric bacterial structure. On the right side is a proposed dimeric DHDPR representing the plant type structure. In the middle is the *C. reinhardtii* DHDPR equilibrium between tetrameric and dimeric forms of the enzyme.

5.7.2 Summary

AUC data showed that *C. reinhardtii* DHDPR exists in an equilibrium between dimer and tetramer. It was also shown that the protein contains a novel disulfide-dependent dimer interface with the reducing agent TCEP forcing an exclusively dimeric species in solution. Kinetic assays showed that the reduced form of the enzyme did not exhibit greatly altered kinetics, indicating that the reduced form of the enzyme is still viable. DSF data indicated that that both nucleotide and the substrate analogue PDC were found to increase the stability of the protein. PDC was also found to induce a one-step melt curve in the presence of NADH in both reduced and unreduced forms of the enzyme but NADPH was found to restore the typical two-step melt profile. This equilibrium along with its location in the evolutionary lineage could mean that *C. reinhardtii* DHDPR exists as the “divergence” point between bacterial and plant forms of the protein.

Chapter 6 – Conclusions

The main aim of this project was to characterise the evolutionary lineage of DHDPS and DHDPR. Firstly, the evolutionary lineage of DHDPS was investigated through the characterisation of *Selaginella moellendorffii* DHDPS, which exists in a substrate-mediated equilibrium between dimeric and tetrameric forms. The allosteric inhibitor lysine pushed the equilibrium towards the dimeric form, whereas the substrate pyruvate pushed the equilibrium towards the tetrameric form. This equilibrium may be the evolutionary divergence point between dimeric and the tetrameric plant quaternary structure. Along with the already discovered dimer-bacterial tetramer equilibrium, this allows the construction of a proposed DHDPS evolutionary pathway.

A similar evolutionary difference in DHDPR is also proposed. In this study, several red, brown and green algal DHDPR enzymes were found to exist as a dimeric species. This reinforced the idea that the DHDPR enzymes from these organisms existed after those of plants on the evolutionary lineage, proposing them as dimers. This included the discovery of *C. reinhardtii* DHDPR whose location on the evolutionary lineage along with its dimer-tetramer equilibrium indicates that it may be a divergence point between the bacterial and plant forms of the enzyme.

Chapter 7 – References

ATKINSON, S. C., DOGOVSKI, C., DOWNTON, M. T., CZABOTAR, P. E., DOBSON, R. C., GERRARD, J. A., WAGNER, J. & PERUGINI, M. A. 2013. Structural, kinetic and computational investigation of *Vitis vinifera* DHDPS reveals new insight into the mechanism of lysine-mediated allosteric inhibition. *Plant Mol Biol,* 81**,** 431-46.

ATKINSON, S. C., DOGOVSKI, C., NEWMAN, J., DOBSON, R. C. & PERUGINI, M. A. 2011. Cloning, expression, purification and crystallization of dihydrodipicolinate synthase from the grapevine *Vitis vinifera*. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 67**,** 1537-41.

BLICKLING, S., RENNER, C., LABER, B., POHLENZ, H. D., HOLAK, T. A. & HUBER, R. 1997. Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. *Biochemistry,* 36**,** 24-33.

BURGESS, B. R., DOBSON, R. C., BAILEY, M. F., ATKINSON, S. C., GRIFFIN, M. D., JAMESON, G. B., PARKER, M. W., GERRARD, J. A. & PERUGINI, M. A. 2008. Structure and evolution of a novel dimeric enzyme from a clinically important bacterial pathogen. *J Biol Chem,* 283**,** 27598-603.

CIRILLI, M., ZHENG, R., SCAPIN, G. & BLANCHARD, J. S. 2003. The three-dimensional structures of the *Mycobacterium tuberculosis* dihydrodipicolinate reductase-NADH-2,6-PDC and -NADPH-2,6-PDC complexes. Structural and mutagenic analysis of relaxed nucleotide specificity. *Biochemistry,* 42**,** 10644-50.

CONLY, C. J., SKOVPEN, Y. V., LI, S., PALMER, D. R. & SANDERS, D. A. 2014. Tyrosine 110 plays a critical role in regulating the allosteric inhibition of *Campylobacter jejuni* dihydrodipicolinate synthase by lysine. *Biochemistry,* 53**,** 7396-406.

COULTER, C. V., GERRARD, J. A., KRAUNSOE, J. A. E. & PRATT, A. J. 1999. *Escherichia coli* dihydrodipicolinate synthase and dihydrodipicolinate reductase: Kinetic and inhibition studies of two putative herbicide targets *Pesticide Science,* 55**,** 887-895.

DEVENISH, S. R., BLUNT, J. W. & GERRARD, J. A. 2010. NMR studies uncover alternate substrates for dihydrodipicolinate synthase and suggest that dihydrodipicolinate reductase is also a dehydratase. *J Med Chem,* 53**,** 4808-12.

DOBSON, R. C., DEVENISH, S. R., TURNER, L. A., CLIFFORD, V. R., PEARCE, F. G., JAMESON, G. B. & GERRARD, J. A. 2005a. Role of arginine 138 in the catalysis and regulation of *Escherichia coli* dihydrodipicolinate synthase. *Biochemistry,* 44**,** 13007-13.

DOBSON, R. C., GIRON, I. & HUDSON, A. O. 2011. L,L-diaminopimelate aminotransferase from *Chlamydomonas reinhardtii*: a target for algaecide development. *PLoS One,* 6**,** e20439.

DOBSON, R. C., GRIFFIN, M. D., DEVENISH, S. R., PEARCE, F. G., HUTTON, C. A., GERRARD, J. A., JAMESON, G. B. & PERUGINI, M. A. 2008. Conserved main-chain peptide distortions: a proposed role for Ile203 in catalysis by dihydrodipicolinate synthase. *Protein Sci,* 17**,** 2080-90.

DOBSON, R. C., GRIFFIN, M. D., JAMESON, G. B. & GERRARD, J. A. 2005b. The crystal structures of native and (S)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. *Acta Crystallogr D Biol Crystallogr,* 61**,** 1116-24.

DOBSON, R. C., PERUGINI, M. A., JAMESON, G. B. & GERRARD, J. A. 2009. Specificity versus catalytic potency: The role of threonine 44 in *Escherichia coli* dihydrodipicolinate synthase mediated catalysis. *Biochimie,* 91**,** 1036-44.

DOBSON, R. C., VALEGARD, K. & GERRARD, J. A. 2004. The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: further evidence for a catalytic triad. *J Mol Biol,* 338**,** 329-39.

DOMIGAN, L. J., SCALLY, S. W., FOGG, M. J., HUTTON, C. A., PERUGINI, M. A., DOBSON, R. C., MUSCROFT-TAYLOR, A. C., GERRARD, J. A. & DEVENISH, S. R. 2009. Characterisation of dihydrodipicolinate synthase (DHDPS) from *Bacillus anthracis*. *Biochim Biophys Acta,* 1794**,** 1510-6.

DOMMARAJU, S. R., DOGOVSKI, C., CZABOTAR, P. E., HOR, L., SMITH, B. J. & PERUGINI, M. A. 2011. Catalytic mechanism and cofactor preference of dihydrodipicolinate reductase from methicillin-resistant *Staphylococcus aureus*. *Arch Biochem Biophys,* 512**,** 167-74.

EGGELING, L. & BOTT, M. 2015. A giant market and a powerful metabolism: L-lysine provided by *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol,* 99**,** 3387-94.

FALCO, S. C., GUIDA, T., LOCKE, M., MAUVAIS, J., SANDERS, C., WARD, R. T. & WEBBER, P. 1995. Transgenic canola and soybean seeds with increased lysine. *Biotechnology (N Y),* 13**,** 577-82.

FUCHS, T. M., SCHNEIDER, B., KRUMBACH, K., EGGELING, L. & GROSS, R. 2000. Characterization of a bordetella pertussis diaminopimelate (DAP) biosynthesis locus identifies dapC, a novel gene coding for an N-succinyl-L,L-DAP aminotransferase. *J Bacteriol,* 182**,** 3626-31.

GALILI, G. & AMIR, R. 2013. Fortifying plants with the essential amino acids lysine and methionine to improve nutritional quality. *Plant Biotechnol J,* 11**,** 211-22.

GALILI, G., AMIR, R., HOEFGEN, R. & HESSE, H. 2005. Improving the levels of essential amino acids and sulfur metabolites in plants. *Biol Chem,* 386**,** 817-31.

GE, X., OLSON, A., CAI, S. & SEM, D. S. 2008. Binding synergy and cooperativity in dihydrodipicolinate reductase: implications for mechanism and the design of biligand inhibitors. *Biochemistry,* 47**,** 9966-80.

GENG, F., CHEN, Z., ZHENG, P., SUN, J. & ZENG, A.-P. 2013. Exploring the allosteric mechanism of dihydrodipicolinate synthase by reverse engineering of the allosteric inhibitor binding sites and its application for lysine production. *Applied Microbiology and Biotechnology*, 97**,** 1963-1971.

GHISLAIN, M., FRANKARD, V. & JACOBS, M. 1990. Dihydrodipicolinate synthase of *Nicotiana sylvestris*, a chloroplast-localized enzyme of the lysine pathway. *Planta,* 180**,** 480-6.

GIRISH, T. S., NAVRATNA, V. & GOPAL, B. 2011. Structure and nucleotide specificity of *Staphylococcus aureus* dihydrodipicolinate reductase (DapB). *FEBS Lett,* 585**,** 2561-7.

GRAHAM, D. E. & HUSE, H. K. 2008. Methanogens with pseudomurein use diaminopimelate aminotransferase in lysine biosynthesis. *FEBS Lett,* 582**,** 1369-74.

GRIFFIN, M. D., BILLAKANTI, J. M., WASON, A., KELLER, S., MERTENS, H. D., ATKINSON, S. C., DOBSON, R. C., PERUGINI, M. A., GERRARD, J. A. & PEARCE, F. G. 2012. Characterisation of the first enzymes committed to lysine biosynthesis in *Arabidopsis thaliana*. *PLoS One,* 7**,** e40318.

GRIFFIN, M. D., DOBSON, R. C., PEARCE, F. G., ANTONIO, L., WHITTEN, A. E., LIEW, C. K., MACKAY, J. P., TREWHELLA, J., JAMESON, G. B., PERUGINI, M. A. & GERRARD, J. A. 2008. Evolution of quaternary structure in a homotetrameric enzyme. *J Mol Biol,* 380**,** 691-703.

GUNJI, Y. & YASUEDA, H. 2006. Enhancement of l-lysine production in methylotroph *Methylophilus methylotrophus* by introducing a mutant LysE exporter. *Journal of Biotechnology,* 127**,** 1-13.

HERMANN, M., THEVENET, N. J., COUDERT-MARATIER, M. M. & VANDECASTEELE, J. P. 1972. Consequences of lysine oversynthesis in Pseudomonas mutants insensitive to feedback inhibition. Lysine excretion or endogenous induction of a lysine-catabolic pathway. *Eur J Biochem,* 30**,** 100-6.

HUDSON, A. O., SINGH, B. K., LEUSTEK, T. & GILVARG, C. 2006. An LL-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants. *Plant Physiol,* 140**,** 292-301.

HUTTON, C. A., PERUGINI, M. A. & GERRARD, J. A. 2007. Inhibition of lysine biosynthesis: an evolving antibiotic strategy. *Mol Biosyst,* 3**,** 458-65.

JANOWSKI, R., KEFALA, G. & WEISS, M. S. 2010. The structure of dihydrodipicolinate reductase (DapB) from *Mycobacterium tuberculosis* in three crystal forms. *Acta Crystallogr D Biol Crystallogr,* 66**,** 61-72.

KAUR, N., GAUTAM, A., KUMAR, S., SINGH, A., SINGH, N., SHARMA, S., SHARMA, R., TEWARI, R. & SINGH, T. P. 2011. Biochemical studies and crystal structure determination of dihydrodipicolinate synthase from *Pseudomonas aeruginosa. Int J Biol Macromol,* 48**,** 779-87.

KEFALA, G., EVANS, G. L., GRIFFIN, M. D., DEVENISH, S. R., PEARCE, F. G., PERUGINI, M. A., GERRARD, J. A., WEISS, M. S. & DOBSON, R. C. 2008. Crystal structure and kinetic study of dihydrodipicolinate synthase from *Mycobacterium tuberculosis*. *Biochem J,* 411**,** 351-60.

KOBASHI, N., NISHIYAMA, M. & TANOKURA, M. 1999. Aspartate kinase-independent lysine synthesis in an extremely thermophilic bacterium, Thermus thermophilus: lysine is synthesized via alpha-aminoadipic acid not via diaminopimelic acid. *J Bacteriol,* 181**,** 1713-8.

KONAREV, P. V., VOLKOV, V. V., SOKOLOVA, A. V., KOCH, M. H. J. & SVERGUN, D. I. 2003. PRIMUS: a Windows PC-based system for small-

angle scattering data analysis. *Journal of Applied Crystallography,* 36**,** 1277-1282.

KOSUGE, T. & HOSHINO, T. 1998. Lysine is synthesized through the alpha-aminoadipate pathway in *Thermus thermophilus*. *FEMS Microbiol Lett,* 169**,** 361-7.

LIU, Y., WHITE, R. H. & WHITMAN, W. B. 2010. Methanococci use the diaminopimelate aminotransferase (DapL) pathway for lysine biosynthesis. *J Bacteriol,* 192**,** 3304-10.

MANUALS, M. 2016. *Nutritional Requirements of Poultry* [Online]. Merck Manuals. Available: http://www.merckvetmanual.com/mvm/poultry/nutrition_and_management_poultry/nutritional_requirements_of_poultry.html [Accessed 8/09/16 2016].

MCCOY, A. J., ADAMS, N. E., HUDSON, A. O., GILVARG, C., LEUSTEK, T. & MAURELLI, A. T. 2006. L,L-diaminopimelate aminotransferase, a trans-kingdom enzyme shared by Chlamydia and plants for synthesis of diaminopimelate/lysine. *Proc Natl Acad Sci U S A,* 103**,** 17909-14.

MIKE TOKACH, J. D., STEVE DRITZ, BOB GOODBAND, JIM NELSSEN. 2013. *Amino Acid Requirements of Growing Pigs* [Online]. The Pig Site. Available: http://www.thepigsite.com/articles/4195/amino-acid-requirements-of-growing-pigs/ [Accessed].

MILLWARD, D. J. & JACKSON, A. A. 2004. Protein/energy ratios of current diets in developed and developing countries compared with a safe protein/energy ratio: implications for recommended protein and amino acid intakes. *Public Health Nutr,* 7**,** 387-405.

MISONO, H., TOGAWA, H., YAMAMOTO, T. & SODA, K. 1979. Meso-alpha,epsilon-diaminopimelate D-dehydrogenase: distribution and the reaction product. *J Bacteriol,* 137**,** 22-7.

MITSAKOS, V., DOBSON, R. C., PEARCE, F. G., DEVENISH, S. R., EVANS, G. L., BURGESS, B. R., PERUGINI, M. A., GERRARD, J. A. & HUTTON, C. A. 2008. Inhibiting dihydrodipicolinate synthase across species: towards specificity for pathogens? *Bioorg Med Chem Lett,* 18**,** 842-4.

MUSCROFT-TAYLOR, A. C., CATCHPOLE, R. J., DOBSON, R. C., PEARCE, F. G., PERUGINI, M. A. & GERRARD, J. A. 2010. Disruption of quaternary structure in *Escherichia coli* dihydrodipicolinate synthase (DHDPS) generates

a functional monomer that is no longer inhibited by lysine. *Arch Biochem Biophys,* 503**,** 202-6.

NAKAYAMA, K., ARAKI, K. & KASE, H. 1978. Microbial Production of Essential Amino Acids with *Corynebacterium Glutamicum* Mutants. *In:* FRIEDMAN, M. (ed.) *Nutritional Improvement of Food and Feed Proteins.* Boston, MA: Springer US.

PEARCE, F. G., SPRISSLER, C. & GERRARD, J. A. 2008. Characterization of dihydrodipicolinate reductase from Thermotoga maritima reveals evolution of substrate binding kinetics. *J Biochem,* 143**,** 617-23.

PETOUKHOV, M. V., FRANKE, D., SHKUMATOV, A. V., TRIA, G., KIKHNEY, A. G., GAJDA, M., GORBA, C., MERTENS, H. D., KONAREV, P. V. & SVERGUN, D. I. 2012. New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr,* 45**,** 342-350.

REBOUL, C. F., POREBSKI, B. T., GRIFFIN, M. D., DOBSON, R. C., PERUGINI, M. A., GERRARD, J. A. & BUCKLE, A. M. 2012. Structural and dynamic requirements for optimal activity of the essential bacterial enzyme dihydrodipicolinate synthase. *PLoS Comput Biol,* 8**,** e1002537.

REDDY, S. G., SACCHETTINI, J. C. & BLANCHARD, J. S. 1995. Expression, purification, and characterization of *Escherichia coli* dihydrodipicolinate reductase. *Biochemistry,* 34**,** 3492-501.

REDDY, S. G., SCAPIN, G. & BLANCHARD, J. S. 1996. Interaction of pyridine nucleotide substrates with Escherichia coli dihydrodipicolinate reductase: thermodynamic and structural analysis of binary complexes. *Biochemistry,* 35**,** 13294-302.

RESEARCH, T. M. 2015. *Global Amino Acids Market to Expand at 9.10% CAGR till 2018, Propelled by Rise in Meat Consumption* [Online]. Available: http://www.transparencymarketresearch.com/pressrelease/lysine-amino-acids.htm [Accessed 7/09/2016 2016].

RICE, E. A., BANNON, G. A., GLENN, K. C., JEONG, S. S., STURMAN, E. J. & RYDEL, T. J. 2008. Characterization and crystal structure of lysine insensitive *Corynebacterium glutamicum* dihydrodipicolinate synthase (cDHDPS) protein. *Arch Biochem Biophys,* 480**,** 111-21.

RICHARDSON, C. R. & HATFIELD, E. E. 1978. The limiting amino acids in growing cattle. *J Anim Sci,* 46**,** 740-5.

RODIONOV, D. A., VITRESCHAK, A. G., MIRONOV, A. A. & GELFAND, M. S. 2003. Regulation of lysine biosynthesis and transport genes in bacteria: yet another RNA riboswitch? *Nucleic Acids Res,* 31**,** 6748-57.

SCAPIN, G., BLANCHARD, J. S. & SACCHETTINI, J. C. 1995. Three-dimensional structure of *Escherichia coli* dihydrodipicolinate reductase. *Biochemistry,* 34**,** 3502-12.

SCAPIN, G., REDDY, S. G., ZHENG, R. & BLANCHARD, J. S. 1997. Three-dimensional structure of *Escherichia coli* dihydrodipicolinate reductase in complex with NADH and the inhibitor 2,6-pyridinedicarboxylate. *Biochemistry,* 36**,** 15081-8.

SCHRUMPF, B., SCHWARZER, A., KALINOWSKI, J., PUHLER, A., EGGELING, L. & SAHM, H. 1991. A functionally split pathway for lysine synthesis in *Corynebacterium glutamicium. J Bacteriol,* 173**,** 4510-6.

SINGH, S. P., BORA, T. C. & BEZBARUAH, R. L. 2012. Molecular Interaction of Novel Compound 2-Methylheptyl Isonicotinate Produced by *Streptomyces sp.* 201 with Dihydrodipicolinate Synthase (DHDPS) Enzyme of *Mycobacterium tuberculosis* for its Antibacterial Activity. *Indian J Microbiol,* 52**,** 427-32.

SKOVPEN, Y. V., CONLY, C. J., SANDERS, D. A. & PALMER, D. R. 2016. Biomimetic Design Results in a Potent Allosteric Inhibitor of Dihydrodipicolinate Synthase from *Campylobacter jejuni*. *J Am Chem Soc,* 138**,** 2014-20.

SOARES DA COSTA, T. P., DESBOIS, S., DOGOVSKI, C., GORMAN, M. A., KETAREN, N. E., PAXMAN, J. J., SIDDIQUI, T., ZAMMIT, L. M., ABBOTT, B. M., ROBINS-BROWNE, R. M., PARKER, M. W., JAMESON, G. B., HALL, N. E., PANJIKAR, S. & PERUGINI, M. A. 2016. Structural Determinants Defining the Allosteric Inhibition of an Essential Antibiotic Target. *Structure,* 24**,** 1282-91.

SOARES DA COSTA, T. P., MUSCROFT-TAYLOR, A. C., DOBSON, R. C., DEVENISH, S. R., JAMESON, G. B. & GERRARD, J. A. 2010. How essential is the 'essential' active-site lysine in dihydrodipicolinate synthase? *Biochimie,* 92**,** 837-45.

SUNDHARADAS, G. & GILVARG, C. 1967. Biosynthesis of alpha,epsilon-diaminopimelic acid in *Bacillus megaterium*. *J Biol Chem,* 242**,** 3983-4.

SVERGUN, D. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography,* 25**,** 495-503.

SVERGUN, D., BARBERATO, C. & KOCH, M. H. J. 1995. CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography,* 28**,** 768-773.

SVERGUN, D. I., PETOUKHOV, M. V. & KOCH, M. H. 2001. Determination of domain structure of proteins from X-ray solution scattering. *Biophys J,* 80**,** 2946-53.

TAMIR, H. & GILVARG, C. 1974. Dihydrodipicolinic acid reductase. *J Biol Chem,* 249**,** 3034-40.

THIERBACH, G., KALINOWSKI, J., BACHMANN, B. & PÜHLER, A. 1990. Cloning of a DNA fragment from *Corynebacterium glutamicum* conferring aminoethyl cysteine resistance and feedback resistance to aspartokinase. *Applied Microbiology and Biotechnology,* 32**,** 443-448.

TOME, D. & BOS, C. 2007. Lysine requirement through the human life cycle. *J Nutr,* 137**,** 1642S-1645S.

TSUJIMOTO, N., GUNJI, Y., OGAWA-MIYATA, Y., SHIMAOKA, M. & YASUEDA, H. 2006. L-Lysine biosynthetic pathway of *Methylophilus methylotrophus* and construction of an L-lysine producer. *J Biotechnol,* 124**,** 327-37.

USHA, V., LLOYD, A. J., ROPER, D. I., DOWSON, C. G., KOZLOV, G., GEHRING, K., CHAUHAN, S., IMAM, H. T., BLINDAUER, C. A. & BESRA, G. S. 2016. Reconstruction of diaminopimelic acid biosynthesis allows characterisation of Mycobacterium tuberculosis N-succinyl-L,L-diaminopimelic acid desuccinylase. *Sci Rep,* 6**,** 23191.

VAZ, F. M. & WANDERS, R. J. 2002. Carnitine biosynthesis in mammals. *Biochem J,* 361**,** 417-29.

VOSS, J. E., SCALLY, S. W., TAYLOR, N. L., ATKINSON, S. C., GRIFFIN, M. D., HUTTON, C. A., PARKER, M. W., ALDERTON, M. R., GERRARD, J. A., DOBSON, R. C., DOGOVSKI, C. & PERUGINI, M. A. 2010. Substrate-mediated stabilization of a tetrameric drug target reveals Achilles heel in anthrax. *J Biol Chem,* 285**,** 5188-95.

WATANABE, N., CLAY, M. D., VAN BELKUM, M. J., CHERNEY, M. M., VEDERAS, J. C. & JAMES, M. N. G. 2008. Mechanism of Substrate Recognition and PLP-induced Conformational Changes in LL-Diaminopimelate Aminotransferase from Arabidopsis thaliana. *Journal of Molecular Biology,* 384**,** 1314-1329.

WATKIN, S. A. J. 2014. *Structure and Function of Dihydrodipicolinate Reductase.* Bachelor of Science with Honours, University of Canterbury.

XU, J., HAN, M., REN, X. & ZHANG, W. 2016. Modification of aspartokinase III and dihydrodipicolinate synthetase increases the production of l-lysine in Escherichia coli. *Biochemical Engineering Journal,* 114**,** 79-86.

ZABRISKIE, T. M. & JACKSON, M. D. 2000. Lysine biosynthesis and metabolism in fungi. *Nat Prod Rep,* 17**,** 85-97.

ZHU, C., NAQVI, S., GOMEZ-GALERA, S., PELACHO, A. M., CAPELL, T. & CHRISTOU, P. 2007. Transgenic strategies for the nutritional enhancement of plants. *Trends Plant Sci,* 12**,** 548-55.