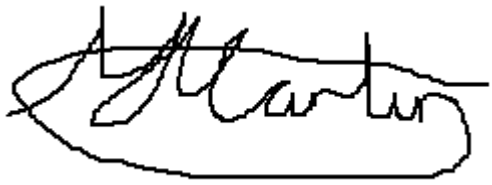
MECHANISTIC STUDIES OF *ESCHERICHIA COLI* TRANSKETOLASE

Submitted by

Mathew Martin

to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Chemistry (December 2008).

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.



Mathew Martin

Abstract

The enzyme transketolase is found in nature as part of the Pentose Phosphate Pathway to rearrange large sugar phosphates. It also is an important enzyme for carbon-carbon bond formation for industrial biocatalysis.

The work presented in this thesis describes the purification, crystallisation, characterisation and structural determination of the recombinant *Escherichia coli* transketolase complexed with the substrate hydroxypyruvate and potential inhibitor fluoropyruvate. The native transketolase and the transketolase-hydroxypyruvate structures were solved to a 1.18 and 1.05 Å resolution respectively. The transketolase structures show a chain of ordered water molecules spanning a distance of 20 Å between the two active sites. The water molecules are linked via a network of hydrogen bonds and they are proposed to facilitate proton transfer between the two-thiamine pyrophosphate molecules, thereby providing a method of communication between the two active sites of the enzyme. The transketolase-hydroxypyruvate structure shows the hydroxypyruvate substrate forming a covalent bond to the thiamine pyrophosphate thereby creating a α,β-dihydroxyethyl–thiamine pyrophosphate complex within the enzyme active site. The novel transketolase-fluoropyruvate structure solved to a 1.60 Å resolution, it produced a snapshot image of the ketol donor prior to formation of the active enamine intermediate. The trapped fluoropyruvate molecule is shown to form an angle that varies from the accepted Burgi-Dunitz angle of 109.5° for nucleophilic attack. However, this is inconclusive due to the low occupancy of the fluoropyruvate. In addition, kinetic studies were performed on the recombinant *E. coli* transketolase to investigate the inhibitory role of fluoropyruvate during the enzymatic reaction.

The active site recombinant *E. coli* transketolase mutants H26Y and D469Y have been also been purified and characterised. The mutant H26Y complexed with fluoropyruvate was crystallised and its structure determined to 1.66 Å resolution. This structure has given an insight into why this mutation results in the formation of the opposite D-enantiomer of erythrulose rather than the L-erythrulose produced by the wild-type transketolase enzyme.

The thesis also includes the purification, crystallisation, characterisation and X-ray diffraction studies of the commercially useful oxygenating enzyme, 2,5-diketocamphane 1,2-monooxygenase from *Pseudomonas putida*. The recombinant dimeric oxygenase component of this enzyme has been crystallised and its structure solved to 1.4 Å resolution.

Acknowledgements

I would like to thank several people for their help and support throughout my PhD years. Firstly, I would like to thank Jenny Littlechild for taking me on, and for all her support throughout my research. I am forever grateful to Misha Isupov who has had the painful task of teaching me crystallography. On the subject of crystallography, I would also like to thank Tsutomu Nakamura for the year he spent in the UK, and Chris Sayer. I would like to thank Kirsty Line for helping me through my purification work, constantly there answering my stupid questions. I would like to thank Paul Dalby and John Ward from UCL; John for supplying me with exceedingly good clones, and Paul for helping me with the kinetic inhibition work.

A big thanks to the people of the Biocatalysis Centre, for all your entertainment at lunch time especially Joe.

Finally, I would like to thank my family for their support; so a big thanks to my Mum, Dad and my little sister Luisa, and of course my girlfriend Claire.

Abbreviations

2,5-DKMO- 2,5-diketocamphane 1,2-monooxygenase

3,6-DKMO- 3,6-diketocamphane 1,2-monooxygenase

A_{280}- Absorbance at 280 nm

Å- Angstrom

Au- Atomic units

APS- Ammonium persulfate

ATP- Adenosine Triphosphate

BAM- Benzamidine

bp- Base pairs

BSA- Bovine serum albumin

BVMO- Baeyer-Villiger monooxygenase

CCP4- Collaborative Computational Project, number 4

CHMO- cyclohexanone monooxygenase

D469Y- Glutamate mutation to tyrosine at position 469

DMSO- Dimethyl Sulfoxide

EC- Enzyme commission

EDTA- Ethylenediaminetetraacetic acid (disodium salt)

FAD- flavin adenine dinucleotide

FMN- flavin mononucleotide

FFQ- Fast flow Q

FPA- Fluoropyruvic acid

FPLC- Fast protein liquid chromatography

g- Acceleration due to gravity

GA- Glycolaldehyde

GAPDH- Glyceraldehyde 3-phosphate dehydrogenase

GF- Gel filtration

H26Y- Histidine mutation to tyrosine at position 26

HEPES- 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

his-tag- Poly histidine tag

HPA- Hydroxypyruvic acid

IPTG- Isopropyl β-D-galactopyranoside

kDa- Kilo Dalton

K_m- Michaelis Constant

LB- Luria-Bertani

MAD- Multiple anomalous diffraction

MES- 2-(*N*-morpholino)ethanesulfonic acid

MW- Molecular weight

MWCO – Molecular weight cut off

NADPH/ NAD^+- Nicotinamide adenine dinucleotide (protonated/deprotonated)

NMR- Nuclear magnetic resonance

OD- Optical density

PAGE- Polyacrylamide gel electrophoresis

PDB- Protein Databank

PEG- Polyethylene glycol

PIPES- 1,4-Piperazinediethanesulfonic acid

PMSF- Phenylmethylsulphonyl fluoride

PPP- Pentose Phosphate Pathway

RMS- Root mean square

rpm- Revolutions per minute

SAD- Single anomalous diffraction

SDS- Sodium dodecyl sulfate

SRS- Synchrotron radiation source

TEMED- N,N,N,N-tetramethylethylene diamide

TFA- Trifluoroacetic acid

TK- Transketolase

TPP- Thiamine pyrophosphate

Tris- Tris(hydroxymethyl)aminomethane

UV- Ultra violet

v/v- Volume to volume

V- Volts

V_{max}- Measure of maximum turnover by an enzyme

w/v- Weight to volume

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction into enzymes

Even though mankind (perhaps unknowingly) has used enzymes for thousands of years, enzymes only began to be fully understood in the last century. Their role within biological systems and living organisms has since been analysed and reviewed by many authors. There have been many scientific advances using enzymes in all walks of life, from the more obvious i.e. food production and medicinal/pharmaceutical viewpoint to the topical bio-fuels industry. Enzymes in their simplest form are biological catalysts, and they perform the same role as the more regular chemical catalysts by lowering the activation energy of a reaction. Scientists have long been admirers of enzymes, which is mainly due to the ability of an enzyme to perform a reaction at great speed in a highly specific manner at ambient temperature and pH. This trait can be observed in the metabolic pathway of a biological cell, where an enzyme's high specificity and rate of reaction offers itself to be selective for a specific substrate. This enables a specific reaction to be carried out at a significant rate, and in turn passing the product of the reaction onto another enzyme, thus forming a pathway.

1.1.1 Brief history of enzymes

Enzymes act as nature's catalysts performing nature's reactions in a concise manner. Enzymes were first used in the fermentation process where the use of yeast extracts in the brewing of hops and other crops dates back before scientific records began. Other microbial enzymes have been used throughout the ages in a number of cases: the leavening of bread, the curdling of cheese and production of cloth. However, no genuine research was associated to these occurrences. The beginnings of research into the mode in which enzymes function started in the 1830's, when Payen and Perzon discovered the first enzyme through observing the formation of a precipitate from aqueous malt extract. Later that decade, three scientists, Cagniard-Latour, Schwann and Kutzing, published their work. Their work helped to understand the role of yeast in

fermentation. Schwann's discovery of the enzyme pepsin alongside Pasteur's study into the role of yeast in fermentation lead to the conclusion that there was a vital force within living cells. In 1836, Swedish chemist Berzelius published findings that led to the word 'catalysis', 40 years on German physiologist Kuhne used the Greek terminology 'enzyme' to describe this process. In 1897, German biochemist Buchner studied the uses of 'non-living' yeast extract in the fermentation activity, were he was later awarded the Nobel Prize in 1907 for his endeavours. This started the quest of isolating enzymes from living cells, which has opened up the field to enzyme purification, crystallisation, and eventually structure determination. This brief history was compiled using the following reviews: (Roseman, 2001), (Krishna, 2002) and (Schoffers *et al.*, 1996).

1.1.2 Enzyme nomenclature and classification

An enzyme's name was often derived from the substrate or reaction it catalysed with the addition of the suffix '–ase' to the end of the substrate or reactions name. Prior to 1956 enzymes were named in a subjective fashion, dependant on no structured system. As more enzymes were identified there became a need for a structured system for nomenclature, and in 1956 the International Union of Biochemistry and Molecular Biology was formed for the naming and classification of enzymes. The E.C. numbering system describes each enzyme by a series of four numbers with the first number relating to the catalytic reaction they conduct. The initial E.C. numbering system is subdivided into six groups: oxido-reductases (E.C.1), transferases (E.C.2), hydrolases (E.C.3), lyases (E.C.4), isomerases (E.C.5) and ligases (E.C.6). Following this subdivision, further divisions of the six classes occur depending on the cofactors and substrates used and the functional groups involved during catalysis.

1.1.3 Enzyme classes

Oxido-reductases (E.C.1) are involved in the transfer of electrons from one molecule to another; oxido-reductase enzymes catalyse the oxidation/reduction reactions.

Transferases (E.C.2) are involved in the transfer of a functional group from one molecule to another.

Hydrolases (E.C.3) are involved in the breaking of a bond using water. Hydrolase enzymes catalyse the hydrolysis of a chemical bond.

Lyases (E.C.4) are involved in the breaking of a chemical bond by an alternative route that doesn't involve using hydrolysis or oxidation.

Isomerases (E.C.5) are involved in catalysing the structural rearrangement of isomers.

Ligases (E.C.6) are involved in the joining of two large molecules with the formation of a new chemical bond.

1.1.4 Enzyme structure

There are four degrees of enzyme structure; primary, secondary, tertiary and quaternary. Enzyme structures consist of 20 different amino acid residues, all with 'L' stereochemistry differing in size and properties. A primary structure consists of a sequence of amino acid residues linked by planar peptide bonds. The secondary structure relates to the α-helices and β-sheets within the overall fold of the enzyme which are stabilised by hydrogen bonds (Anfinsen, 1973). The tertiary structure is the three-dimensional arrangement of the secondary structure elements, which is stabilised via weak interactions, such as hydrogen bonds, van der Waals interactions, hydrophobic interactions, and ionic bonds. Some structural proteins and globular proteins that are exported from the cell are stabilised by formation of covalent disulfide bonds. The quaternary structure is related to the association of two or more protein subunits. It is for example, common for individual polypeptides to interact together forming dimer, trimer, tetramers and higher aggregates. A protein structure consists of approximately 50% water. This consists of internal ordered water molecules often involved with enzyme mechanism and an ordered external shell of waters interacting with the positively charged amino acid side chains. In some cases enzymes contain cofactors that are involved in the catalytic process. An enzyme's structure can be unfolded by denaturing the protein, which entails the breaking of the interactions within the fold by heating, change in pH or some chemical method.

1.1.5 Enzyme selectivity

The interest in enzymes within the industrial world has risen significantly over the past 50 years. This is also because of our 'recent' understanding of enzyme function and our ability to technically utilise enzymes due to the enzymes ability to carry out a reaction at ambient temperatures, within a neutral pH condition, whilst producing fewer waste products. These key factors all relate to the enzymes selectivity, as enzymes can be very selective for the substrates they select for catalysis and in turn the product formed. Enzymes have high levels of stereospecificity, regioselectivity and chemoselectivity endearing them to the chemical industry (Roberts, 1998). The stereospecific nature of an enzyme allows it to act specifically upon one stereoisomer of a molecule. The enzymes regioselective character permits the enzyme to differentiate between identical functional groups in different position on the same molecule. The chemoselective quality of an enzyme can be observed by its ability to act upon a specific functional group in a molecule distinguishing it from similar entities.

1.1.6 Enzyme mechanism

In any chemical reaction, there is an activation energy to overcome before the reaction will take place. Any chemical or biological catalyst acts by lowering the energy barrier, allowing the reaction to proceed quicker in comparison to an un-catalysed reaction. There are a variety of mechanisms in which enzymes operate to reduce the activation energy.

One such way is for the enzyme to bind the substrate into its transition state form. With the catalytic residues of the enzyme stabilising the substrate in its strained transition state, the energy for the reaction to occur is reduced as the energy required for the transition stage is removed. Another method in which an enzyme can reduce the energy of the transition state, is to generate an environment that is of opposite charge to that of the transition state. By doing this the enzyme need not require to alter the substrate's conformation, as it stabilises the intermediate through the offset of the substrate's charge. Enzymes may also reduce the activation energy by providing an alternative pathway in which the reaction can take place. A way that an enzyme may achieve this is through binding directly with the substrate forming an enzyme-substrate

intermediate, which in turn reacts again forming the product. In some cases this may not be achievable without the presence of a coenzyme. As a general rule, enzymes create an environment that stabilises the transition state during catalysis in a way that is not achievable in a regular solvent based reaction.

1.1.7 Enzyme cofactors

When enzymes catalyse a reaction they may form a transient covalent bond to change the reaction pathway. Some enzymes require a non-protein component within their active site to achieve full activity. These non-protein components are called cofactors; they range from small inorganic molecules such as metals to larger organic compounds. The organic compounds are seen in two forms, a prosthetic group molecule which is tightly bound covalently in the active site playing a key role in the catalysis reaction, remaining bound throughout the whole process. The other type of organic cofactor is the coenzymes, which unlike the prosthetic group are not covalently bound. Some cofactors are released from the active site during catalysis. The coenzymes act by transferring chemical moieties between enzymes.

1.1.8 Enzyme kinetics

Enzyme kinetics concerns itself with the rate at which an enzyme catalyses a reaction. Study of enzyme kinetics can reveal the catalytic mechanism of the enzyme, showing the mechanistic steps during a reaction pathway. Enzyme kinetics is routinely determined by following the concentration of the substrate or product as the reaction proceeds. From the data obtained binding and inhibition constants may be calculated for the substrate and enzyme. A variety of techniques maybe used in the measurement of kinetic data. Spectroscopic methods are commonly used observing the change in absorbance from the product and substrate. Chromatography methods have also been used to measure the amount of product formed or substrate used. Other methods used include radioactive labelling, which can monitor the formation of product over time.

1.1.8.1 Single-substrate biotransformations

Single substrate reactions generally relate to one substrate entering the active site, creating some form of enzyme-substrate intermediate, whilst being transformed into a single product, before being released out of the enzyme (Fig. 1.1)

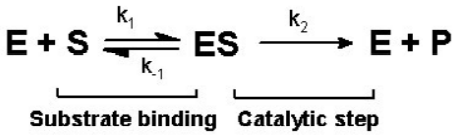


Figure 1.1 The single step mechanism for a single substrate reaction.

All enzyme-catalysed reactions are saturable, thus on increasing concentrations of substrate the enzyme will reach a level of maximum rate, known as the V_{max}. The Michaelis Menten kinetic model of a single substrate reaction is derived from the understanding of the equilibrium concentrations of substrate and enzyme throughout the reaction. This led to the Michaelis-Menten equation (Michaelis and Menten, 1913) that describes how the reaction rate (v) depends on the position of the substrate-binding equillibrium and the rate constant (k_2). It was observed that equation (Eqn. 1.1) was achieved when k_2 is far less than k_1.

$$v = V_{max}[S] / K_m + [S]$$

Equation 1.1 The Michaelis - Menten equation for single substrate reaction.

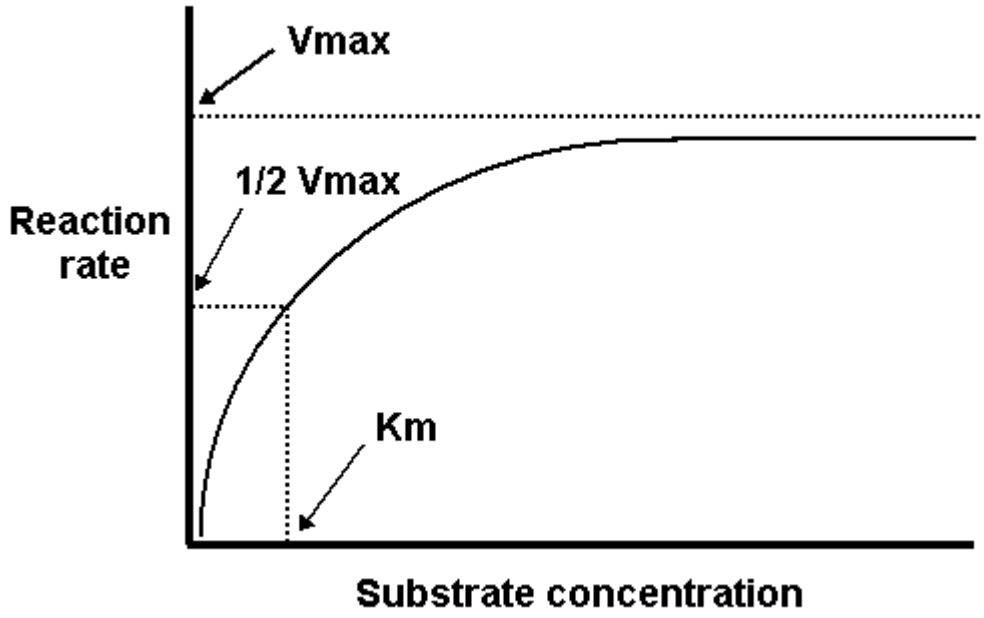


Figure 1.2 The saturation curve for a single substrate reaction.

The Michaelis-Menten constant K_m is defined, as the concentration of substrate when the enzyme reaction is half V_{max} (Fig. 1.2). Lineweaver-Burk plots are also used to illustrate kinetic data; this is formed by taking the reciprocal of both sides of the Michaelis -Menten equation (Lineweaver and Burk, 1934). The resulting calculations create a straight-line graph in the form y=mx+c, with a y-intercept of $1/V_{max}$ and an x-intercept of $-1/K_m$ (Eqn. 1.2) (Fig. 1.3).

$$1/\boldsymbol{v} = (\boldsymbol{K_m} / \boldsymbol{V_{max}} [S]) + (1 / \boldsymbol{V_{max}})$$

Equation 1.2 The Lineweaver-Burk double reciprocal of the Michaelis Menten equation for single substrate reaction.

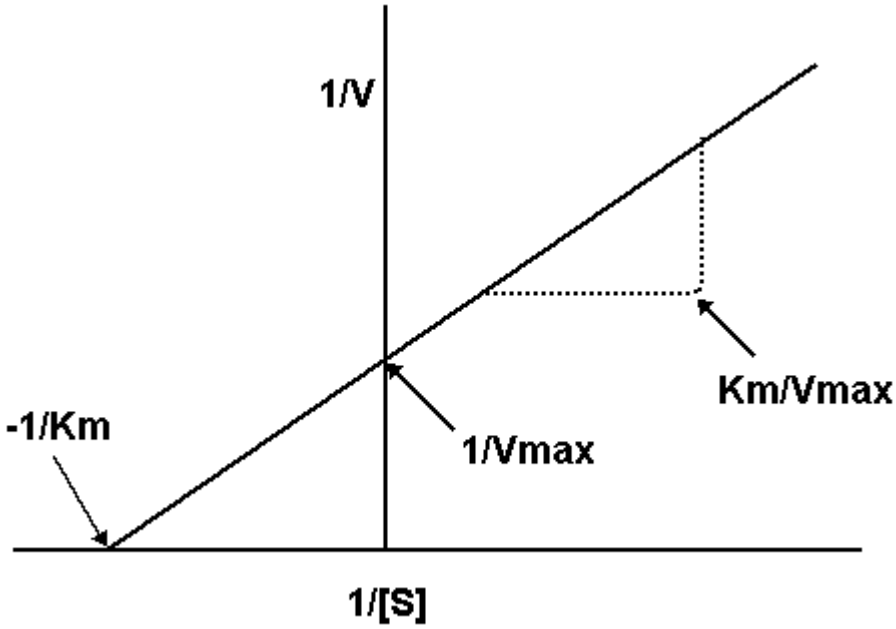


Figure 1.3 The Lineweaver-Burk plot of the Michaelis Menten equation for a single substrate reaction.

1.1.8.2 Multi-substrate reactions

Multi-substrate reactions differ from single substrate reactions due to the complexity of their rate equations. Multi-substrate reactions may be followed in the same manner as a single-substrate reaction, if one of the substrates is kept constant whilst the other is varied. In this case the V_{max} and K_m can be calculated in the same manner as before. For enzymes that accept two substrates and form two products there are two mechanisms, the ternary complex mechanism and the ping-pong mechanism. The ternary complex mechanism is used when both substrates bind to the enzyme producing a ternary complex. The order of binding of the substrates can either be of random nature, or by an expected sequence (Fig. 1.4). The ternary complex mechanism was used to describe the reaction catalysed by dihydrofolate reductase from (Morrison and Stone, 1988)).

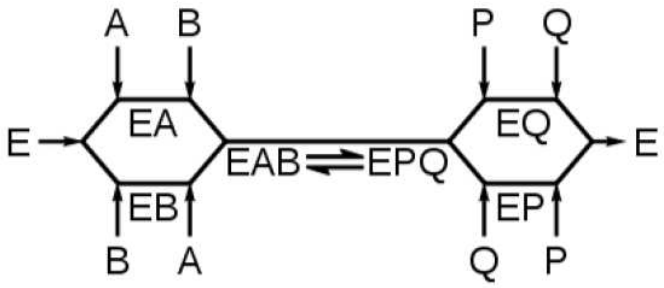


Figure 1.4 The random order of ternary complex mechanism.

The ping-pong mechanism refers to enzymes that can exist in two states, the ground state and a chemically modified state. In these cases, the first substrate binds to the enzyme, chemically modifying it for example, by transferring a chemical group to the enzyme, and is then released. Once the initial substrate is released, the second substrate can bind and react with the modified enzyme, reforming the ground state enzyme and final product (Fig. 1.5). The rate equation for the ping-pong mechanism was discussed by William Wallace Cleland (Cleland, 1973). In a Lineweaver-Burk plot of data for a ping-pong mechanism, parallel lines are produced when the initial substrate is kept contant and the second substrate is varied (and vice versa).

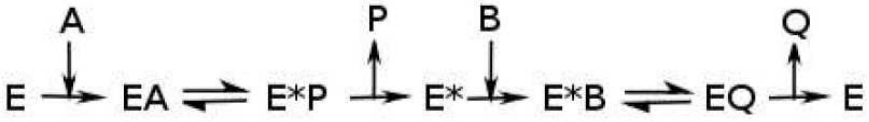


Figure 1.5 The ping-pong mechanism.

1.1.9 Enzyme inhibition

Enzyme inhibitors are small molecules that reduce or completely remove the enzyme's activity. There are two types of inhibition, reversible and irreversible. Reversible inhibitors can be referred to as competitive, non-competitive, or a mixture of the two depending on the effect they have on the reaction's V_{max} and K_m. These reversible inhibitors usually work by binding to the enzyme in place of the substrate. The actual type of reversible inhibition can be observed by analysis of the kinetic data

for the reaction. Irreversible inhibition operates in a separate manner to reversible inhibition. The most common form of irreversible inhibition is the modification of the active site of the enzyme, usually by altering an active site residue. Suicide substrates may be referred to as a subset of irreversible inhibition, as they form an irreversible complex with the enzyme through a covalent bond during the catalytic reaction.

1.2 Introduction to biocatalysis and biotransformation

As discussed, the ability of a substance to speed up a transformation, without itself being altered, is known as a catalyst. The most well known catalysts are chemical catalysts, and have been used for years in important industrial processes. An obvious example is iron, which is used in the Haber process to produce ammonia, and catalysts such as platinum and rhodium that are used in car catalytic converters.

Biological catalysts have been performing biotransformations for centuries. An example was the use of papaya juice, which was used in the tendering of leather utilising the protease papain that removed animal hairs from the hides (Muench, 1934). Another example can be seen in the use of rennet from the stomach of calves for cheese making. Little of the theory was understood, but humans persisted in using the extract of rennet, renin, a protease that broke down the proteins in milk, which prevented its coagulation (McMahon and Brown, 1982). These are examples of original biocatalysts used by man, but it was not until chemists came up against a series of problems that they looked back at nature, and today the modern biocatalyst is present throughout our lives.

For years straightforward wet organic chemistry has been the main route of synthesis for organic compounds. However, in the past half-century chemists have encountered a range of issues in attaining their final compound. These problems include: the separation of racemic mixtures, the detrimental effect upon the environment caused during the industrial process, and economic problems due to the increased price of substrate. In a racemic mixture, both enantiomers of the compound are present within a defined ratio of each other. The racemates could display similar chemical and physical properties, but have totally different biological properties. An example of this comes from the pharmaceutical industry, in the case of the drug Thalidomide. The drug was used to treat morning sickness in pregnant women during the late 50’s and early 60’s.

Thalidomide (α-phthalimidoglutarimide) has a chiral centre and was used as a racemate (1:1 mixture) of dextrorotatory (R)- and levorotatory (S)-thalidomide. Only the R-enantiomer had the desired calming and sleep-inducing properties whereas the S-enantiomer caused the teratogenic effects (Brent, 2004).

Chemists have been successful in introducing techniques to tackle the problem of racemic mixtures. With the development of a number of organic chiral synthesis reactions, yielding a high enantiomer excess, and using chemical resolution to separate the racemates. These methods have been proven to be an expensive and a non eco-friendly way to achieve an optically pure compound. These problems have forced chemists to look elsewhere.

The introduction of enzymes to catalyse reactions in organic synthesis (biotransformation / biocatalysis reactions) contributed to the resolution of both environmental and economic problems in the search of an optically pure compound (Roberts, 1998). Enzymes are highly efficient catalysts with the rates of mediated processes being typically accelerated by a factor 10^8-10^{10} compared to the corresponding non-enzymatic reactions (Storm and Koshland, 1970). These values are far greater than the values that chemical catalysts are capable of achieving.

1.2.1 Industrial biocatalysis

The industrial biocatalysis was born from two patterns of thought. Firstly, the curiosity of scientists with their endeavour to understand the process by which an enzyme exhibits its catalytic nature. Research is undertaken into the enzyme's function and mechanism, which in turn leads to the full classification. The second driving force is the financial reward through the enzyme's commercial usage. The ability for the enzyme to replace an expensive chemical procedure in a more efficient semi-natural manner is endearing to both the companies bank balance and PR division. Naturally, for newly discovered pharmaceuticals and other fine chemicals, a synthetic route to these new products designed around a key biocatalytic step must nowadays deserve serious consideration (Schmid *et al.*, 2001).

The introduction of a biocatalyst into the industrial world was first exhibited by the use of whole cell microorganisms; an example of this was in the production of acetic acid from alcohol. The first use of an isolated enzyme in industrial biocatalysis was

observed in 1969, where the acylase method was used in the resolution of a racemic mixture of amino acids (Reviewed by (Sheldon *et al.*, 2007)). Then in the 1970's, hydrolase enzymes were used in the cleavage of amide and ester bonds (Jones *et al.*, 1976) Over the next decade more and more biocatalysts were beginning to play a role in industrial organic synthesis (Fig. 1.6). We now live in a world where many different biotransformations are carried out in industry, mostly for the production of pharmaceutical and agrochemical precursors. In most cases, chiral compounds (fine chemicals) are obtained. Biotransformations are also successful for the production of commodities such as acrylamide (Wandrey *et al.*, 2000). The increase in importance of biocatalysis with the chemical industry has been reflected in its economic importance to the world (Manfred, 2006). The production of acrylamide is one of the largest applications of biocatalysis in the industrial world, annually producing several thousands tons of product (Thomas *et al.*, 2002). Another example is the production of (*S*)-methoxyisopropylamine which involves the use of a lipase biotransformation during the 1000 tons per year production (Schmid *et al.*, 2001).

An example of where a chemical company has moved from chemical to biocatalysis can be seen in the production of the β-lactam antibiotic cephalexin (van de Sandt and De Vroom, 2000). The Dutch company DSM made the switchover in production after it was clear that the academic and industrial research investment would make a return in profits. Other companies have incorporated enzymes into the production of their chemical compounds, BASF (Ludwigshafen) has exploited the substrate specificity of the *Pseudomonas* lipase in order to gain enantiomerically pure (R) and (S) amines. This process is being performed on a 100 ton a year scale (Schulze and Wubbolts, 1999).

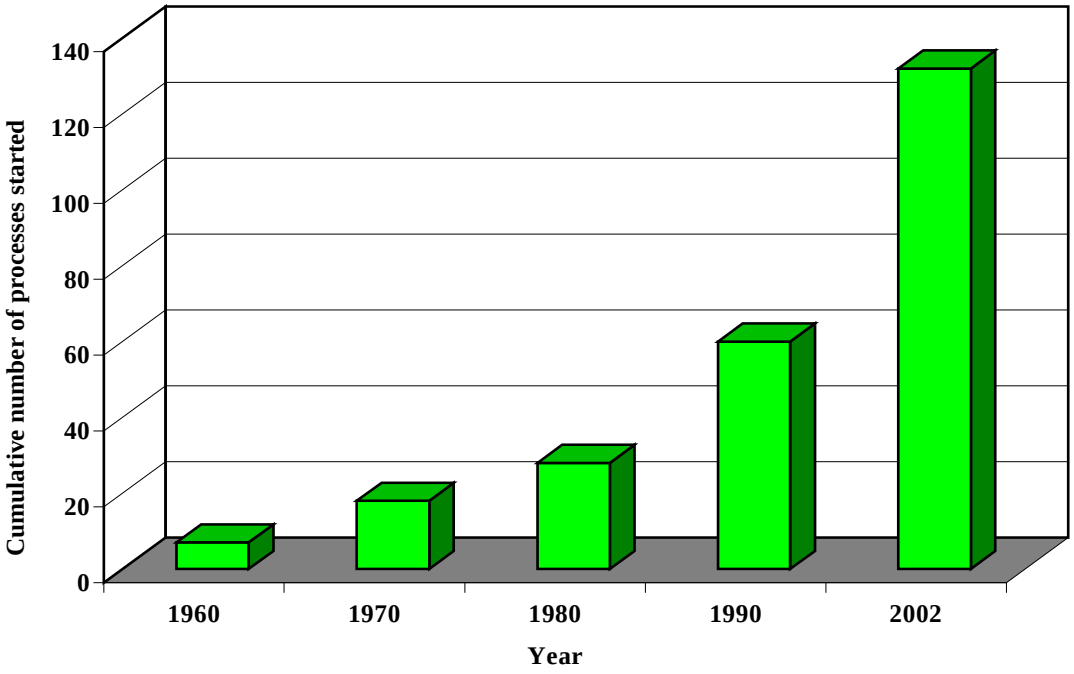


Figure 1.6 Bar chart showing the cumulative number of industrial processes using biotransformations from 1960 to 2002 (Straathof *et al.*, 2002).

The uses of enzymes as biocatalysts now stretch far and wide and are involved in many areas of modern day commerce (Fig. 1.7). The majority of their use still remains in the production of pharmaceuticals; an example is the anti-leukaemia agent being produced by GSK (UK) by a purine acylation biotransformation. The reaction is almost impossible to achieve by normal chemical acylation due to the preference for N-acylation, but the lipase used in the biotransformation produces a relatively high percentage of regioselectivity in creation of the desired produced (Rasor and Voss, 2001). The food industry continues to utilise biotransformations in the production of dairy products and general food manufacture. This is a growing area as the world faces a future of possible food shortages. The use of biotransformations in agro-industries encompass the production of chemicals and the waste management aspect.

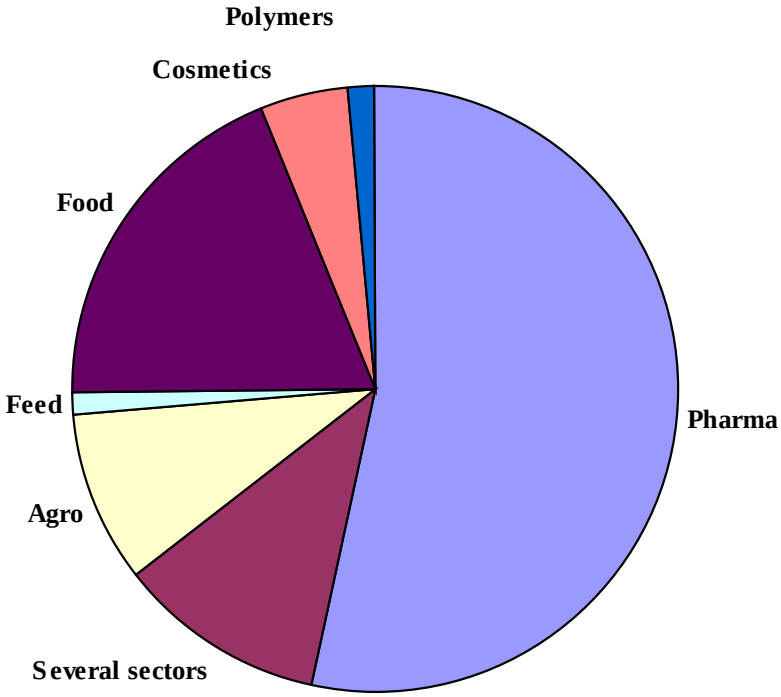


Figure 1.7 Pie chart showing the proportion of different industrial processes using biotransformation (Straathof *et al.*, 2002).

Although there has been a huge growth in the use of enzymes as industrial catalysts, there still remains many prejudices regarding biocatalysis within industry (Rasor and Voss, 2001). These prejudices have slowed the integration of biocatalysts into the industrial world, but the increased research and development are beginning to remove such patterns of thought.

Previously, enzymes were widely believed to act only on their natural substrates. Although it is true that enzymes exhibit higher activity towards the natural substrates, they can still conduct related transformations on a broader range of non-natural substrates to a lesser degree.

Another issue that has caused apprehension regarding the use of enzymes is the belief that enzymes only work in aqueous conditions. For the majority of enzymes this is true, but the enzymes that are commonly used in organic synthesis such as lipases can work just as effectively in a range of organic solvents. This may come as no surprise since the environment where lipases act in nature is on the interface between lipids and

water (Rasor and Voss, 2001). There have been many documented cases in industry of the increased use of enzymes in non-aqueous conditions (Morgan *et al.*, 1997).

Two further difficulties relate to the low substrates loading concentrations, and the sensitive operating environment of enzymes themselves. Once more both these topics are key, but there have been many industrial advances that show these factors need not stunt the growth of biocatalysis within the industrial sector. An example of this is GSK's (UK) production of abacavir, where one stage is catalysed by an alkaline protease using substrate concentrations up to 100 g/l. The sensitivity of an enzyme is a very subjective matter, regarding the process that the industry requires, however, this is also the case with a chemical catalyst. Enzymes found in organisms that live under extreme conditions have been used in biotransformations to take advantage of their geochemical properties. Enzymes from thermophiles found in the hot springs of Japan, such as the alcohol dehydrogenase from *Aeropyrum pernix* have been used for the synthesis of chiral alcohols (Guy *et al.*, 2003).

The initial cost of developing a biocatalytic process in comparison to the chemical alternative can be high, as the infrastructure for the production of the chemical catalysts is already well established. The recent increase in biotechnology and advances in the over-expression of enzymes has brought biocatalysis up alongside the chemical catalyst in economic terms. The questions asked from an industrial perspective centre upon, 'can the product afford the price of the catalyst?' and 'is there a cheaper alternative?'(Rasor and Voss, 2001).

1.3 Formation of C-C bonds

C-C bond formation is fundamental to all of organic chemistry, and is used in a variety of processes. In the production of a rather chemically complex molecule, a series of reactions involving small molecules must occur. These reactions have a broad range, but it is likely that some form of C-C bond formation reaction will have to take place to increase the carbon backbone of the molecule. There have been many documented examples of catalysed C-C bond formation, both chemically and biochemically.

1.3.1 Chemical C-C bond formation

In conventional organic chemistry many routes towards the formation of the C-C bond have been devised. The aldol condensation is a synthesis where the nucleophilic addition of a ketone enolate to an aldehyde forms an aldol in the presence of a base (Fig. 1.8a). Another name for this form of synthesis is the Claisen condensation. One of the most commonly used synthesis in industry is the Diels-Alder reaction that involves the addition of an alkene to a 1,4 diene forming a 6-membered carbon ring (Diels and Alder, 1928) (Fig. 1.8b). Another way of creating a C-C bond is in the use of a Grignard reagent, which acts as a nucleophile by attacking a carbonyl group carbon. The reaction often yields a tertiary alcohol (Grignard, 1900) (Fig. 1.8c). The Wittig reaction is an alternative carbon bond forming synthesis focussed around the reactivity of the triphenyl phosphonium yilde. An alkene is produced from the reaction between an aldehyde or a ketone with the triphenyl phosphonium yilde (Wittig and Schöllkopf, 1954) (Fig. 1.8d). A more recent development in C-C bond chemical synthesis was the use of a palladium catalyst in the Heck reaction (Heck and Nolley, 1972) (Fig. 1.8e). This reaction often involves an aryl halide and an alkene in the presence of strong base with an organopalladium catalyst. This reaction is stereospecific producing a trans product, making the Heck reaction a widely available tool to industry in the production of fine chemicals (de Vries, 2001).

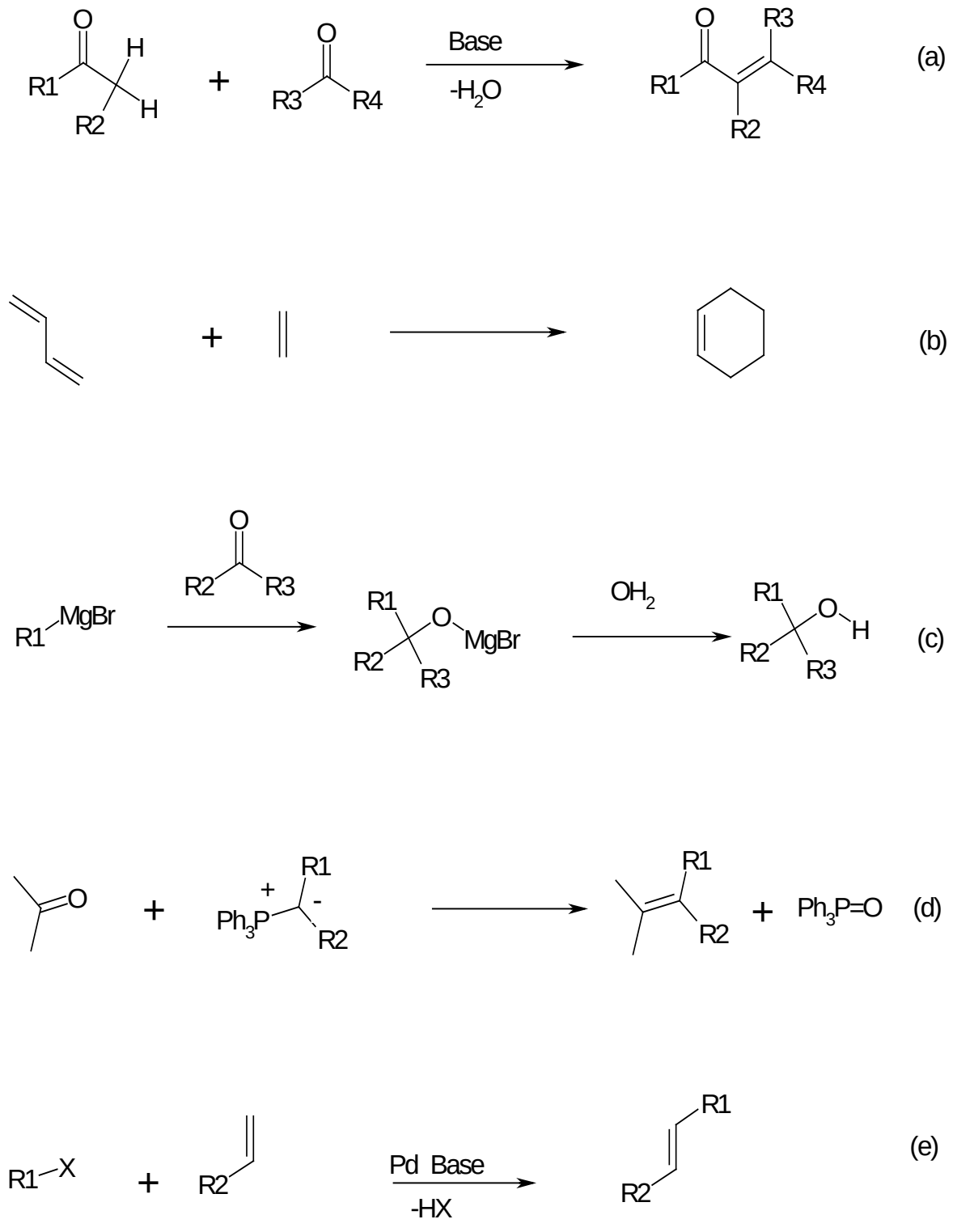


Figure 1.8 The chemical C-C bond formation reactions. (a) Aldol condensation, (b) Diels-Alder reaction, (c) Grignard reaction, (d) Wittig reaction, and (e) Heck reaction.

1.3.2 Biological C-C bond formation

The use of a biocatalyst in the formation of C-C bonds is an attractive proposition, as a biocatalyst can offer a larger degree of stereoselectivity compared to a regular chemical catalyst without the added inconvenience of the use of protecting groups. There are 3 classes of enzymes involved industrially in the formation of C-C bonds. Several enzymes have been documented to give rise to C-C bonds, including pyruvate-dependant and hydroxynitrile lyases (Fessner, 1998). With the latter being used in microreactor technology (Koch *et al.*, 2008). The hydroxynitrile lyases act by inserting a cyanide across the carbonyl of an aldehyde (Fig. 1.9a). Aldolase is a ligase enzyme that possesses the ability to form C-C bonds. The aldolase enzyme plays an important role in gluconeogenesis, where two 3-carbon compounds are joined to form the 6-carbon fructose molecule, fructose 1,6-*bis*phosphate (Fig. 1.9d). The two transferase enzymes transaldolase and transketolase act by transferring 3-carbon and 2-carbon subunits respectively in the form of a ketone to an aldehyde. Both these transferases are biologically present in the pentose phosphate pathway. Transketolase and transaldolase catalyse an aldol reaction as seen in section 1.3.1, whereby an aldehyde is added to another carbonyl extending the carbon chain length. The transaldolases are a group of enzymes that have the ability to extend the carbon chain length by 3-carbon units, this makes them very adept in carrying out a series of reactions which would normally prove rather difficult by traditional chemical means (Fig. 1.9b). The transketolases are a group of enzymes that have the ability to extend the carbon chain length by 2-carbon units (Fig. 1.9c). Compared to transaldolases, transketolase enzymes operate without the requirement of a phosphorylated substrate. In addition, transketolase have an increased stereospecificity towards non–native substrates, making it a better industrial catalyst in terms of racemic resolution.

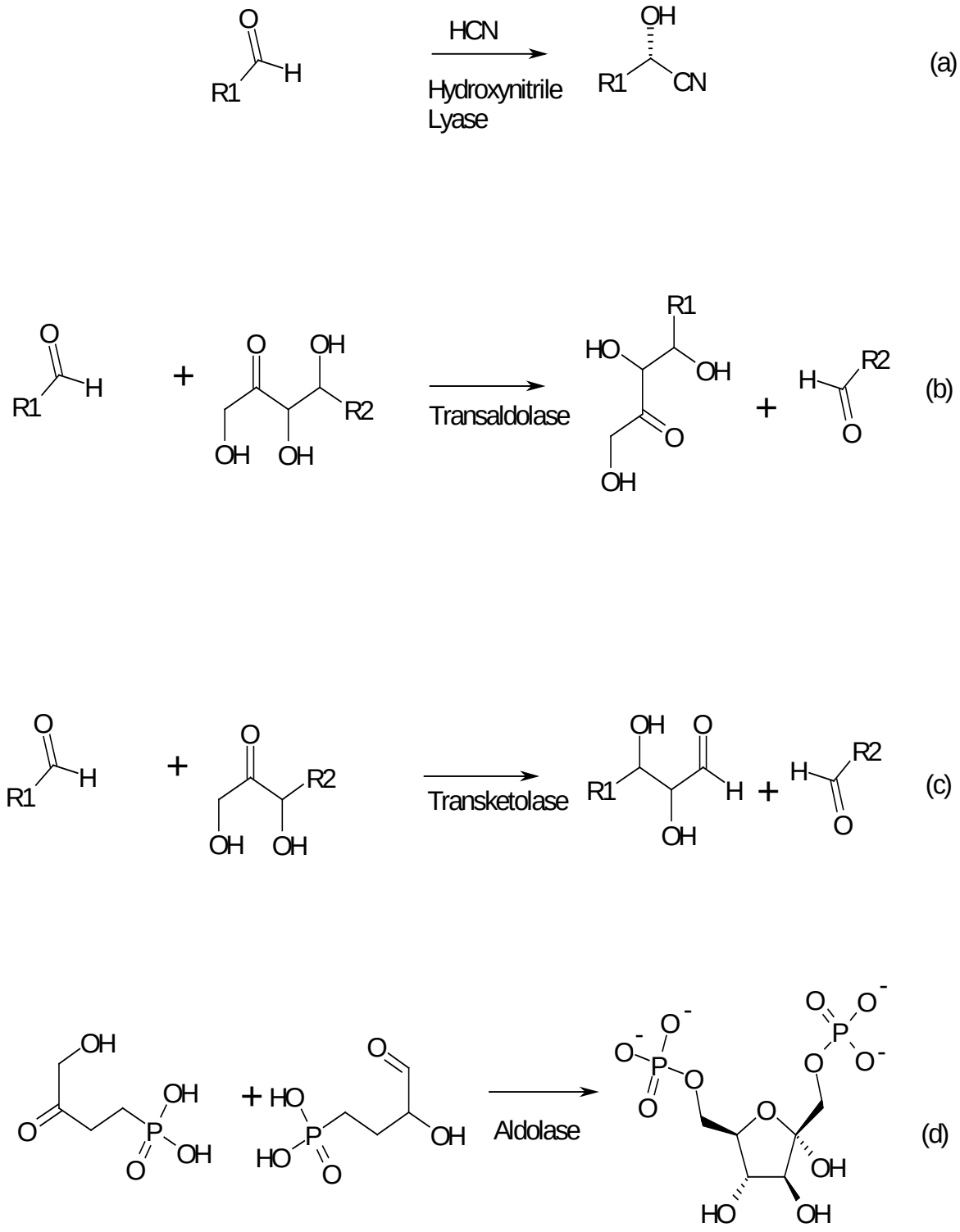


Figure 1.9 Examples of the schematic biocatalytic C-C bond formation reactions. (a) hydroxynitrile lyases, (b) transaldolase, (c) transketolase, and aldolase (d)

1.4 Transketolase

Horecker and Smyrniotis first observed the transketolase enzyme in the 1950's as an extract from rat's liver. They were able to utilise the transketolase ability to catalysis the reaction of sedoheptulose phosphate from pentose phosphate (Horecker and Smyrniotis, 1952) (Fig. 1.10). Later that decade Racker and De la Haba purified the transketolase enzyme from yeast. In doing so they discovered the presence of the cofactor thiamine pyrophosphate (TPP) and the metal ion magnesium. They found that the presence of these two small molecules were vital for the enzymatic activity of the enzyme (Racker *et al.*, 1953). In 1954, Horecker proposed the mechanism for the conversion of pentose phosphate to hexose monophosphate (Horecker *et al.*, 1954). This research allowed greater understanding of the role that transketolase and transaldolase have in the creation of NADPH in the pentose phosphate pathway. Racker and De la Haba conducted further experimentation on transketolase in 1954, producing the first crystalline transketolase enzyme with observation of the 'active glycoaldehyde' using the hydroxypyruvate substrate (de la Haba *et al.*, 1955) (Fig. 1.10).

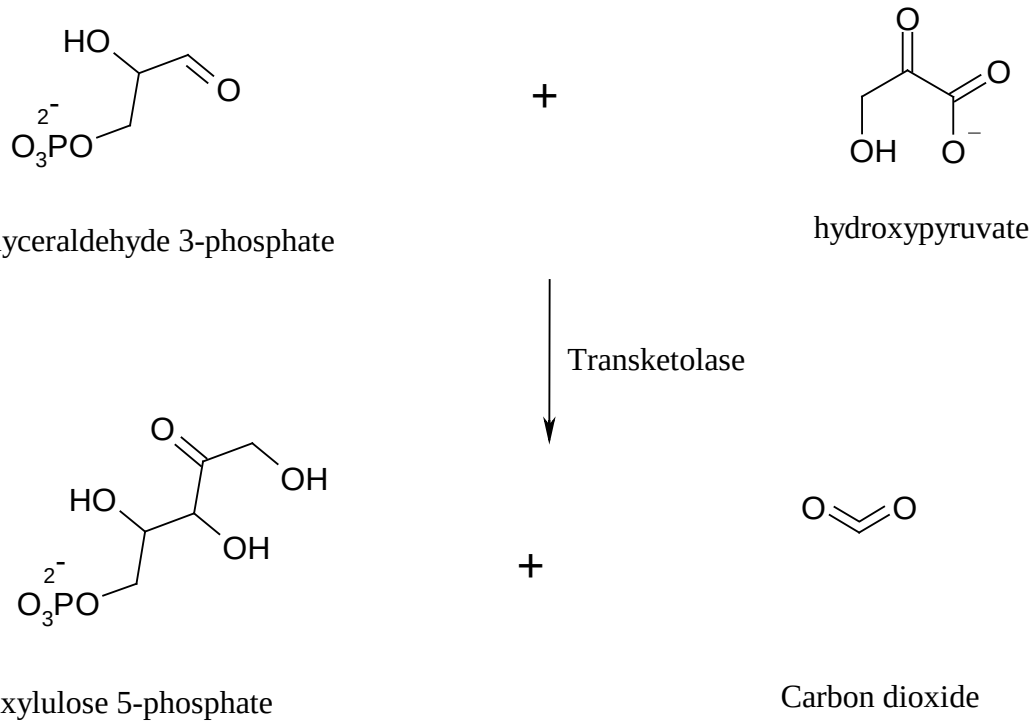


Figure 1.10 The transketolase reaction that occurs with the industrial substrate hydroxypyruvate.

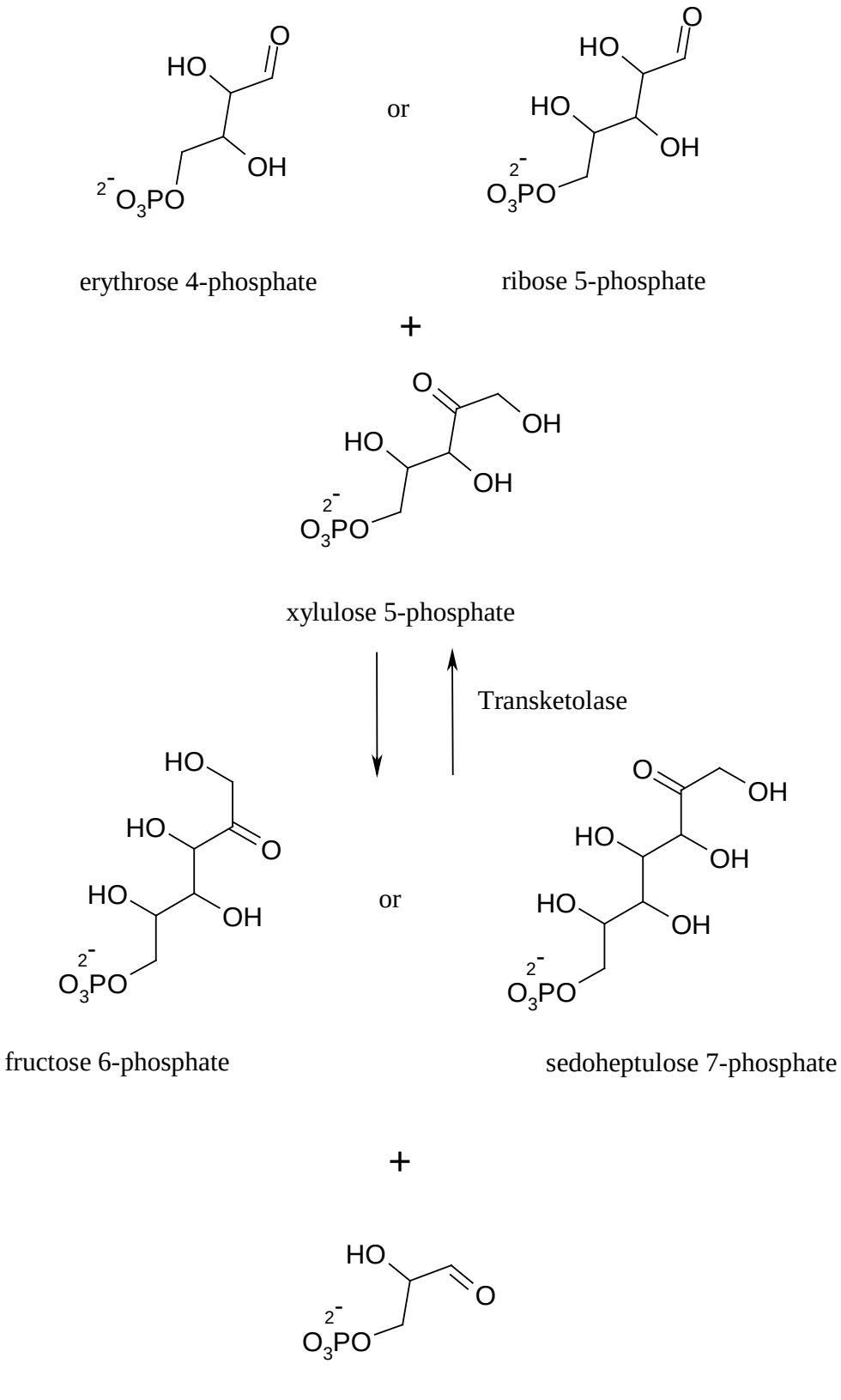


Figure 1.11 The transketolase reaction that occurs in nature, as part of the PPP

1.4.1 Cellular functions

The transketolase enzyme is present in all known organisms, and exists mainly in the cytosol. As discussed previously, the transketolase enzyme catalyses the exchanges of a 2-carbon subunit from one sugar to another. In mammals the products of the transketolase reaction may be used as a metabolite in glycolysis or a precursor for biosynthesis. Thus linking the main pathways of carbohydrate metabolism (production of the energy molecule ATP from glycolysis), and the pentose phosphate pathway (generation of reducing agent NAPDH from glucose 6-phosphate, and the production of the ribose 5-phosphate sugar, the nucleotide precursor). The non-oxidative stage of the pentose phosphate pathway can be seen in figure 1.12.

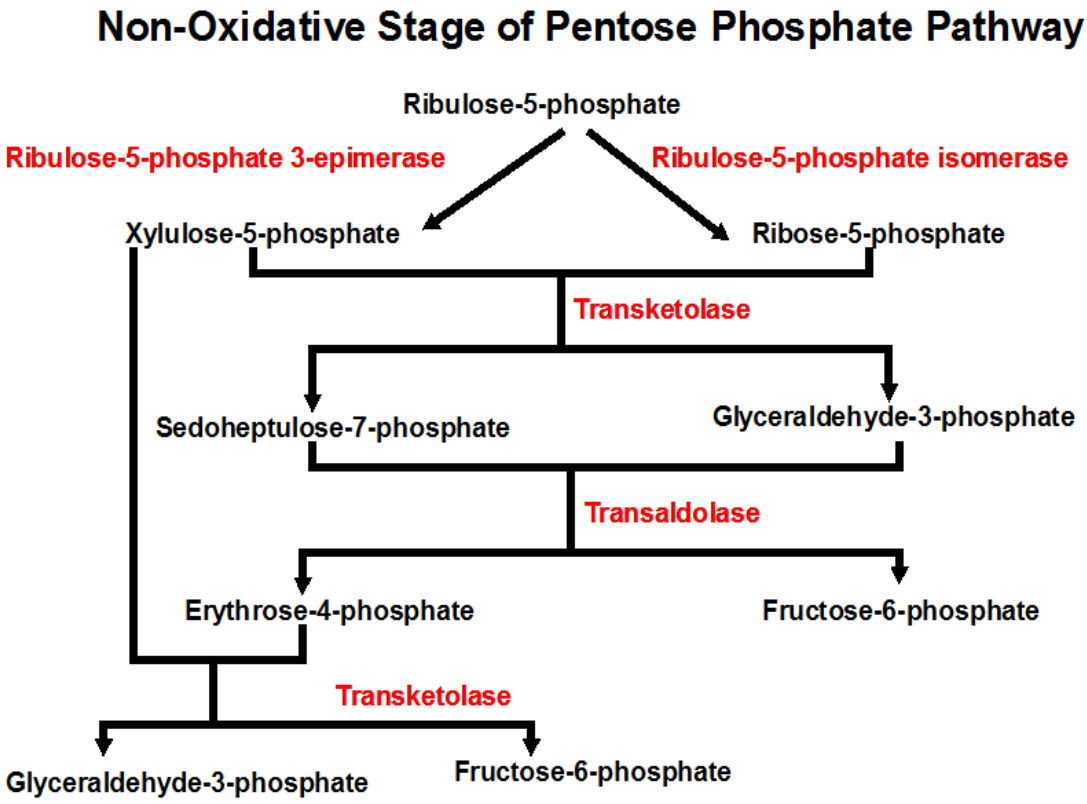


Figure 1.12 The non-oxidative stage of the pentose phosphate pathway.

The transketolase enzyme is believed to be the rate-determining enzyme in the non-oxidative stage of this pathway as shown in the work by Datta and Racker (Datta and Racker, 1961). They reported that the equilibrium constants for the transketolase reaction were close to unity, implying that the relative concentrations of the reactants and products controlled the direction of the reaction, hence satisfying the condition for metabolite flexibility. Dependant on the necessity of the cell, transketolase can play a key role in metabolite flexibility. During cell replication an increased amount of nucleotides are required, hence the transketolase enzyme is sanctioned to produce more ribose 5-phosphate. Transketolase is again regulated by the cell in the breakdown of fat tissues, where the requirement for the increased production of NADPH is necessary to facilitate the energy release. The role transketolase plays in linking the pentose phosphate pathway with glycolysis in mammals is also seen in plants, where transketolase links the Calvin cycle with carbohydrate metabolism. In the Calvin cycle transketolase initially mediates the conversion of fructose 6-phosphate to xylulose 5-phosphate via glyceraldehyde 3-phosphate, and later the production of ribose 5-phosphate and another xylulose 5-phosphate (Bassham *et al.*, 1950). Products of the transketolase enzyme in the pentose phosphate pathway are essential in many other pathways. In particular erythrulose 4-phosphate, which is a precursor for the production of vitamins and aromatic amino acids in the shikimic acid pathway.

1.4.2 Substrate specificity

The transketolase enzymes of plants, yeast and bacteria have all been screened against a range of compounds similar to their natural substrates. It has been found in all sources that the transketolase enzyme has the ability to accept a broad range of substrates. Transketolase from yeast *Saccharomyces cerevisiae* have been seen to accept an array of donor substrates: sugars such as D-xylulose 5-phosphate, D-sedoheptulose 7-phosphate, D-fructose 6-phosphate and D-erythrulose 4-phosphate, and smaller organic molecules like dihydroxyacetone phosphate, dihydroxyacetone and hydroxypyruvate. All of these compounds have a glycolaldehyde group transferred to the acceptor substrate. The yeast transketolase also has have a wide variety of acceptor substrates in the form of D-ribose 5-phosphate, D-glyceraldehyde 3-phosphate, D-erythrose 4-phosphate and glycolaldehyde (Usmanov and Kochetov, 1983). It has been found that transketolase found in plants and bacteria exhibit similar broad substrate specificity to that of the yeast transketolase. This can be observed in spinach leaves where non-phosphorylated sugars were used in conjunction with hydroxypyruvate yielding a transketolase product (Villafranca and Axelrod, 1971). The similarities between the bacteria and yeast transketolase may also be seen in reactions catalysed by *E. coli* (Sprenger *et al.*, 1995). However, mammalian transketolases show a much higher degree of specificity to that seen for the bacteria, plant and yeast enzymes. They use only D-xylulose 5-phosphate, D-fructose 6-phosphate and D-sedoheptulose 7-phosphate as donors, and D-ribose 5-phosphate, D-erythrose 4-phosphate, D-glyceraldehyde 3-phosphate and glycolaldehyde as acceptor substrates (Schenk *et al.*, 1998). The activity in all transketolase enzymes varies depending upon their affinity towards the selected substrate; table 1.1 and 1.2 show the K_m values for a select few transketolase substrates. It can be observed that the affinity for the non-phosphorylated sugar substrates used by the spinach transketolase is rather low. These results may be understood by the fact that the phosphate tail of the substrate enhances stability through additional bonding in the cleft of the active site. Hence the removal of this phosphate tail would therefore lead to a reduced affinity for the substrate, as discussed by (Nilsson *et al.*, 1997).

Source	Human	*S. cerevisiae*	Spinach	*E. coli*
Xylulose 5-phosphate	0.49 mM	0.21 mM	Not available	0.16 mM
Ribose 5-phosphate	0.53 mM	0.4 mM	0.4 mM	1.4 mM
Fructose 6-phosphate	7 mM	1.8 mM	3.2 mM	1.1 mM
Glyceraldehyde 3-phosphate	Not available	4.9 mM	Not available	2.1 mM
Erythrose 4-phosphate	0.36 mM	Not available	Not available	0.09 mM
Hydroxypyruvate	No activity	33 mM	Not available	18 mM

Table 1.1 The K_m values for selected substrates for TK from various sources, representing mammalian, fungal, plant and bacterial enzymes. Data taken from a table compiled by (Schenk *et al.*, 1998).

Substrate	D-ribose	L-lyxose	L-arabinose	D-xylose
K_m value	45 mM	55 mM	120 mM	230 mM

Table 1.2 The K_m values for non-phosphorylated substrates of transketolase from spinach leaves. Data taken from table compiled by (Villafranca and Axelrod, 1971).

1.4.3 Structure

Although transketolase was first classified in the 1950's it was not until the early 1990's that scientists were able to analyse the enzyme's structure. The first crystallographic diffraction data set was collected for the yeast transketolase in the late 1980's (Schneider *et al.*, 1989). The yeast enzyme was the first source of transketolase to have its structure solved (Nilsson *et al.*, 1993; Kochetov, 2001). The structural solution of the yeast transketolase, was followed 2 years later by the *E. coli* transketolase structure (Littlechild *et al.*, 1995). The transketolase enzyme found in maize was isolated leading to the crystal structure in 2003 (Gerhardt *et al.*, 2003). A year later a joint collaboration between groups in Scotland, Argentina and Sweden solved the structure for the transketolase found in the parasite of *Leishmania mexicana* (Veitch *et al.*, 2004). The most recent transketolase structure reported is the structure of the *Thermus thermophilus* transketolase (Yoshida *et al.*, 2007).

All of the known structures are well conserved, and show striking similarity between the various sources. All of the oligomeric transketolase structures exist as homodimers, consisting of two subunits of approximately 70 kDa. Each subunit contains one molecule of thiamine pyrophosphate (TPP) and a metal ion in the form of magnesium or calcium. Even though the structure of the human transketolase has yet to be resolved, it is believed that it also exhibits the same basic structure as the known transketolase enzymes. The general transketolase monomer structure consists of three domains, the first domain referred to as the PP-domain, the second domain dubbed the Pyr-domain, and the third domain known as C-terminal domain. These three domains are α/β folded, and are consistent with that of the general transketolase family. The PP-domain and the Pyr-domain both participate in the binding during the formation of the dimer. Two identical active sites are formed on the interface between the subunits, the active sites contain one TPP molecule and one metal cation. This can be seen in figures 1.13 and 1.14, where the first described transketolase structure from yeast is shown.

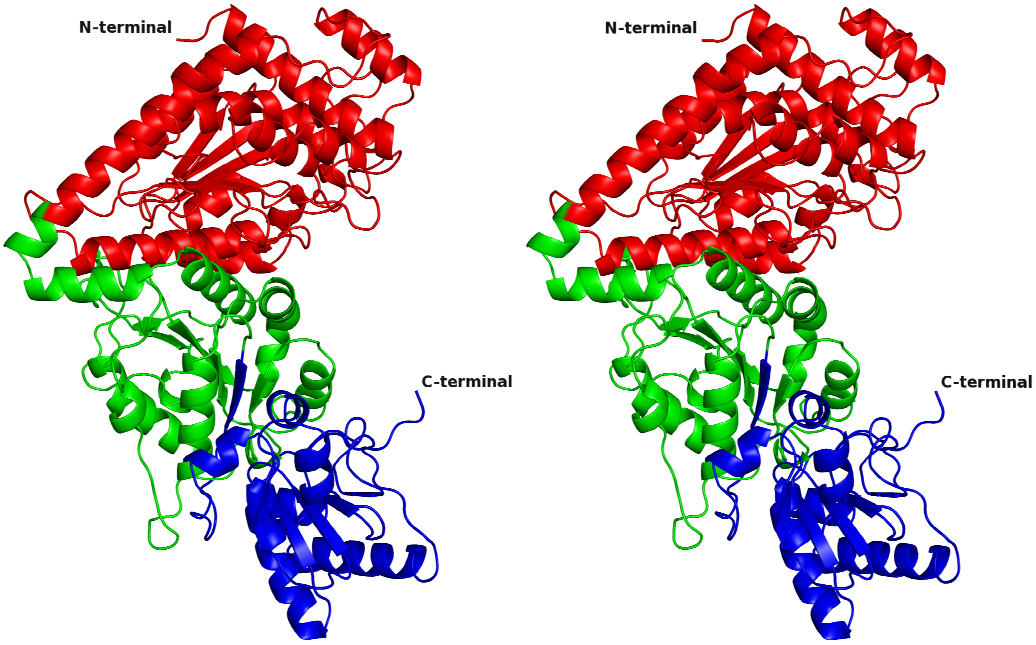


Figure 1.13 Divergent stereo diagram showing the secondary structure of the yeast transketolase monomer (1TRK). The secondary structure domains are displayed, PP-domain (red), Pyr-domain (green) and C-domain (blue). The α-helices are shown as helices and the β-sheets as ribbons. Figure produced in PYMOL (DeLano, 2006).

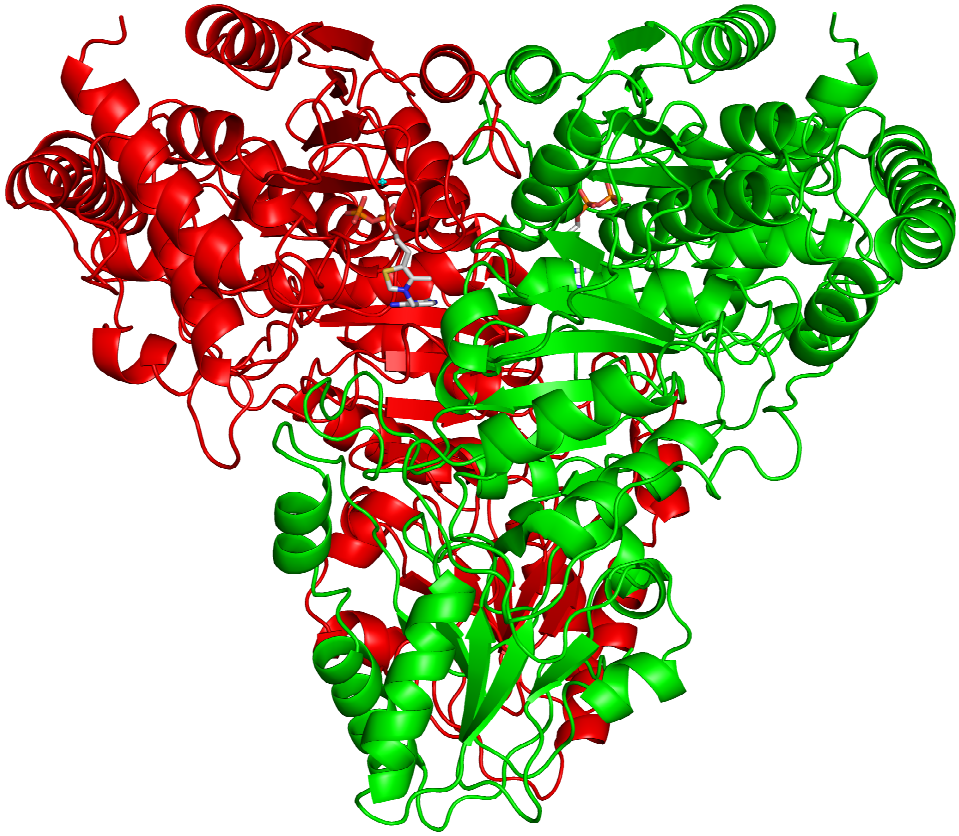


Figure 1.14 Diagram showing the yeast transketolase dimer with two thiamine pyrophosphate molecules. α-helices as helices and the β-sheets as ribbons (1TRK), subunit A (red) and B (green). Figure produced in PYMOL (DeLano, 2006).

1.4.4 Cofactors

All homodimer structures of the transketolase enzyme contain two molecules of thiamine pyrophosphate (Fig. 1.15), and two divalent metal ions. These cofactors are necessary for activity (Racker *et al.*, 1953). Cofactors have been seen to be lost during purification, however addition of these cofactors to the purification buffers has been used to maintain optimal activity of the transketolase during the purification process (Datta and Racker, 1961; Heinrich *et al.*, 1972). The importance of the divalent metal ion has been extensively researched, and has been found to be a prerequisite for the effective binding of the coenzyme, as the affinity of the active site towards the coenzyme is greatly decreased in the absence of the cation (Heinrich *et al.*, 1972; Selivanov *et al.*, 2003). The five known transketolase structures all contain either an Mg^{2+} or Ca^{2+} metal ion. Studies have suggested that the affinity for the apo-transketolase towards the thiamine pyrophosphate is greater in the presence of Ca^{2+} ion than an Mg^{2+} ion (Kovina *et al.*, 1997; Selivanov *et al.*, 2003).

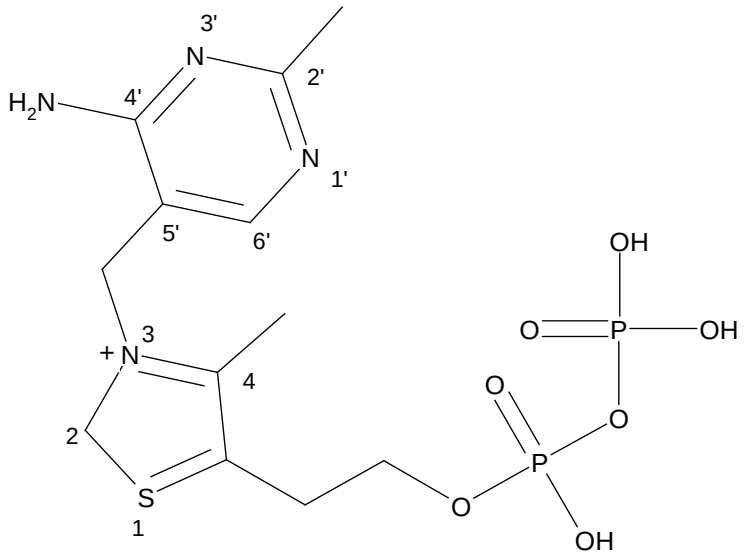


Figure 1.15 Shows a representation of the thiamine pyrophosphate molecule. Thiazolium ring number 1-5. Aminopyrimidine number 1'-6'.

1.4.5 Stability and optimum pH

The transketolase enzyme is relatively stable in comparison to other enzymes. The stability of the enzyme centres around the presence of the cofactors, which reside in the active site cleft that is formed on the interface of the subunits. The thermal stability of transketolase enzymes has been assessed finding the enzyme to be notably more stable in the holo-form opposed to the apo-form and even more so in the presence of Ca^{2+} ion than Mg^{2+} ion (Esakova *et al.*, 2005). The *T. thermophilus* transketolase can operate at a higher temperature than any other transketolases. The majority of transketolase enzymes have optimum activity between pH 7 and 9 depending on the buffer used.

1.4.5 Catalytic mechanism

As discussed, transketolase mediates the transfer of a glycolaldehyde moiety from a ketose donor to an aldose acceptor. The proposed mechanism for the catalytic pathway of transketolase has been constructed based largely on the studies of the yeast transketolase enzyme. The thiamine pyrophosphate and the divalent metal ion are bound in a hydrophobic pocket, with only the C2 of the thiazolium ring accessible by the substrate. During the transfer of the two-carbon unit moiety an 'active glycolaldehyde' intermediate is formed, which marks the halfway point of the reaction. There are several key steps involved in the reaction. (Fig. 1.16) (Schneider and Lindqvist, 1993; Wikner *et al.*, 1995; Gyamerah and Willetts, 1997; Nilsson *et al.*, 1997).

1) The invariant glutamate residue at 418 is in a position to form a hydrogen bond with the N1' of the aminopyrimidine ring, this in turn promotes resonance within the aminopyrimidine ring forming the positively charged imino-group at the 4' ring position.

2) The thiazolium ylide is now primed. The donor substrate (fructose-6-phophate) enters through the active site opening, and proceeds down the channel toward the C2 position of the thiazolium ylide.

3) With the donor substrate intermediate formed, the release of the first product is facilitated by the removal of the proton from the Cγ hydroxyl group of the donor substrate by the histidine residue 263. The product is released and passes back out of the channel and exits via the entrance the substrate has entered.

4) The halfway intermediate is now formed, named 'active glycolaldehyde' or α,β-dihydroxyethyl–TPP.

5) The acceptor substrate enters through the active site opening, and proceeds down the channel towards the 'active glycolaldehyde' intermediate.

6) The 'active glycolaldehyde' is a carbanion/enamine that performs nucleophilic attack upon the Cα carbonyl of the acceptor substrate, with the Oα of the carbonyl being in turn protonated by the histidine residue 263. A reverse of the donor reaction.

7) The acceptor substrate is now covalently bound to the α,β-dihydroxyethyl–TPP intermediate. The reverse process now continues, the aminopyrimidine moiety deprotonates the Cβ hydroxyl group of the donated two-carbon subunit. This causes the

covalent bond between the C2 of the thiazolium ring and the Cβ of the substrate to break.

8) This yields the second product that has a longer carbon backbone by two carbons.

9) The thiazolium ylide is regenerated and may catalyze further reactions, or return to its non-resonance state.

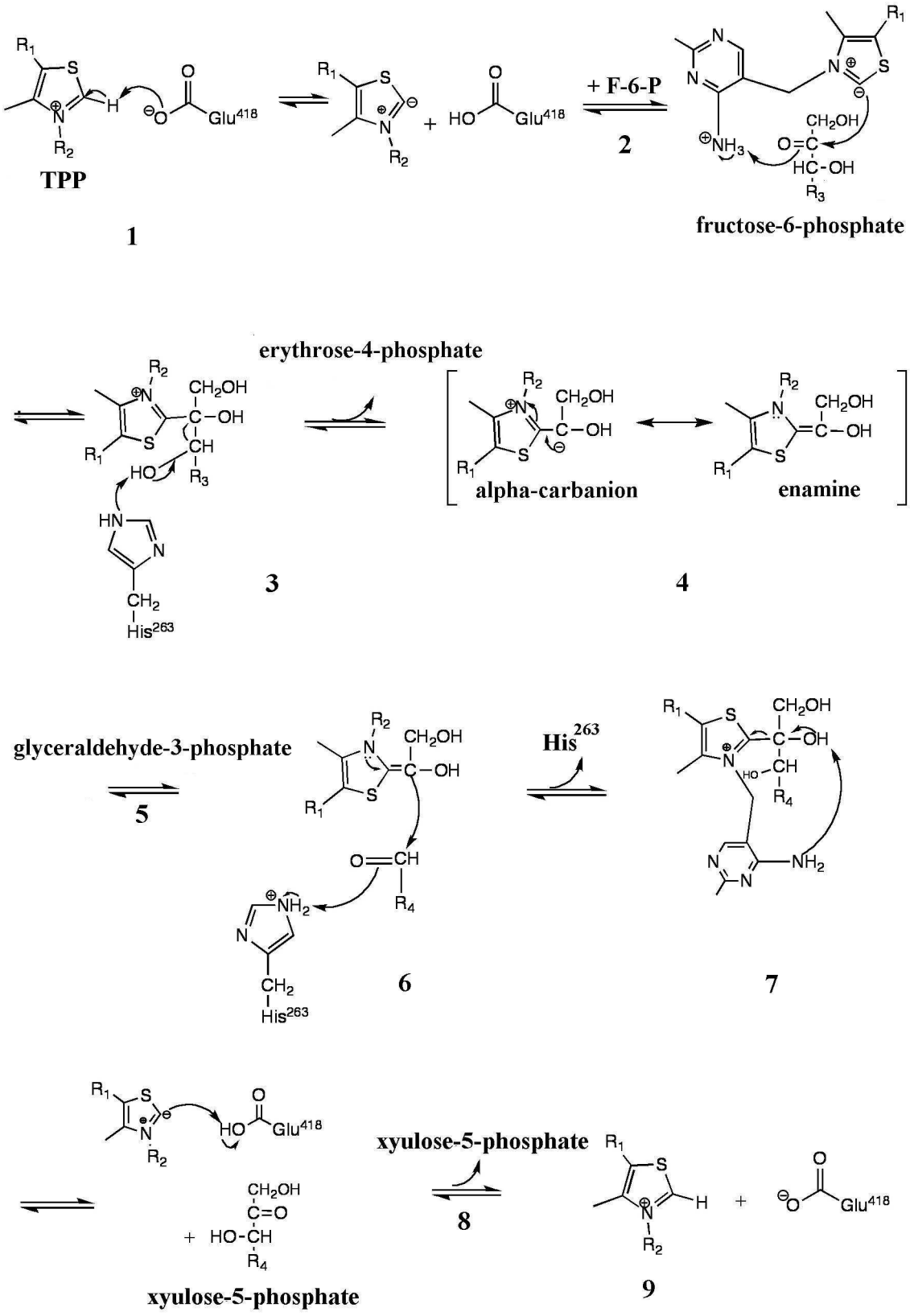


Figure 1.16 Shows the proposed transketolase mechanism, with the donor substrate and acceptor substrate (Nilsson *et al.*, 1997).

1.4.6 Inhibition

It has been observed how important transketolase is in the cellular functions of all organisms. If inhibition of transketolase were to occur, a number of precursor molecules would not be synthesised, and this in turn would bring about the standstill of the particular reaction pathway. Although it is true that cells have more than one route of synthesis for the required precursor molecules, the inhibition of transketolase would have a consequential effect on the pentose phosphate pathway.

One of the roles of transketolase in the pentose phosphate pathway is in the production of nucleic acid precursors during cell division. In cancerous tumor cells, transketolase uses glucose through the elevated non-oxidative pentose phosphate pathway to produce these nucleic acids. Identification of inhibitors that could specifically target transketolase and prevent the non-oxidative pentose phosphate pathway from generating the RNA ribose precursor, ribose-5-phosphate, would afford a suitable and novel approach for an effective anticancer therapeutic agent (Du *et al.*, 2004). Du *et al.* identified active scaffolds of small molecule inhibitors for the transketolase enzyme, as a starting point for further research in the design of a pharmaceutical for use in anticancer therapy. Comin-Audriux *et al.* used Avemar (Biromedicina, Co. (Budapest, Hungary)) in studies of fermented wheat extract, which indicate that Avemar is a powerful inhibitor of *de novo* nucleic acid synthesis (Comin-Anduix *et al.*, 2002). The structure of Avemar is protected by a chemical transfer agreement. Another study was undertaken by Rias *et al* into the inhibition of the transketolase in an effort to restrict the growth of tumor cells and were based around the two compounds oxythiamine (Fig. 1.17) and dehydroepiandrosterone (Fig. 1.18). They concluded that both oxythiamine and dehydroepiandrosterone reduce the growth of cancerous cell with a non-toxic effect towards the host mice. The oxythiamine was seen to be more effective than dehydroepiandrosterone, however both were deemed suitable as possible target scaffolds for future drug development (Raïs *et al.*, 1999).

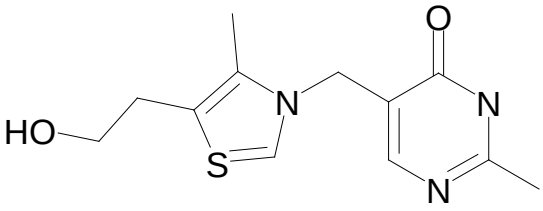


Figure 1.17 The structure of the possible anticancer drug oxythiamine (Raïs *et al.*, 1999).

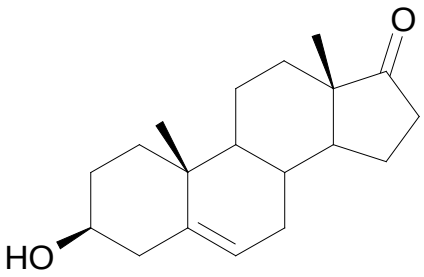


Figure 1.18 The structure of the possible anticancer drug dehydroepiandrosterone (Raïs *et al.*, 1999).

In 1988 N-acetylimidazole (Fig. 1.19) was studied as a possible inhibitor for transketolase. It was observed that the activity of the baker's yeast transketolase was rapidly reduced in the presence of the proposed inhibitor. It was interpreted from the kinetic data that the acetylation of one amino acid per active site was enough to remove the enzyme's activity (Kuimov *et al.*, 1988).

An earlier example of the use of a pharmaceutical in the inhibition of transketolase can be seen by the studies of the Astra Zeneca drug omeprazole (Fig. 1.20) by Nixon, Diefenbach and Duggleby (Nixon *et al.*, 1992). They found that the inhibition of transketolase caused by omeprazole was a complex affair, compared with the drug's inhibition of pyruvate decarboxylase that acted in a competitive manner with the coenzyme.

Kochetov *et al.* utilised the substrate analogue *p*-hydroxyphenylpyruvate (Fig. 1.21), which proved to be a reversible and competitive inhibitor of transketolase with regards to the substrate used. It was also able to displace thiamine diphosphate from the holo-transketolase. The data suggest that *p*-hydroxyphenylpyruvate participates in the regulation of tyrosine biosynthesis by influencing the catalytic activity of transketolase (Solovjeva and Kochetov, 1999).

The large range of inhibitors used against the transketolase enzyme has shown many mechanisms in which they may inhibit. Some have shown competitive inhibition towards the thiamine-based cofactor such as omeprazole and oxythiamine, whilst others have acted in a manner to alter the shape of the active site via acetylation (N-acetylimidazole). The other mode of inhibition seen is to mimic the substrate as with *p*-hydroxyphenylpyruvate, which operates by competitive inhibition against the donor substrate. This is a form of reversible inhibition as on removal of the *p*-hydroxyphenylpyruvate full activity of the transketolase is regained.

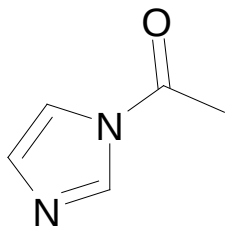


Figure 1.19 The structure of the inhibitor N-acetylimidazole (Kuimov *et al.*, 1988).

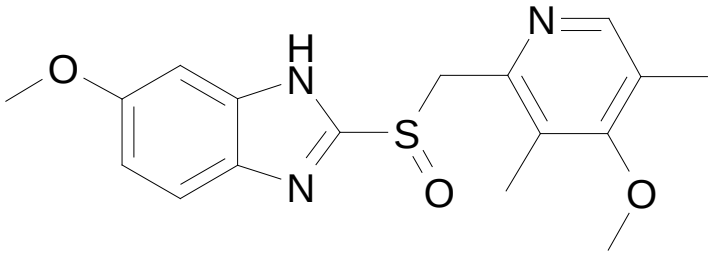


Figure 1.20 The structure of the inhibitor the Astra Zeneca drug omeprazole (Nixon *et al.*, 1992).

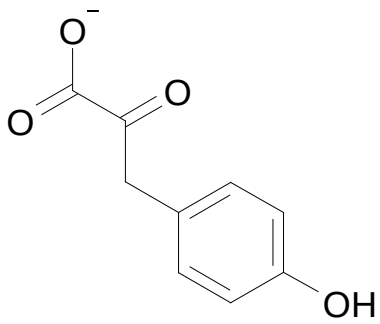


Figure 1.21 The structure of the inhibitor substrate analogue *p*-hydroxyphenylpyruvate (Solovjeva and Kochetov, 1999).

1.4.7 Use as a biocatalyst

When deciding on whether an enzyme is suitable to be used as a biocatalyst, a number of parameters must be satisfied in order to assure reproducible results. Factors that influence the suitability of an enzyme's potential to become a biocatalyst focuses around substrate specificity, the over-expression of the gene in a suitable host, and the potential for scale-up.

Transketolase over the past 20 years has become a useful biocatalyst, which is due to a number of reasons. One of the key factors regards the chemistry that transketolase undertakes; the stereospecific formation of C-C bonds is a useful tool in any chemical synthesis. Another advantage is that the enzyme is very specific, but can accept a broad range of substrates (Sprenger *et al.*, 1995). The recent advances in microbiology have given rise to an efficient over-expression system of the *E. coli* enzyme, which has allowed this transketolase to be produced on much larger scale (French and Ward, 1995).

Transketolase has recently been used to good effect in organic synthesis. There has been great interest in the use of the transketolase in biocatalysis for preparing flavour and fragrance components due to their ability to produce 'natural' molecules that command a premium price as food additives (Turner, 2000). Transketolase from spinach was employed in a chemoenzymatic synthesis for 6-deoxy-L-sorbose, which is a known precursor of furaneol, a compound that possesses a caramel-like flavour (Hecquet *et al.*, 1996). An application with the use of hydroxypyruvate as the ketol donor can be seen for the *E. coli* transketolase, which facilitates the pairing of racemic 3-*O*-benzylglyceraldehyde and the hydroxypyruvate, forming 5-*O*-benzyl-D-xylulose, on a multi-gram scale (Morris *et al.*, 1996) (Fig. 1.22). Another synthesis that utilises the release of carbon dioxide to drive the process to completion is seen in the coupling of hydroxypyruvate to ribose 5-phosphate, producing highly pure D-sedoheptulose-7-phosphate in a one step reaction on a multigram scale (Charmantray *et al.*, 2009) (Fig. 1.23). It may be noted that the transketolase enzyme has been immobilised upon a commercial support. The immobilisation of the enzyme has seen an increase in the stability of the enzyme (Brocklebank *et al.*, 1999).

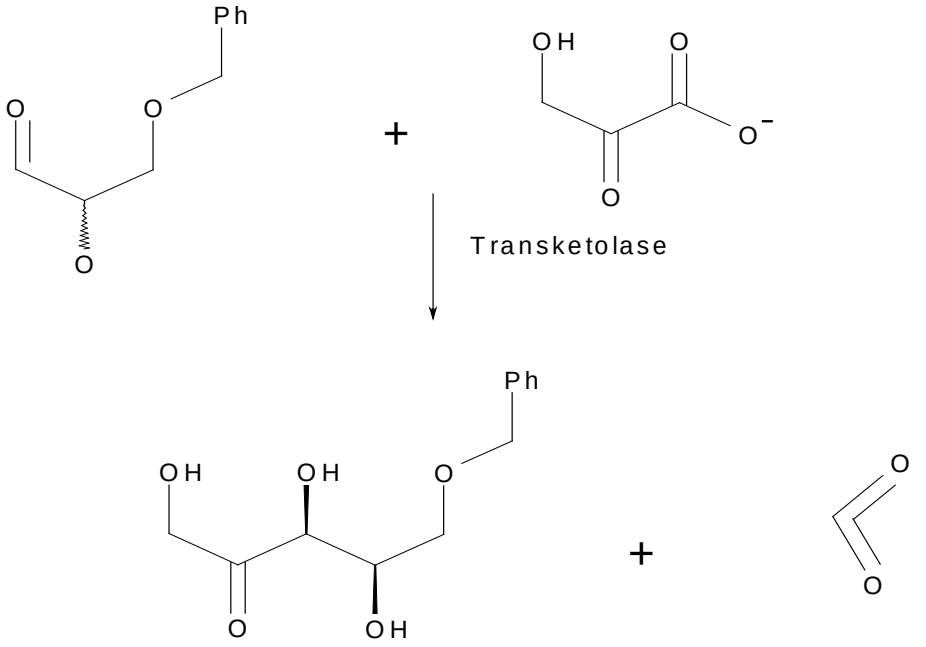


Figure 1.22 The conversion of the racemic 3-*O*-benzylglyceraldehyde into 5-*O*-benzyl-D-xylulose using hydroxypyruvate by transketolase with the release of carbon dioxide (Morris *et al.*, 1996).

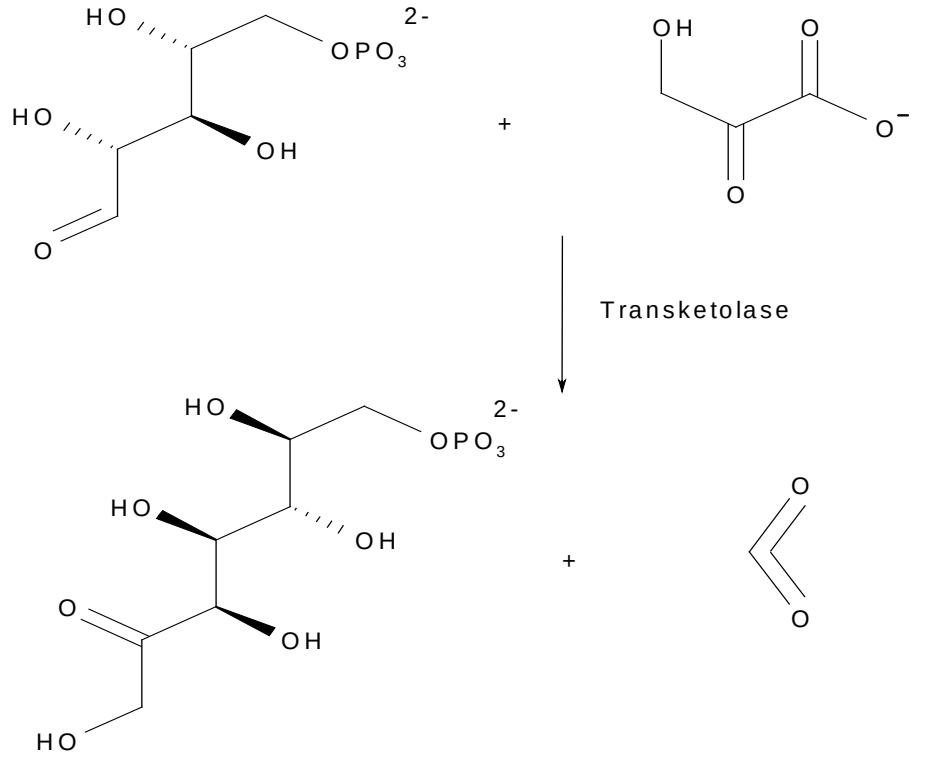


Figure 1.23 The coupling of hydroxypyruvate to ribose 5-phosphate producing sedoheptulose-7-phosphate (Charmantray *et al.*, 2008).

1.4.7.1 Industrial applications

In order for an enzyme to make the leap from a conventional bench top biosynthetic reaction to a high performing industrial biocatalyst, it must conform to the unwritten rules of industry. Industry requests that the biocatalytic reaction must be able to be performed on a large scale, thus reducing the cost of the final product. For this to happen for transketolase the scale of production had to change. French and Ward were able to improve the stability of the transketolase for use in large-scale production (French and Ward, 1995). Then Hobbs *et al.* devised an efficient method for the production and recovery of transketolase on a kilogram scale (Hobbs *et al.*, 1996). With the new found large-scale production of transketolase a method for the biotransformation had to be construed, which was undertaken initially by Lilly *et al.* Layond and Woodley suggested a method where the transketolase product should be removed on production, so as to avoid unnecessary side reaction that could cause inhibitory effects (Lye and Woodley, 1999).

The role of transketolases in the pentose phosphate pathway makes it a prime candidate to mediate the possible conversion of abundant carbohydrate substrates to a more lucrative product such as ethanol via some form of fermentation methodology. So far all known efforts for this biotransformation have failed (Walfridsson *et al.*, 1995; Zhang *et al.*, 1995). Transketolase has found a number of areas of industrial processes that require the biocatalytic route. An example is the biosynthesis of the aromatic amino acids L-phenylalanine, L-tryptophan and L-tyrosine from D-glucose (Ganem, 1978). These amino acids have been found to be valuable precursors in the synthesis of organic product. The amino acid L-phenylalanine formed in the biotransformation can be converted into the artificial sweetener aspartame (Vinick and Jung, 1982). In summary, the transketolase enzyme may be used as an industrial biocatalyst for the production of fine chemicals. This is possibly due to the physical size of the active site and enzyme inability to accept much larger molecules. The syntheses that the transketolase enzymes are occupied in mainly involve the production of precursors, as opposed to the final product. The productivity of transketolases within the industrial environment has only just begun. The production of mutant transketolase will lead to the further applications and adaptation of the enzyme. Additionally, the progress made within the bioreactor field, combined with increased scale-up techniques (Charmantray *et al.*, 2008;

Matosevic *et al.*, 2008) makes transketolase an exciting target for further routes of biotransformation.

1.5 Project aims

The aim of the project is to study the *E. coli* transketolase enzyme. Although the transketolase enzyme has been studied previously, complete understanding of the mechanistic pathway requires further probing. This would be of benefit for both academia and industry, as with further knowledge in understanding the mechanism comes greater awareness of feasible substrate molecules, inhibitor design and potential active site mutations to redesign the enzyme. The aim of the study is to probe the mechanistic details of the *E. coli* transketolase enzyme by structural and kinetic analysis and to analyse the interactions made by the substrates/inhibitors during this process. A detailed structural investigation of the *E. coli* transketolase enzyme and active site variants that produces an alternative enantiomer will be initiated in order to understand the route by which the transketolase enzyme achieves its enantio-selectivity. The transketolase enzyme produces the L-enantiomer product, where as the transketolase variants produce the D-erythulose product. The results of these studies will lead to new potential substrates of the enzyme, and allow its optimisation for use in the fine chemical industry. An improved understanding of the enzyme mechanism will assist in the design of new therapeutic inhibitors as anti-cancer drugs.

Chapter 2

Materials and Methods

2.1 Purification

General procedures

The general procedures described in this chapter were used for all enzymes studied in this thesis. Where the procedures differ with a specific enzyme the details are given in the appropriate sections.

2.1.1 Reagent grade materials

Unless stated otherwise all chemicals used are of the highest purity available supplied from Sigma-Aldrich Chemical Company (Poole, UK). The water used throughout this work was of a double distilled quality purified by a Purite™ System (Oxon, UK). The SDS-PAGE gel apparatus was obtained from Bio-Rad Laboratories (Hertfordshire, UK). The Dialysis tubing was obtained from Medicell International Ltd.

2.1.2 Buffers

Unless stated, Buffer A was used in the general purification processes. All buffers were filtered and degassed prior to use in the purification experiments.

Buffer A
10 mM Tris-HCl buffer pH 7.5.
1 mM EDTA.
6 mM β-Mercaptoethanol.
1 $x10^{-5}$ M PMSF.
2 $x10^{-5}$ M BAM.

Buffer B

10 mM Tris-HCl buffer pH 7.5.

0.5 M NaCl.

1 mM EDTA.

6 mM β-Mercaptoethanol.

1 $x10^{-5}$ M PMSF.

2 $x10^{-5}$ M BAM.

2.1.3 Protein concentration estimation

The protein concentration was measured at every stage of the purification by the following methods:

2.1.3.1 Spectrophotometric method (Warburg and Christian method)

A 1 ml aliquot of protein solution was placed in a quartz cuvette of 1 cm pathlength and measured against the corresponding blank buffer at 280 nm. All measurements were made using a Shimadzu double beam UV-2100. The extinction coefficient of the protein was obtained using Protpram (Gasteiger *et al.*, 2008). The protein concentration was then calculated using the Beer-Lambert equation ($A = \varepsilon[c]l$).

2.1.3.2 Bio-rad standard assay procedure (Dye binding method)

The dye reagent (Coomassie Blue G-250) required for this method was purchased from Bio-Rad Laboratories Ltd (Hertfordshire, UK). The Coomassie Blue G-250 was diluted to a ratio of 1:3 with distilled water.

Protein solution (0.1 ml) was added to the diluted solution of Coomassie Blue G-250 (0.3 ml) and allowed to mix for 10 minutes. The sample was then measured at 595 nm using BMG Lab-Tech FLUOstar OPTIMA plate reader. The values obtained were compared against a standard curve of increasing concentrations of bovine serum albumin in a standard curve graph.

2.1.4 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Denaturing SDS-PAGE was used as a method for measuring the purity of the protein samples. These gels allow identification of the apparent molecular weight of protein polypeptide chains and indicate the presence of contaminating proteins and other impurities such as degradation products. The apparatus used in the separation was the Bio-Rad Mini-Protean II gel apparatus (Bio-rad Laboratories Ltd., Hertfordshire, UK.)

2.1.4.1 Stock solutions for SDS PAGE

1. Acrylamide solution: 30% (*w/v*) acrylamide, 1% (*w/v*) N,N'-methylene bisacrylamide in water.
2. Separating gel buffer: 1.5 M Tris-HCl, pH 8.8.
3. Stacking gel buffer: 0.5 M Tris-HCl, pH 6.8.
4. SDS-PAGE running buffer: 0.05 M Tris-HCl, 0.38 M Glycine, 0.1% (*w/v*) SDS, adjusted to pH 8.8.
5. SDS solution: 10% (*w/v*) SDS in distilled water.
6. Ammonium persulfate: 10% (*w/v*) in distilled water.
7. TEMED.
8. Sample-loading Buffer: 125 mM Tris-HCl, pH 6.8, 10 mM EDTA, pH 6.8, 0.1% (*w/v*) SDS, 20% (*w/v*) glycerol, 0.05% (*w/v*), 0.001% *(w/v)* Bromophenol blue tracking dye, 0.005% (*w/v*) 2-mercaptoethanol.

2.1.4.2 SDS-PAGE Gel preparation

The gels were assembled according to the manufacturer's instructions, using gel preparative glass plates, measuring 73 mm x 102 mm (inner) and 83 mm x 102 mm (outer), with spacers of 1mm thickness inserted between the plates.
The separating gel consisted of Acrylamide solution (4.2 ml), Separating gel buffer (2.5 ml), Purite water^{TM} (2.3 ml), 10% SDS (1.0 ml), 10% ammonium persulfate (100 µl) and TEMED (10 µl). The solution was mixed thoroughly, avoiding frothing, and 3.25 ml was immediately pipetted between the gel plates. The gel was overlaid with a 1 ml layer of water-saturated isobutanol, and left to polymerise at room temperature. The

1 ml layer of water-saturated isobutanol was poured off and the gel washed with distilled water.

The stacking gel consisted of acrylamide solution (2.0 ml), stacking gel buffer (2.5ml), Purite water^{TM} (4.5 ml), 10% SDS (1.0 ml), 10% ammonium persulfate (100 µl) and TEMED (10 µl). The solution was mixed thoroughly, avoiding frothing, and was immediately pipetted to the top of the plates. A 10 well comb was inserted to the top of the plates, and the gel was left to polymerise at room temperature.

Once the stacking gel had set, the comb was removed and the gel installed into the electrophoresis chamber apparatus according to the manufacturer's instructions. The chamber was then filled with the reservoir buffer solution.

2.1.4.3 Molecular weight protein standards

Precision Plus protein unstained standards (Bio-Rad Laboratories Ltd) containing recombinant proteins of 10 kDa, 20 kDa, 25 kDa, 37 kDa, 50 kDa, 75 kDa, 100 kDa, 150 kDa and 250 kDa molecular weight. These were used as reference for the proteins molecular weight.

2.1.4.4 Sample preparation

20 µl of a protein sample was added to 10 µl of the sample loading buffer in a 1.5 ml Eppendorf tube. To denature the protein, the samples were heat treated at 105 °C for 5 minutes, using a Techni Dri-block DB-2A.

2.1.4.5 Gel electrophoresis

The 20 µl protein samples and 10 µl of the LMW markers (Bio-Rad Laboratories Ltd, Hertfordshire, UK) were loaded into the wells of the gel using a 50 µl Hamilton microsyringe. A constant voltage of 200V was applied across the gel for approximately 45 minutes, until the bromophenol blue tracking dye reached the bottom of the gel.

2.1.4.6 Gel staining and destaining

The gel was place in a plastic container with 100 ml of gel staining solution containing Coomassie brilliant blue G250 (0.25 g), Methanol (250 ml), water (200 ml), and acetic

acid (50 ml). The container was then heated in a microwave oven for 3 minutes. The gel was then removed and placed in a plastic container containing Purite™ water (1 L) and heated for a further 10 minutes. The resultant gel gave qualitative data about the protein samples.

2.1.5 Dialysis tubing

The dialysis tubing was prepared by boiling in sodium bicarbonate for a period of 15 minutes, and then the tubing was boiled in 0.1 M EDTA for 10 minutes. The tubing was rinsed in water and stored in 0.5% sodium azide at 4 °C to prevent microbial contamination. The dialysis of protein samples was carried out to remove salts and other low molecular weight impurities from the protein solution. When required the tubing was cut to length and rinsed thoroughly with water. The tubing was tied at one end and fastened with a clip. The protein solution was poured in, after removal of air, a knot was tied at the other end and a clip was applied. The tubing was then submerged in a low salt buffer (2 x 5 L) and left for 24 hrs (2 x 12 hrs) at 4 °C, changing the buffer once halfway through the dialysis.

2.1.6 Concentration of protein solutions

Proteins were concentrated using Vivaspin concentrators obtained from Vivascience (UK). The concentrators were of a range of sizes, 2 ml (30,000 MWCO), 5 ml (10,000 MWCO) and 20 ml (10,000 MWCO). The protein solution was placed in the appropriate concentrator and centrifuged at 5000 g until the required concentration was attained. The membrane compartment retained the protein, whilst the excess buffer solution was collected in the lower compartment and discarded.

2.1.7 pH Measurement

The pH of solutions were tested using a Whatman PHA 2000 system calibrated with BDH ‘colourkey’ buffer solutions of pH 4.0, 7.0 and 10.0. HCl or NaOH was used to adjust the buffer solution to its correct pH.

2.1.8 Conditions for the growth of microorganisms

The *E. coli* transketolase gene was over expressed in the plasmid pQr 711 in *E. coli* strain K12 bacterial JM107 along with the active site variant proteins H26Y and D469Y (Fig.2.1). It was provided by Prof. John Ward of University College London (UK), as a glycerol stock, which was stored at -80°C until needed.

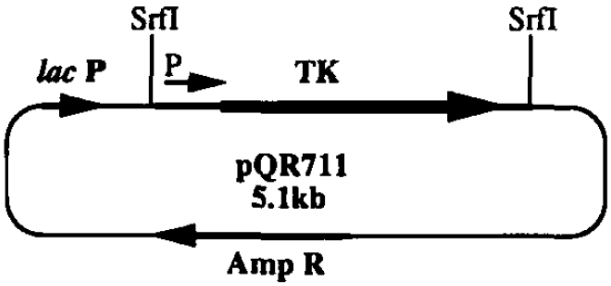


Figure 2.1 Shows the restriction enzyme map of transketolase-expressing plasmid. The arrow represents the orientation of the gene and P represents the *tkt* promoter (Ward and French, 1996).

In expression of the TK gene, the lacZ gene of the *lac* operon is replaced by the TK gene. The IPTG is used to mimic allolactose, which induces the *lac* operon. IPTG acts by inducing the activity of beta-galactosidase, this enzyme promotes the consumption of lactose, by binding and inhibiting the lac repressor thus inactivating it and allowing the TK gene to be expressed. IPTG is used in place of allolactose as it can't be broken down in the cell, so the TK gene is continued to be expressed.

2.1.9 Growth of microorganisms

Into a 250 ml conical flask containing 100 ml of Liquid Broth (LB)(Tryptone 10 g/l, Sodium Chloride 10 g/l, Yeast extract 5 g/l) (0.24 g/ml), ampicillin (0.1 mg/ml) was added. The flask was then inoculated with *E. coli* over expressing the enzyme of interest and incubated at 37 °C for 24 hrs in an orbital shaker at 150 r.p.m. The contents were then aseptically added to two sterile 2L conical flasks containing 1L of Liquid Broth (LB) (0.24 g/ml) and ampicillin (0.1 mg/ml). The flasks were incubated for approximately 6 hrs until the optical density (A_{595}) reached 0.7-1.0, then isopropyl-beta-D-thiogalactopyranoside (IPTG) (1 mmol/L) was added and incubated for a further 6

hrs. The cells were then harvested using a Sigma Howe 6K10 centrifuge for 20 minutes at a speed of 13500 rpm. The pellet was retained whilst supernatant was autoclaved and discarded.

2.1.10 Protein purification

2.1.10.1 Sonication of cells

The harvested cells were resuspended in Buffer A (250 ml) on ice. Using a Sanyo Soniprep 150 MSE the cells were sonicated in ice at an amplitude of 20 microns for 30 seconds. The suspension was left to cool for a further 60 seconds, which was then followed by a further 30 seconds of sonication. This was repeated approximately 8 times, frothing was avoided throughout the process. The resulting homogenate was centrifuged using a Beckman GS-15R centifuge at 12500 rpm for 20 minutes at 6 °C. The cell debris pellet was discarded and supernatant retained.

2.1.10.2 Ammonium sulfate fractionation

Ammonium sulfate fractionation separates on the principles of the solubility of the protein at high ionic strengths of solution by competing with water. At high ammonium sulfate concentrations the natural tendency for proteins not to aggregate is overcome, as the surface charge of the protein has become neutral, hence allowing protein aggregation, which then can be removed by centrifugation.

Initial trials were carried out varying the ammonium sulfate saturation 20 %, 40 %, 60 % and 80 % to establish the precipitation point of the transketolase enzyme. Enzyme grade ammonium sulfate was added gradually to a stirring solution of crude enzyme extract, until the solution became 40 % saturated. The resulting solution was stirred for a further 30 minutes and then centrifuged at 14000 g for 20 minutes. The pellet was discarded, and to the supernatant enzyme grade ammonium sulfate was added gradually to the stirring solution until the solution became 80% saturated. The solution was again stirred for another 30 minutes and centrifuged at 14000 g for 20 minutes. The supernatant was discarded and the pellet retained and stored at 4 °C.

2.1.10.3 Ion-exchange chromatography

Ion-exchange media such as Fast Flow-Q™ separate on the principles of the different net charges on proteins at a certain pH and interact with the proteins, principally by electrostatic attraction. The bound proteins are eluted off the column by increasing ionic strength of the buffer, which competes with the protein for the sites on the ion-exchanger. Protein separation is based on their ability to bind to the ion-exchanger.

Ion-exchange chromatography was performed using a Pharmacia FFQ anion exchange column (column volume 80 cm^3) with a flow rate of 4 ml/min.

The pellet from the ammonium sulfate precipitation was re-dissolved in buffer A (5 ml) and dialysed for 24 hrs in 2 x 5 L aliquots of Buffer A as described in 2.1.5, and loaded onto the FFQ column pre-equilibrated with two column volumes of Buffer A. The column was then washed further, with 100 ml of buffer A to remove all unbound materials. A gradient of increasing concentration of Buffer B was applied to the column over 500 ml. 10 ml fractions were collected, and a sample of each fraction run on a SDS-PAGE gel. The cleanest fractions were combined and subjected to an 80 % ammonium sulfate precipitation. The suspension was centrifuged at 14000 g the supernatant discarded and the pellet stored at 4 °C.

2.1.10.4 Gel-filtration chromatography

Gel filtration media separates on the principles of differing molecular sizes. The separation between molecules of different sizes is based on the ability of the molecule to enter the pores in the media beads. The smaller molecules can enter the pores in the beads, therefore their passage through the column is slowed down. The larger molecules cannot fit into the pores, hence they are eluted before the smaller molecules.

Gel Filtration chromatography was performed using a Pharmacia Superdex 200 gel filtration column (column volume 120 cm^3) with a flow rate of 1 ml/min. The pellet from the anion exchange was redisolved in Buffer A (1 ml), and loaded onto the gel filtration column pre-equilibrated with two column volumes of Buffer A. The protein

was eluted with 120 ml of buffer A and collected in 1 ml fractions. The homogeneity of the purified enzyme was assessed using SDS-PAGE and the active protein fractions were combined. The Pharmacia Superdex 200 Gel filtration column had been previously calibrated using protein standards. The elution profile was used to determine the approximate molecular mass of the transketolase enzymes (appendix Ia)

2.1.10.5 Handling and storage of protein solutions

To preserve the maximum activity of the enzyme, all samples were stored at 4 °C. For long-term storage the protein was subject to an 80 % ammonium sulfate precipitation and the resulting pellet was stored at 4 °C. All purification steps were carried out at 4 °C. Disposable gloves were worn whenever handling the proteins.

2.2 Characterisation

2.2.1 Colorimetric activity assay

The activity of the TK enzyme was assessed using the colormetric assay (Smith *et al.*, 2006). The reaction mixture (90 µl) containing propanal (50 mM), lithium hydroxypyruvate (50 mM), TPP (2.4 mM), $CaCl_2$ (9 mM) and TK sample (50 % total volume) in glycylglycine (50 mM, pH 7.0) was incubated at 20 ^{o}C for 17 hrs. 10 µl of the reaction mixture was then transferred to a micro-well containing MP-carbonate resin (Biotage AB) (10 mg) and glycylglycine buffer (90 µl, 50 mM, pH 7.0) and the mixture were incubated at 20 ^{o}C for 3 hrs. 50 µl of this mixture (without the beads) was then diluted with further glycylglycine buffer (50 µl, 50 mM, pH 7.0), then tetrazolium red solution (20 µl, 0.2 % 2,3,5-triphenyltetrazolium chloride in methanol) and finally 3 M NaOH (aq) (10 µl) with good mixing. A standard curve was produced (see appendix Ib) using stock solutions of (3S)-1,3-dihydroxypentan-2-one, which were made up at 50, 25, 10, 5, and 2.5 mM concentrations in glycylglycine buffer (50 mM, pH 7.0). Each solution (5 µl) was diluted with glycylglycine buffer (95 µl, 50 mM, pH 7.0) prior to the addition of tetrazolium red solution (20 µl) and finally 3 M NaOH (aq) (10 µl) with good mixing. All reactions were monitored after 2 min and OD_{485} measurements were carried out using an infinte 200 Tecan i-control.

2.2.2 Dynamic light scattering analysis

The homogeneity of the final purified enzyme was assessed. A 25 µl sample (1 mg/ml) in buffer A was placed in a Wyatt Technology cuvette (Wyatt Technology, UK) and analysed at room temperature using the Wyatt technologies DynaPro titan Dynamic Light Scattering instrument (Wyatt Technology, UK). The Dyanamics 6.7.3 software was used to view the data.

2.3 Crystallisation

2.3.1. Sample preparation

The purified TK protein was concentrated to between 10-20 mg/ml (depending on the particular experiment) in a 2 ml Vivaspin centrifugal concentrators (Vivascience, UK) with a 10,000 Da cut-off membrane in buffer A. The concentrated protein sample was then buffer exchanged into a series of crystallisation buffers (varying on the crystallisation buffers used by Littlechild et al (1995).

Buffer C 50 mM PIPES buffer pH 6.4 containing 20 mM thiamine pyrophosphate and 90 mM $CaCl_2$.

Buffer D 50 mM PIPES buffer pH 6.4 containing 10 mM thiamine pyrophosphate and 45 mM $CaCl_2$.

Buffer E 50 mM PIPES buffer pH 6.4 containing 2 mM thiamine pyrophosphate and 9 mM $CaCl_2$.

2.3.2 Crystallisation methods

Crystallisation trials were set up using sitting drop vapour diffusion and the microbatch method. All purified TK samples in crystallisation buffer were centrifuged at 14,000 g to remove any insoluble materials prior to crystallisation.

2.3.2.1 Sitting drop vapour diffusion method

A 24 reservoir well Linbro plate (17 mm diameter and 16 mm deep) was used. The type and concentration of precipitant (1 ml) was then pipetted into the reservoir well. The protein sample (9 µl) was pipetted onto the centre of the micro-bridge that had been placed in the centre of the well using forceps. An aliquot (1 µl) of the precipitant was pipetted on top of the protein droplet. The top of the well was lined with a thin layer of silicon grease, which was dispensed using a 10 cm^3 syringe. A square plastic cover slip (22 mm) was placed on the silicon grease lined reservoir well to complete a sealed system (Fig. 2.2). This was repeated on the remaining reservoir wells with a range in concentration of precipitant. The Linbro plates were then incubated at 19 ^oC and checked on a regular basis.

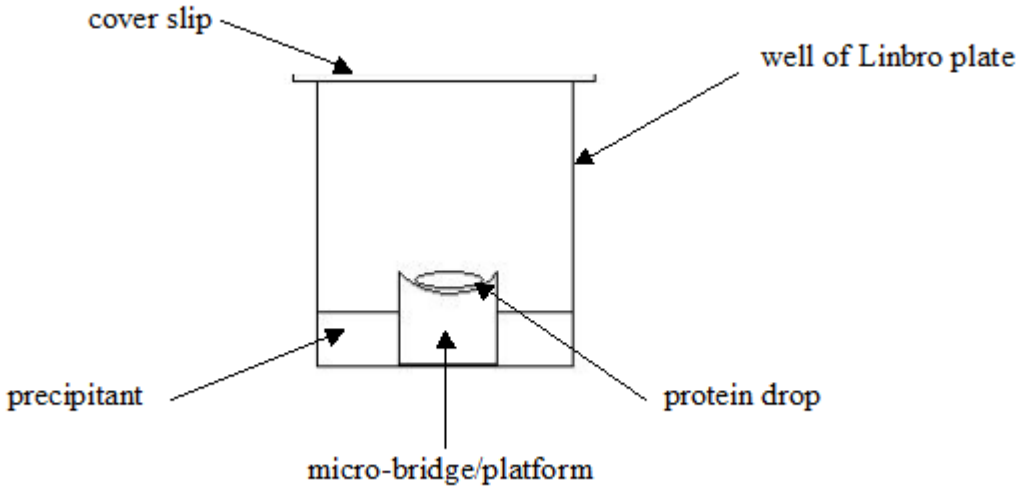


Figure 2.2 Diagram showing the set-up of a vapour diffusion crystallisation experiment.

2.3.2.2 Microbatch method

The Oryx6 system crystallisation robot (Douglas Instruments, Cambridge, UK) was used to perform all of the microbatch crystallisation experiment. The protein sample (96 µl) (using 1 µl per well) was pipetted into an eppendorf, which was placed in position on the Oryx6 system operations table. The desired screen of precipitants were pipetted into a 96 well Nunc HLA plate (Nunc, Denmark), which was placed in position on the operations table. Finally the 96 well hydrophobic microbatch crystallisation plate (Douglas Instruments, Cambridge, UK) was placed in position on the operations table. A 50/50 mix of paraffin oil (2.5 ml) and silicon oil (2.5 ml) (Al's

oil)(D'Arcy *et al.*, 2003) was made up and deposited into the vial on the operations table. The crystallisation robot was set up with the desired parameters (Douglas-Instruments-Ltd, 2008). The two-channel pipette of crystallisation robot arm X pipetted up 96 µl of protein sample into channel A. The precipitant (1 µl) was then pipetted up into channel B. The two-channel pipette then deposited the mixture (1 µl protein, 1 µl precipitant) into the well and was covered with a layer of Al's oil (5 µl) (Fig. 2.3). The process was repeated until the whole screen had been completed. A layer of Al's oil (2 ml) was pipetted on to the microbatch crystallisation plate before covering the lid. The microbatch plate was then incubated at 19 ^{o}C and checked on a regular basis.

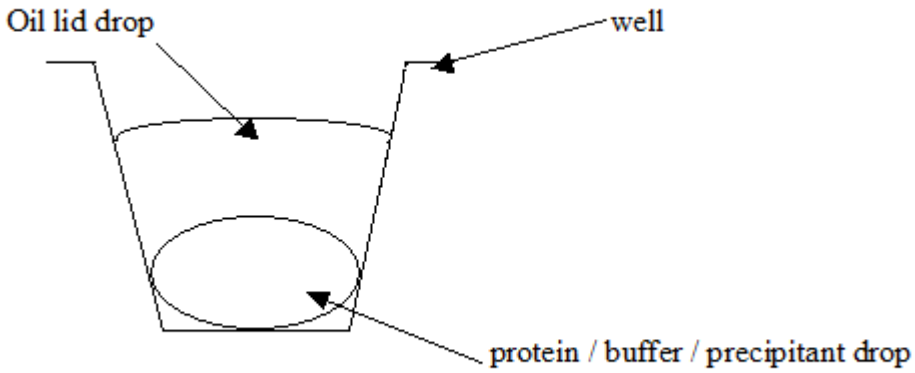


Figure 2.3 Diagram showing the set-up of a microbatch crystallisation experiment.

2.3.3 Crystallisation of TK

2.3.3.1 Crystallisation trials

2.3.3.1.1 Microbatch trials of TK

The TK samples were prepared as described in section 2.3.1 using a protein concentration of 20 mg/ml in crystallisation buffers C, D and E. The protein samples were screened with NeXtal'The pHclear™' (Qiagen, UK) (appendix) using the Oryx6 system crystallisation robot (Douglas Instruments, Cambridge, UK) and the method described in section 2.3.2.2.

2.3.3.1.2 Vapour diffusion trials of TK

The TK samples were prepared as described in section 2.3.1 using protein concentrations of 10 and 20 mg/ml in crystallisation buffers C, D and E. Sitting drop vapour diffusion crystallisation trials were set up using the conditions of Littlechild et al with the Linbro 24 well plates described in section 2.3.2.1. The reservoir solutions contained a range of ammonium sulfate concentrations (40 %-60 % saturation) in 50 mM PIPES, pH 6.4. 9 µl of protein sample was dispensed onto the micro-bridge and 1 µl of 100 % (saturation) ammonium sulfate solution was added to it. The wells were sealed and incubated as described in section 2.3.2.1. See appendix for full list of conditions.

2.3.3.1.3 Optimisation of vapour diffusion crystallisation of TK

Subsequent crystals were grown using a reservoir solution of ammonium sulfate concentrations (44 %-52 % saturation) in 50 mM PIPES, pH 6.4. The drop contained 9 µl protein (20mg/ml in Buffer E) and 1 µl 100 % (saturation) ammonium sulfate solution.

2.3.4 Co-crystallisation of TK with hydroxypyruvate (HPA)

Co-crystallisation experiments were conducted using the optimised vapour diffusion conditions used in section 2.3.3.1.3 with ranging β-hydroxypyruvic acid (Fluka, UK) concentrations (0.3 mM – 40 mM). Crystals were grown using a reservoir solution of ammonium sulfate concentrations (42 %-56 % saturation) in 50 mM PIPES, pH 6.4. The drop contained 9µl protein (20mg/ml in Buffer E, with 0.3 mM – 40 mM β-hydroxypyruvic acid, pH 6.4) and 1µl 100 % (saturation) ammonium sulfate solution.

2.3.5 Co-crystallisation of TK with fluoropyruvate (FPA)

Co-crystallisation experiments were conducted using the optimised vapour diffusion conditions used in section 2.3.3.1.3 with ranging β-fluoropyruvic acid (sodium salt)(Sigma, UK) concentrations (0.3 mM – 40 mM). Crystals were grown using a reservoir solution of ammonium sulfate concentrations (42 %-56 % saturation) in 50

mM PIPES, pH 6.4. The drop contained 9 µl protein (20 mg/ml in Buffer E, with 0.3 mM – 40 mM β-fluoropyruvic acid, pH 6.4) and 1 µl 100 % (saturation) ammonium sulfate solution.

2.3.6 Crystallisation of H26Y

The H26Y samples were prepared as described in section 2.3.1 using protein concentrations of 15 and 20 mg/ml in crystallisation buffer E. Sitting drop vapour diffusion crystallisation trials were set up with the Linbro 24 well plates described in section 2.3.2.1. The reservoir solutions contained a range of ammonium sulfate concentrations (45 %-55 % saturation) in 50 mM PIPES, pH 6.4. 9 µl of protein sample was dispensed onto the micro-bridge and 1 µl of 100 % (saturation) ammonium sulfate solution was added to it. The wells were sealed and incubated as described in section 2.3.2.1. See appendix II for full list of conditions.

2.3.7 Crystallisation of D469Y

The D469Y samples were prepared as described in section 2.3.1 using protein concentrations of 15 and 20 mg/ml in crystallisation buffer E. Sitting drop vapour diffusion crystallisation trials were set up with the Linbro 24 well plates described in section 2.3.2.1. The reservoir solutions contained a range of ammonium sulfate concentrations (45 %-55 % saturation) in 50 mM PIPES, pH 6.4. 9 µl of protein sample was dispensed onto the micro-bridge and 1 µl of 100 % (saturation) ammonium sulfate solution was added to it. The wells were sealed and incubated as described in section 2.3.2.1. See appendix II for full list of conditions.

Chapter 3

Protein Purification

3.1 Introduction

In order to gain accurate results for enzyme characterisation and for optimal success for protein crystallisation it is important that the enzyme is purified to homogeneity. To obtain a homogenous protein a variety of protein purification techniques can be utilised to enable separation of the proteins on their individual characteristic properties. Proteins can be separated on the basis of their hydrophobicity, ionic charge and size using a range of chromatography methods. It is also possible to use affinity techniques that exploit specific binding properties of the protein (such as a co-factor or His-tag). Other methods of protein purification exploit the intrinsic properties of the protein, using selective denaturation, or differential precipitation.

A purification protocol was established for the recombinant *E. coli* transketolase and the two variant proteins H26Y and D469Y.

3.2 Purification of recombinant *E. coli* transketolase.

The *E. coli* transketolase enzyme gene was over expressed in the plasmid pQr 711 in *E. coli* strain K12 bacterial JM107, and was provided by Prof. John Ward of University College London (UK), as a glycerol stock, which was stored at –80 °C until needed.

3.2.1 Materials and methods

3.2.1.1 Preparation of lysed cell extract

The *E. coli* transketolase cell paste was suspended in Buffer A. The cells were lysed by sonication using a Sanyo Soniprep 150 MSE at full power for six-30 seconds

intervals on ice, the resultant extract was centrifuged at 12500 rpm for 20 minutes to remove the cellular debris, as described in section 2.1.10.1.

3.2.1.2 Ammonium sulfate fractionation

Initial trials were carried out varying the ammonium sulfate saturation 20 %, 40 %, 60 % and 80 % to establish the precipitation point of the transketolase enzyme. Enzyme grade ammonium sulfate was added gradually to a 60 ml solution of crude transketolase enzyme extract, and a 40 % and 80 % precipitation was performed, as described in section 2.1.10.2.

3.2.1.3 Ion-exchange chromatography

The pellet (transketolase 80% ammonium sulfate) from section 3.2.1.2 was subjected to ion-exchange chromatography and was performed using a Pharmacia FFQ anion exchange column (column volume 80 cm^3) with a flow rate of 4 ml/min, as described in section 2.1.10.3.

3.2.1.4 Gel filtration chromatography

The pellet (transketolase 80 % ammonium sulfate post FFQ) from section 3.2.1.3 was subjected to gel filtration chromatography and was performed using a Pharmacia Superdex 200 gel filtration column (column volume 120 cm^3) with a flow rate of 1ml/min, as described in section 2.1.10.4.

3.2.2 Purification results for recombinant *E. coli* transketolase

3.2.2.1 Ammonium sulfate fractionation

Initial trails of ammonium sulfate fractionation showed that the transketolase began to precipitate at 40 % saturation and was completely precipitated at 80 % saturation. Although some degree of transketolase was lost prior to the 40 % saturation point, it was considered to be minimal after consultation of SDS-PAGE gel.

3.2.2.2 Ion-exchange chromatography

FFQ Sepharose anion exchange chromatography was the first purification column used in the homologous purification of the transketolase enzyme. The elution of the column produced two significant peaks, the first peak was eluted at approximately 0.1 M NaCl and the other at 0.165 M NaCl (Fig. 3.1). The latter of the two contained the transketolase enzyme, which was established by the tetrazolium red colorimetric assay (section 2.2.1) (Fig. 3.1) and SDS-PAGE gel electrophoresis showing the transketolase monomer with a molecular weight of 70 kDa (Fig. 3.5).

3.2.2.3 Gel-filtration chromatography

Gel filtration chromatography was the final purification column used in the homologous purification of the transketolase enzyme. The column produced one significant peak, eluting at approximately 69 ml with a UV_{280} absorbance of 0.840 Au. (Fig. 3.2) The peak of the transketolase enzyme was assessed for purity on an SDS-PAGE gel. Calculation of the K_{av} and the log molecular weight from the elution volume of the transketolase enzyme (see appendix I) estimates that the protein with an elution of 69 ml had a molecular mass of approximately 158 kDa, which would indicate that the transketolase protein was in a dimeric state.

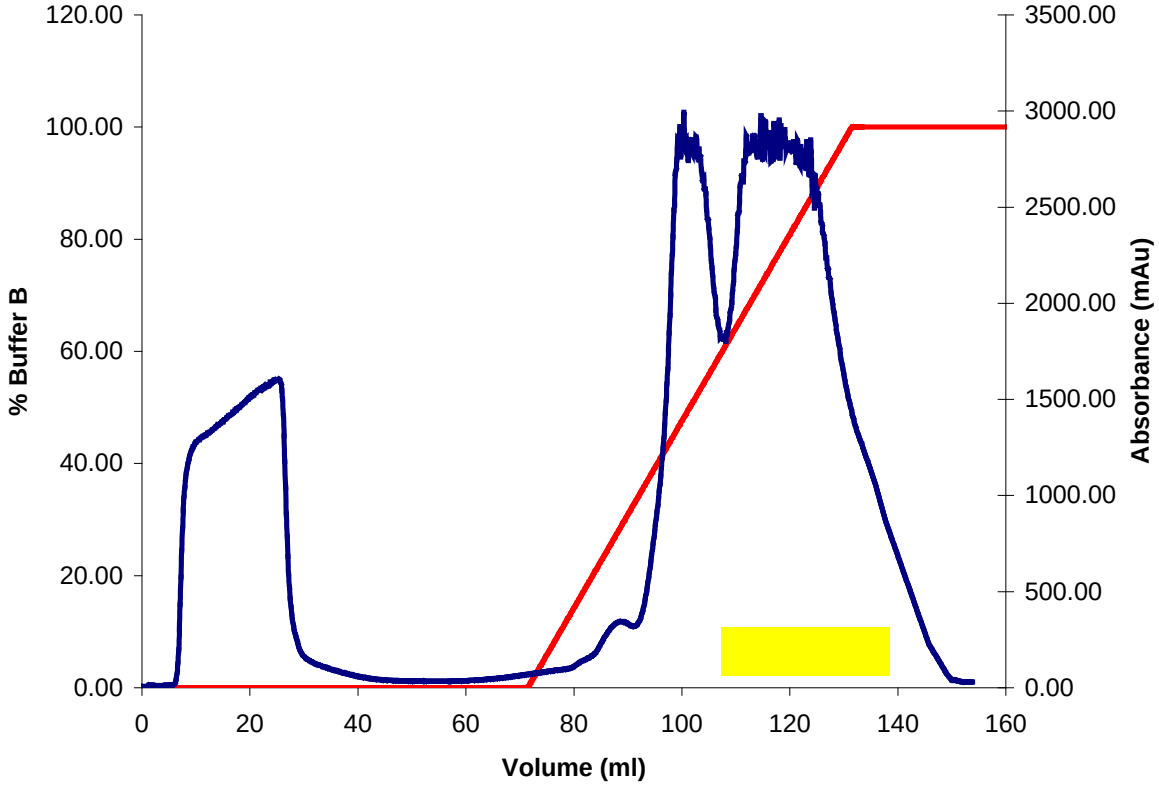


Figure 3.1 Chromatography trace of the FFQ column for TK. The absorbance at 280 nm is shown in blue, and fractions with TK activity shown in yellow. The percentage concentration of Buffer B is shown in red.

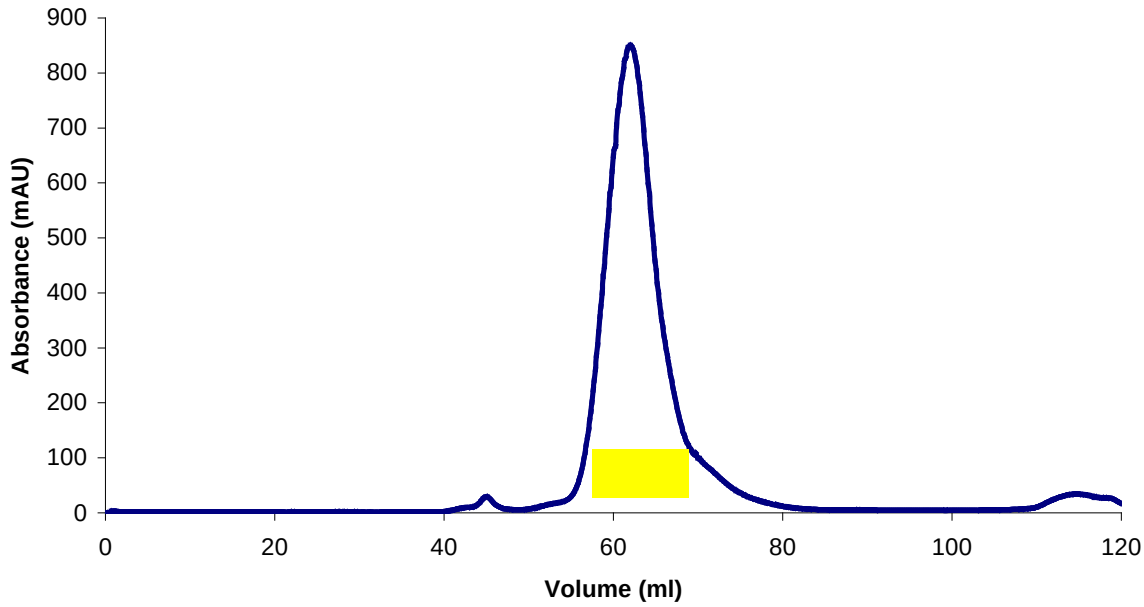


Figure 3.2 Chromatography trace of Superdex 200 column for TK. The absorbance at 280 nm is shown in blue, and fractions with TK activity shown in yellow.

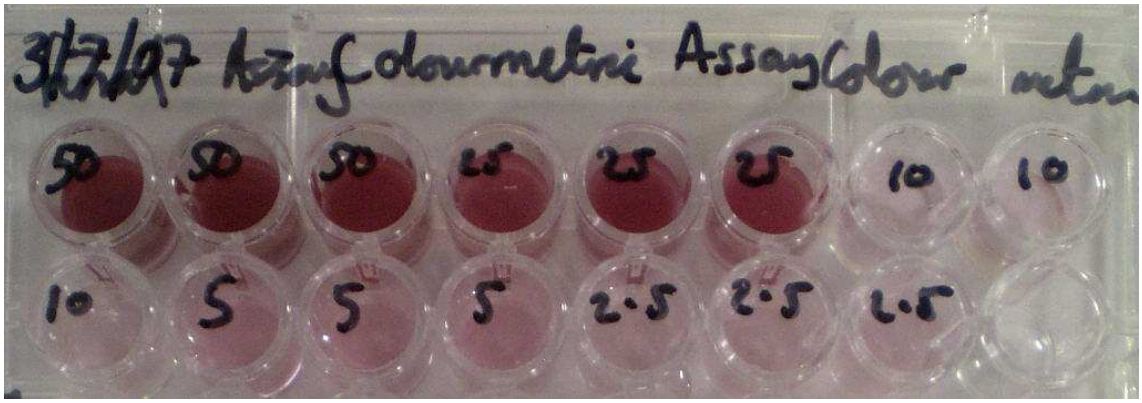


Figure 3.3 Figure showing the tetrazolium colorimetric assays standards 50 mM, 25 mM, 10 mM, 5 mM and 2.5 mM.

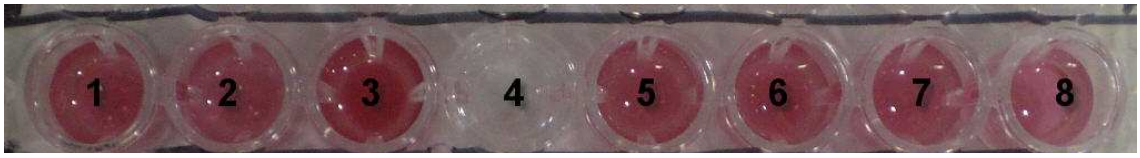


Figure 3.4 Figure showing the tetrazolium red colorimetric assays conducted during the purification of the recombinant *E. coli* transketolase. Crude (1,2), 80 % $(NH_4)_2SO_4$ (3,4), FFQ (5,6) and GF (7,8).

3.2.3.1 SDS PAGE analysis of purification

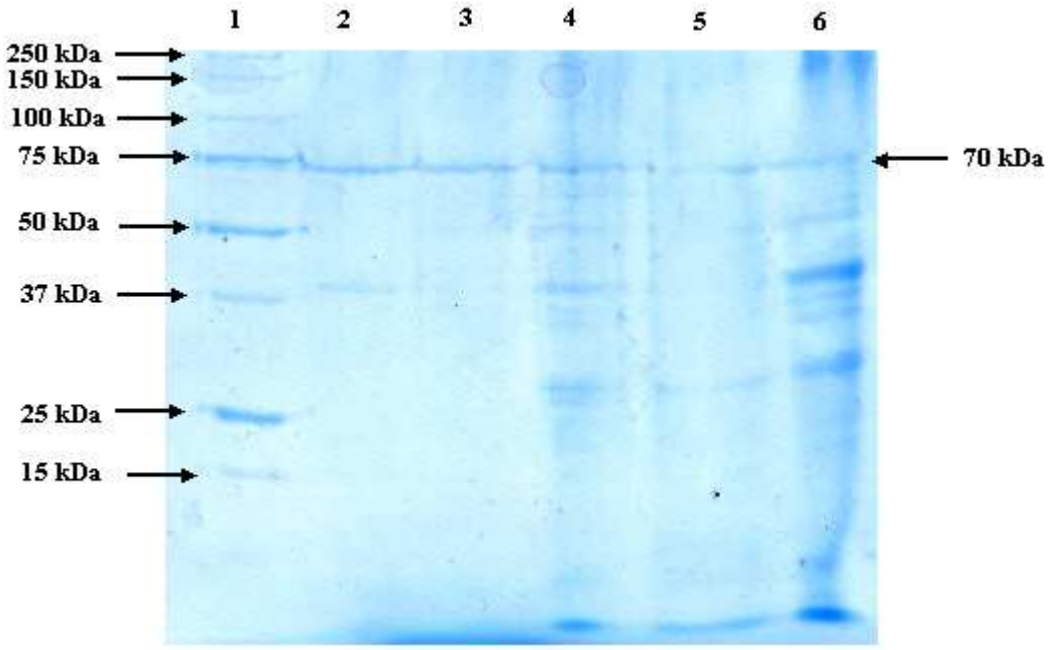


Figure 3.5 SDS PAGE gel showing the purification of the recombinant *E. coli* transketolase. Lane 1 is the molecular weight marker (Bio-Rad), lane 2 is pure TK post gel filtration, lane 3 is after 80 % ammonium sulfate fraction post FFQ, lane 4 is pooled active fractions post FFQ, lane 5 is after ammonium sulfate fractionation (40-80% saturation), and lane 6 is the crude sonicate.

3.2.4 Discussion

Purification of the recombinant *E. coli* transketolase was achieved in four stages, an initial ammonium sulfate fractionation (40 % - 80 % saturation), FFQ anion exchange chromatography, a further ammonium sulfate precipitation (80 % saturation), and gel filtration chromatography. The recombinant *E. coli* transketolase was found to contribute over 10 % of the initial protein mass of the crude extract. Analysis of the purification table (Figure 3.5) shows a drop in specific activity between the final ammonium sulfate and gel filtration chromatography step, this result could be explained by the possible displacement of the TPP co-factor by ammonium sulfate, as ammonium sulfate has been known to displace the co-factor. The molecular mass of the soluble enzyme was found to be 158 kDa, which would suggest that the quaternary structure is a dimer, as the monomer is approximately 70 kDa. The method is radius dependant, so the accuracy of this method depends on the shape of the protein. If the protein is more globular then the results will be better correlated with the actual oligomeric state of the protein. The protein was seen to behave in a stable manner throughout the entire purification process, and maintained activity after being stored in –80 °C. The pure protein was then subjected to crystallisation and characterisation experiments.

3.3 Purification of recombinant *E. coli* transketolase variant H26Y

The *E. coli* transketolase variant enzyme H26Y gene was over expressed in the plasmid pQr 711 in *E. coli* strain K12 bacterial JM107 and was provided by Prof. John Ward of University College London (UK), as a glycerol stock, which was stored at –80 °C until needed.

3.3.1 Materials and methods

The purification of the H26Y variant enzyme was conducted in the same manner as the native TK performed in section 3.2.

3.3.2 Purification results for recombinant *E. coli* transketolase variant H26Y

3.3.2.1 Ammonium sulfate fractionation

Initial trials of ammonium sulfate fractionation showed that the H26Y transketolase variant enzyme began to precipitate at 40% saturation and was completely precipitated out at 80% saturation. Although some degree of the enzyme was lost prior to the 40% saturation point, it was considered to be minimal after consultation of SDS-Page Gel.

3.3.2.2 Ion-exchange chromatography

FFQ Sepharose anion exchange chromatography was the initial purification column used in the homologous purification of the H26Y variant enzyme. The column produced two major peaks similar to that of the transketolase in section 3.2.2.2, and a smaller peak on the shoulder of the first, again the first peak was eluted at approximately at 0.075M NaCl and the broadest peak at 0.1625M NaCl (Fig. 3.6). As with the transketolase the broader peak was confirmed to be the H26Y variant enzyme using the tetrazolium red colorimetric assays (section 2.2.1) (Fig. 3.6). The SDS-PAGE gel electrophoresis determined that the fractions from the broadest peak were of higher purity than the first, again showing the H26Y monomer to have a molecular weight of 70 kDa (Fig. 3.8).

3.3.2.3 Gel-filtration chromatography

Gel filtration chromatography was the second and final purification column used in the homologous purification of the H26Y variant enzyme. The column produced one significant peak, eluting at approximately 71 ml with a UV_{280} absorbance of 2.70 Au (Fig. 3.7). The peak of the H26Y variant enzyme was assessed for purity on an SDS-PAGE gel. Calculation of the K_{av} and the log molecular weight were conducted using the same calibration figures as in section 3.2.2.3, the estimate size of the H26Y variant enzyme with an elution of 71 ml had a molecular mass of approximately 136 kDa, which again like the transketolase, indicated that the variant transketolase protein was in a dimeric state.

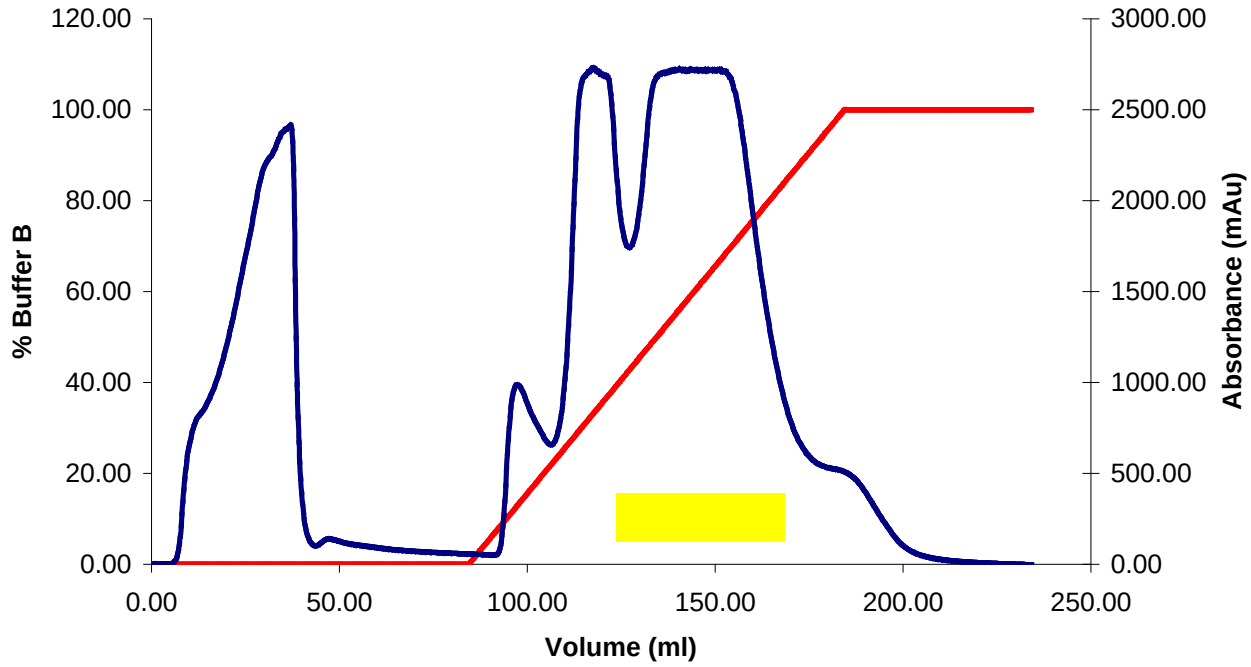


Figure 3.6 Chromatography trace of FFQ column for the TK variant enzyme H26Y. The absorbance at 280 nm is shown in blue, and fractions with TK activity shown in yellow. The percentage concentration of Buffer B is shown in red.

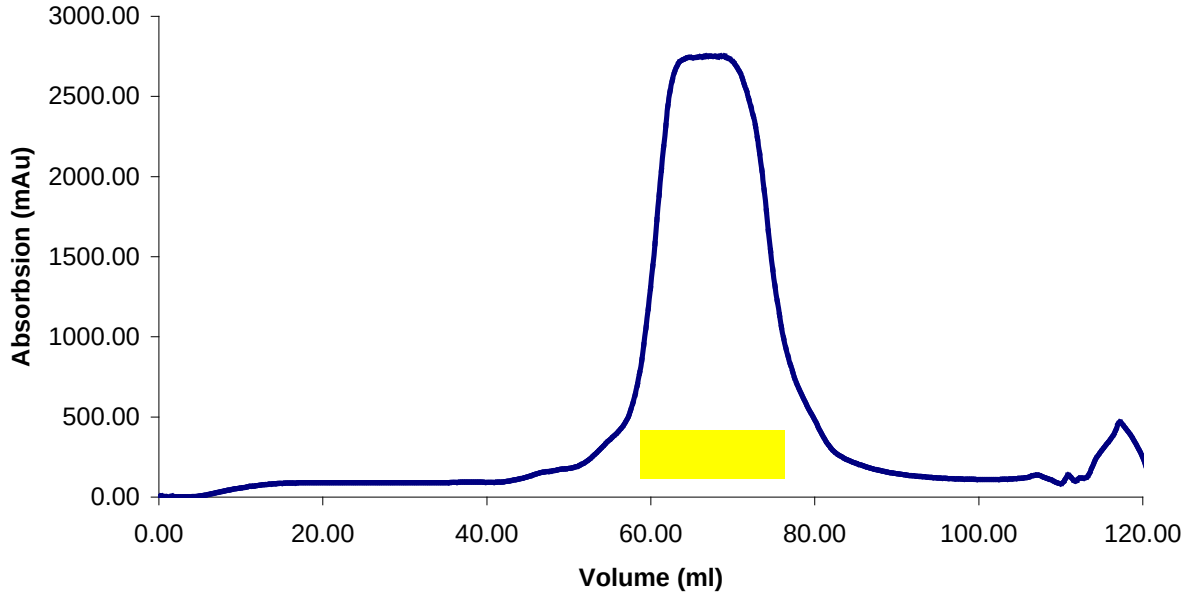


Figure 3.7 Chromatography trace of Superdex 200 column for TK H26Y variant enzyme. The absorbance at 280 nm is shown in blue, and fractions with TK activity shown in yellow.

3.3.3.1 SDS PAGE analysis of purification

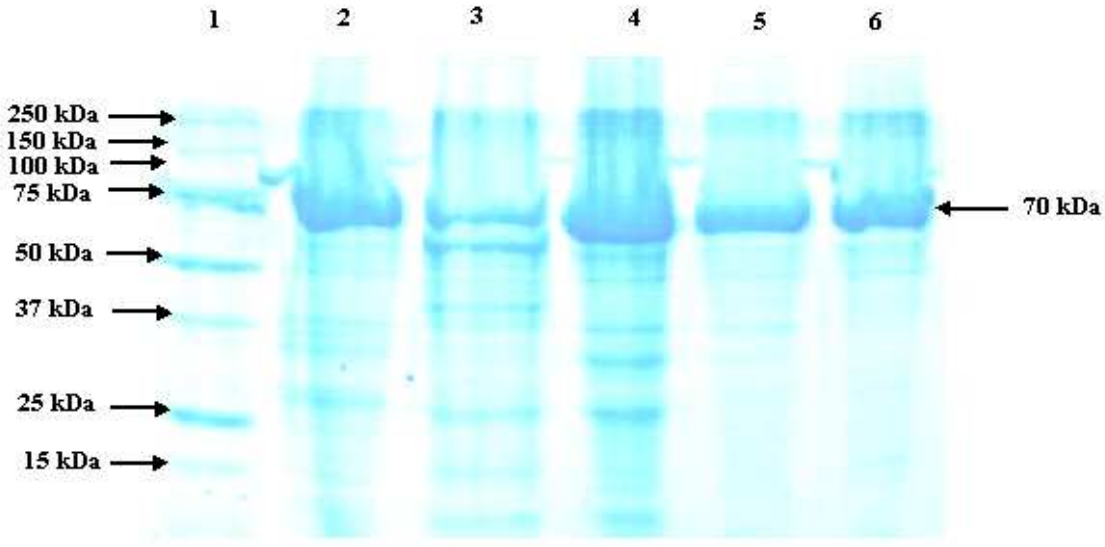


Figure 3.8 SDS PAGE gel showing the purification of the H26Y variant enzyme. Lane 1 is the molecular weight markers (Bio-Rad), lane 2 is crude sonicate, lane 3 is after ammonium sulfate fractions, lane 4 is pooled active fractions post FFQ, lane 5 is after 80% ammonium sulfate fractionation post FFQ and lane 6 is pure H26Y post gel filtration.

3.3.4 Discussion

Purification of the transketolase variant enzyme H26Y was achieved in four stages as with the recombinant TK enzyme, an initial ammonium sulfate precipitation (40 % - 80 %), FFQ anion exchange chromatography, a further ammonium sulfate fractionation (80 %), and gel filtration chromatography. The H26Y transketolase was found to contribute nearly 20 % of the initial protein mass of the crude extract. The molecular mass of the soluble enzyme was found to be 136 kDa, which would again suggest that the variant enzyme quaternary structure is the same as the TK of a dimer. The protein was seen to behave in a stable manner throughout the entire purification process, and maintained activity after a period of frozen storage at –80 °C.

The pure protein was then subjected to crystallisation and characterisation experiments.

3.4 Purification of recombinant *E. coli* transketolase variant D469Y

The *E. coli* transketolase variant enzyme D469Y gene was over expressed in the plasmid pQr 711 in *E. coli* strain K12 bacterial JM107 and was provided by Prof. John Ward of University College London (London, UK). This was provided as a glycerol stock, which was stored at -80°C until needed.

3.4.1 Materials and methods

The purification of the variant enzyme D469Y was conducted in the same manner as the native TK performed in section 3.2.

3.4.1 Purification results for recombinant *E. coli* transketolase variant D469Y

3.4.1.1 Ammonium sulfate fractionation

Initial trials of ammonium sulfate fractionation were identical to that of the H26Y variant enzyme, precipitation of the enzyme started at round 40 % saturation and was completely precipitated out at 80 % saturation. Although some degree of D469Y variant enzyme was lost prior to the 40 % saturation point, it was considered to be minimal after consultation of SDS-PAGE Gel.

3.4.1.2 Ion-exchange chromatography

FFQ Sepharose anion exchange chromatography was the first purification column used in the homologous purification of the D469Y variant enzyme. The column produced two significant peaks in an identical manner to that of the H26Y variant enzyme; the first one was eluted at approximately 0.09 M NaCl and the other at 0.1625 M NaCl (Fig. 3.9). The broader peak was confirmed to be the D469Y enzyme using the tetrazolium red colorimetric assays (section 2.2.1) (Fig. 3.9). The SDS-PAGE gel assessed the purity and confirmed the D469Y variant enzyme monomer to have a molecular weight of 70 kDa (Fig. 3.11).

3.4.1.3 Gel-filtration chromatography

Gel filtration chromatography was the second and final purification column used in the purification of the D469Y variant enzyme. The column produced one significant peak, eluting at approximately 71 ml with a UV_{280} absorbance of 2.75 Au (Fig. 3.10). The peak of the D469Y variant enzyme was assessed for purity on an SDS-PAGE gel. Calculation of the K_{av} and the log molecular weight were conducted using the same calibration figures as in 3.2.2.3, the estimate size of the D469Y variant enzyme with an elution of 71 ml had a molecular mass of approximately 136 kDa, which again like the H26Y variant enzyme, indicated that the transketolase protein was in a dimeric state.

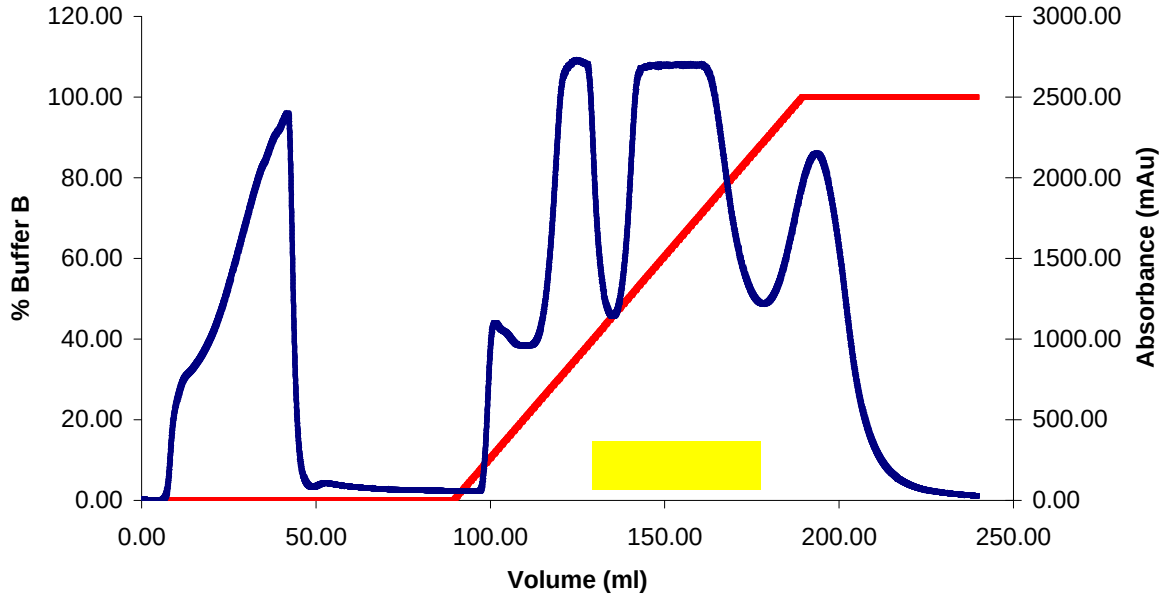


Figure 3.9 Chromatography trace of FFQ column for TK variant enzyme D469Y. The absorbance at 280 nm is shown in blue, and fractions with TK activity shown in yellow. The percentage concentration of Buffer B is shown in red.

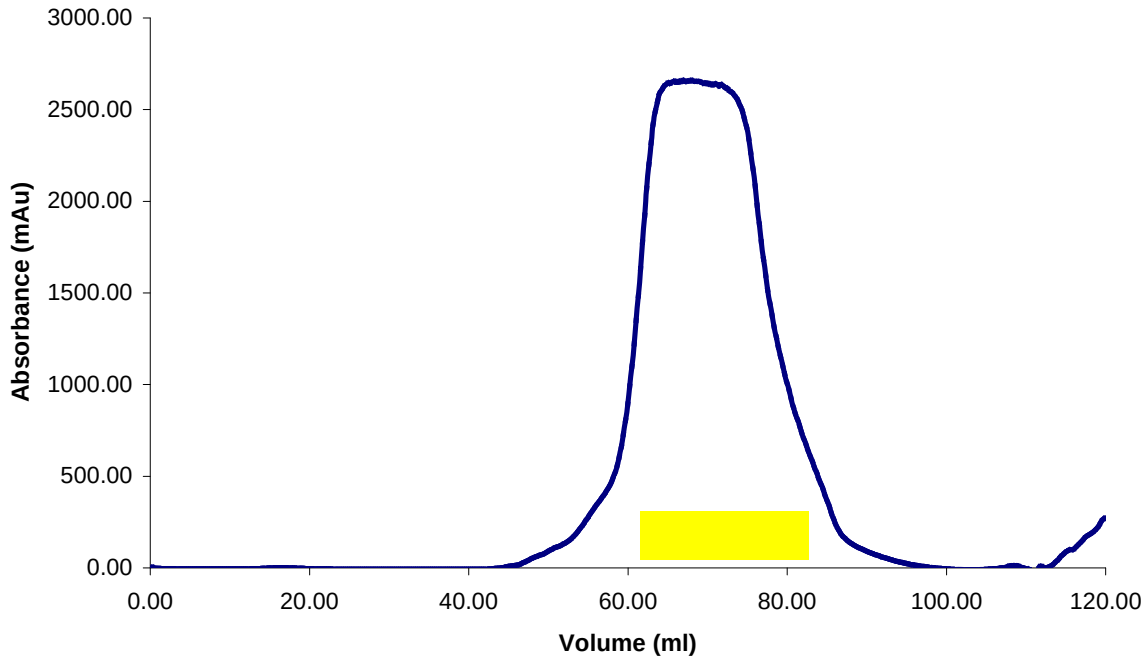


Figure 3.10 Chromatography trace of Superdex 200 column for TK D469Y variant enzyme. The absorbance at 280 nm is shown in blue, and fractions with TK activity shown in yellow.

3.4.3.1 SDS PAGE analysis of purification

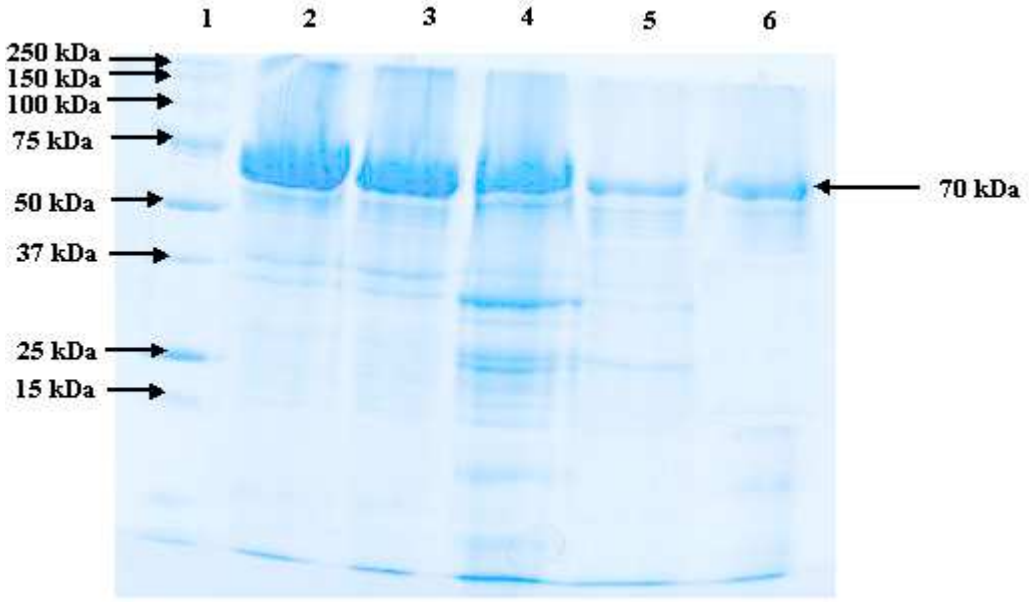


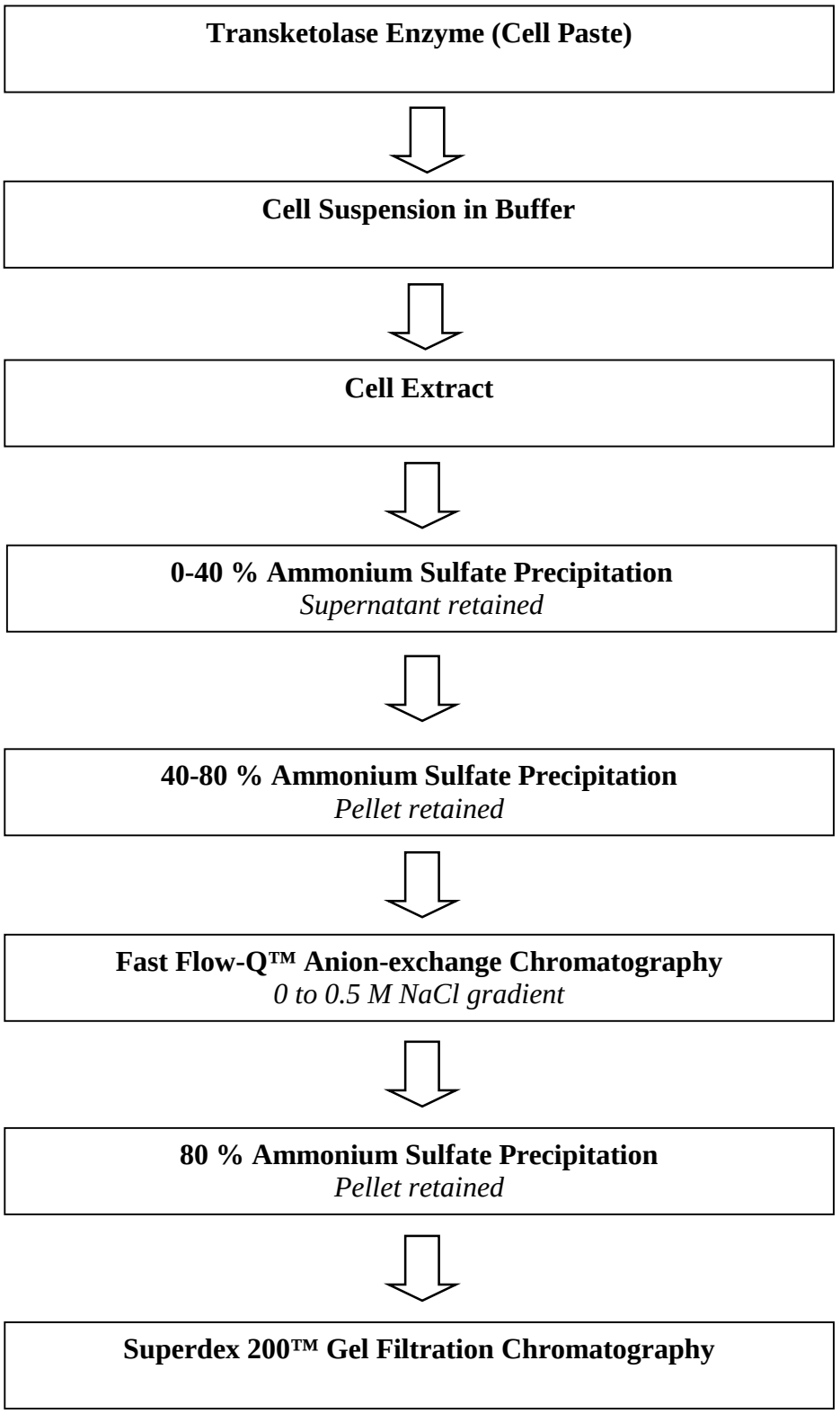
Figure 3.11 SDS PAGE gel showing the purification of the D469Y variant enzyme. Lane 1 is the molecular weight marker (Bio-Rad), lane 2 is pure D469Y post gel filtration, lane 3 is after 80 % ammonium sulfate fractionation post FFQ, lane 4 is pooled active fractions post FFQ, lane 5 is after ammonium sulfate fractions, and lane 6 is crude sonicate.

3.4.4 Discussion

Purification of the transketolase variant enzyme D469Y was achieved in four stages as with the recombinant TK enzyme, an initial ammonium sulfate fractionation (40% - 80%), FFQ anion exchange chromatography, a further ammonium sulfate fractionation (80%), and gel filtration chromatography. The D469Y variant transketolase was found not to be of homogenous purity after the final gel filtration chromatography step as with the recombinant TK and H26Y variant enzyme. Even so, the molecular mass of the soluble enzyme was found to be 136 kDa, the same as the H26Y variant enzyme, which again would suggest that the quaternary structure is a dimer. The protein was seen to behave in a stable manner throughout the entire purification process, and maintained activity after a period of frozen storage at –80 °C.

The protein from the gel filtration column was then subjected to crystallisation and characterisation experiments.

3.5 Purification protocol of *E. coli* transketolase



Chapter 4

Characterisation

The studies conducted during this chapter were aided and performed under the supervision of Dr. Paul Dalby of University College London (London, UK).

4.1 Introduction

Enzyme characterisation involves the study of the biological functions and catalytic activity of the enzyme. Examination of protein biochemistry opens up the field of enzyme characterisation, which includes the direct study of the reactivity of the enzyme, where the ability of the enzyme to perform a catalytic reaction is monitored within an enzyme assay. In essence an enzyme assay measures the production of a product or the consumption of a substrate over time and can be split into four groups; initial rate experiments are the simplest form of assay, where the substrate is in excess with the enzyme and when steady state is reached, rates are measures from the accumulation of product. The second, progressive curve experiments are used in analysis of enzyme assays, where the concentration of the substrate or product is recorded after the initial transient stage and over a long period of time approaching equilibrium. Thirdly, the most labour intensive assay group is the transient kinetic experiments, which are performed during the initial transient period as steady state is reached. The final class of assay are the relaxation experiments. This method involves monitoring a pre-equilibrated mixture (enzyme, substrate and product) return to equilibria after altering an external factor such as pH or temperature. Continuous assays can be used to follow a reaction and are measured using spectrophotometric, fluorimetric, calorimetric, chemiluminescent techniques; where-as discontinuous assays measure an end-point of a reaction using radiometric or chromatographic methods.

The transketolase enzyme is used in an industrial biocatalyst process to convert sugars to longer chain sugars with the transfer of a two-carbon subunit from the ketol donor. A common method of characterising the transketolase enzyme was to use a enzyme coupled linked spectrophotometric method developed by Heinrich *et al*., in 1972, which measures residual HPA using an NADH dependent assay. The linked assay

involves α-glycerophosphate dehydrogenase, triosephosphate isomerase, phosphoriboisomerase and D-ribulose-phosphate-3-epimerase, which could be followed by the oxidation of NADH at 340 nm (Heinrich *et al.*, 1972). A decade later Kochetov developed a similar assay that measured the rate of NAD^+ reduction in a coupled system with GAPDH (Kochetov, 1982). Two different discontinuous assays were developed at UCL (London, UK) within a ten-year period. A HPLC assay was developed in the mid 90's by Mitra and Woodley, (Mitra and Woodley, 1996), they studied the conversion of glycolaldehyde to L-erythrulose using lithium hydroxypyruvate as the ketol donor with a reverse phase column (Mitra *et al.*, 1998). Ten years later in 2006, Smith *et al*., published a highly sensitive colorimetric assay that could be used in a high-throughput manner (Smith *et al.*, 2006). Tetrazolium red oxidises the product of the TK reaction, 2-hydroxyketone to the diketone. This oxidation coincides with the reduction of the colourless tetrazolium red to the corresponding formazone creating an intense red colour (λ_{max} = 485 nm). In between the two assay developed at UCL (London, UK), a group from Université Blaise Pascal (France) developed a fluorimetric assay in 2003. The work by Sevestre *et al.,* is based on the principles that an umbelliferone product is released during the catalysed reaction (Sevestre *et al.,* 2003). The umbelliferone product was initially attached to the two-carbon subunit ketol donor by a chemoenzymatic route and is then released upon the elimination of the ketol donor that in turn fluoresces.

One of the research objectives was to monitor the transketolase reaction with the substrate analogue fluoropyruvate, to investigate the possible inhibition properties the compound may possess. Knowledge of the kinetic results of this reaction could lead to a greater insight into the catalytic mechanism that occurs during the transketolase reaction with industrial substrates. An investigation of the recombinant *E. coli* transketolase and two variant forms, H26Y and D469Y, with substrate analogs were conducted to ascertain the likely mode of inhibition. Inhibition incubation/time studies were performed on all three, however investigations of the kinetics were carried out with only the recombinant *E. coli* transketolase, as the mutants were received at the end of the project.

4.2 Materials and methods

4.2.1 Dynamic light scattering analysis

The Dynamic Light scattering analysis was conducted using TK after purification as a further test for homogeneity. Refer to section 2.2.2 in the materials and methods chapter for protocol.

4.2.2 Colorimetric activity assay

The colorimetric activity assay was conducted using TK and the two variants H26Y and D469Y after purification to ensure activity. Refer to section 2.2.1 in the materials and methods chapter for protocol.

4.2.3 Transketolase HPLC assay

A HPLC method was used to investigate the catalytic activity of TK. The conversion of glycolaldehyde to L-erythrulose by the TK enzyme was monitored. The cofactor solution (170 µL) (2 mM TPP, 9 mM $CaCl_2$, in Buffer A) was added to the purified TK sample (30 µL) and incubated for 20 mins at room temperature. The substrate solution (100 µL) (0-100 mM glycolaldehyde (GA), 0-100 mM lithium hydroxypyruvate (HPA), in Buffer A) was added to start the reaction. Samples (20 µL) were taken at regular intervals, and added to 0.1 % trifluoroacetic acid (TFA) in water (180 µL). Samples were applied to a Bio-Rad Aminex HPX-87H Reverse Phase Column, 300 x 7.8 mm, using mobile phase 0.1 % TFA in water at 60 ^{o}C and analysed using the UV detector of the HPLC at 210 nm for HPA and L/D-erythrulose.

4.2.4 Inhibition studies using fluoropyruvic acid (FPA)

A series of inhibition studies were performed on the TK enzyme and the two variants H26Y and D469Y using the substrate analogue fluoropyruvic acid. Fluoropyruvic acid acts by mimicking the substrate hydroxypyruvic acid.

4.2.4.1 Time/inhibition studies of TK with FPA

The inhibition relationship between TK and FPA was first studied using the transketolase HPLC assay method (section 4.2.3). The TK (1 mg/ml, 30 µL) samples were added to the cofactor solution (170 µL) (2 mM TPP, 9 mM $CaCl_2$, in Buffer A) and incubated for 20 mins at room temperature. The resultant solutions were incubated with and without FPA (50 mM). The resultant solutions (200 µL) were taken and added to the substrate solution (100 µL) containing GA (50 mM), HPA (50 mM), in Buffer A. Samples (20 µL) were taken at set intervals at 0, 1, 3, 5, 10, 30, 45, and 60 mins, and transferred to a solution of 0.1 % TFA in water (180 µL) and analysed using HPLC as described in section 4.2.3. All assays were repeated in triplicate. The consumption of reagent HPA with the production of product L-erythrulose was monitored and recorded and the results were analysed and used to calculate the optimum incubation time.

4.2.4.2 Time/inhibition studies of the variant H26Y with FPA

The inhibition relationship between H26Y and FPA was studied as described in section 4.2.4.1, using the TK variant H26Y (1 mg/ml, 30 µL). Results were recorded, analysed and used to calculate the optimum incubation time.

4.2.4.3 Time/inhibition studies of the variant D469Y with FPA

The inhibition relationship between D469Y and FPA was studied as described in section 4.2.4.1, using the TK variant D469Y (1 mg/ml, 30 µL). Results were recorded analysed and used to calculate the optimum incubation time.

4.2.4.4 Kinetic studies of TK with FPA

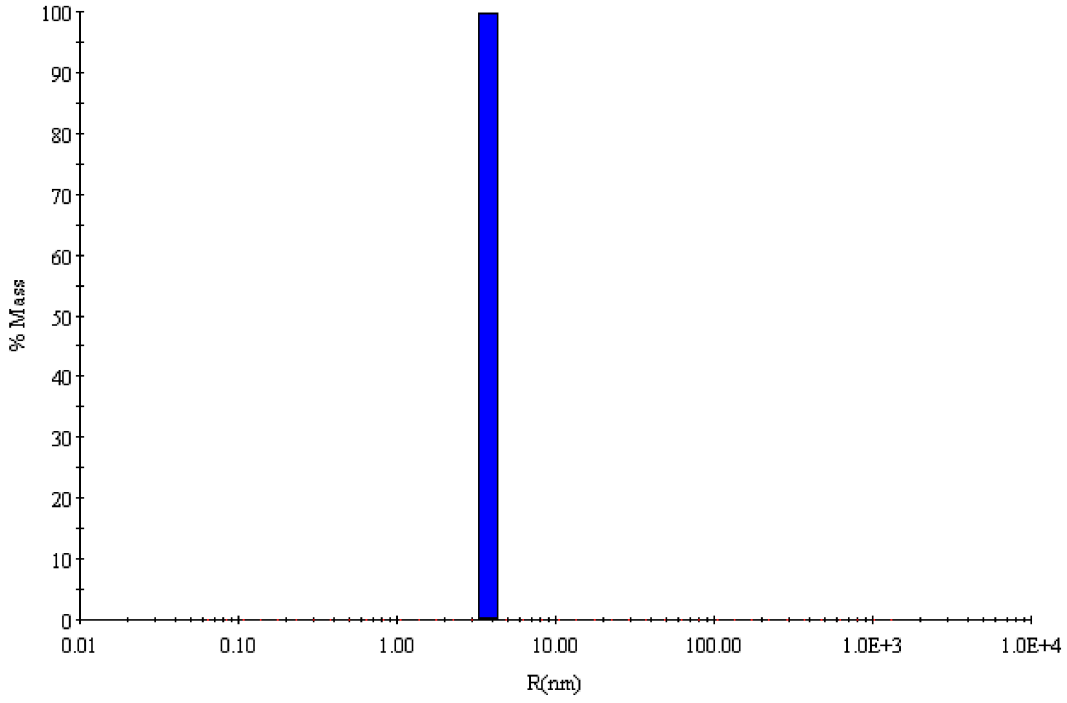
The kinetic relationship between TK and FPA was studied using the transketolase HPLC assay method (section 4.2.3). As with the time/inhibition studies, the consumption of HPA with the production of the product L-erythrulose was monitored. The approach used was to vary the concentration of either of the two substrates, HPA and GA, from 0-100 mM whilst keeping the other constant, with and without FPA. The TK (1 mg/ml, 30 µL) samples were added to the cofactor solution (170 µL) (2mM TPP, 9mM $CaCl_2$, in Buffer A) and incubated for 20 mins at room

temperature. The resultant solutions were incubated with and without FPA (50 mM). The resultant solutions (200 µL) were added to the substrate solution (100 µL) containing a set concentration of GA (100mM), with varying concentrations of HPA (5, 10, 25, 75 mM), in Buffer A. Samples (20 µL) were taken after 10, 20, 30, and 60 mins, and transferred to a solution of 0.1 % TFA in water (180 µL) and were analysed using HPLC as described in section 4.2.3. A second set of experiments was carried out simultaneously, using a set concentration of HPA (100 mM) and varying concentrations of GA (5, 10, 25, 75 mM). Results were analysed and used to calculate the kinetic relationship between TK and the inhibitor FPA.

4.3 Results and discussion

4.3.1 Dynamic light scattering analysis

The analysed *E. coli* transketolase sample showed a monomodal distribution with no polydispersity (Fig. 4.1). This implies that the *E. coli* transketolase sample is homogeneous. The other data sourced from the analysis shows the *E. coli* transketolase to have an estimated size of 70 kDa, which would imply the enzyme is in its monomer form. However, the size estimation capabilities of this apparatus are debatable due to the methods used in the estimation process. The data does reveal the protein to be homogeneous, with no aggregation that could cause problems crystallisation studies.



R (nm)	Pd (nm)	%Pd	MW-R (kDa)	%Int	%Mass
3.8	0.0	0.0	76.210	46.9	99.7

Figure. 4.1 Graph produced from dynamic light scattering for the *E. coli* transketolase sample. The graph plots the percentage mass against the hydrodynamic radius (nm) of the sample.

4.3.2.1 Colorimetric activity assay of the *E. coli* transketolase, and the variants H26Y and D469Y.

Investigations into *E. coli* transketolase using the colorimetric activity assay designed by (Smith *et al.*, 2006) yielded acceptable results. The initial calibration curve was created using ranging concentrations of (3S)-1,3-dihydroxypentan-2-one (Fig. 4.2), and the assay results were very similar to the published work by Smith *et al.*, observing the greatest amount of red precipitate formed from the formazane production in the 50 mM sample of (3S)-1,3-dihydroxypentan-2-one. The calibration curve was used to calculate the percentage conversion of substrate by the recombinant *E. coli* transketolase and the two variant forms H26Y and D469Y.

Figure. 4.2 The Tetrazolium red colorimetric assay showing the formation of the formazane complex for the three samples. TK (9), H26Y (10) and D469Y (11).

Using the data provided from the colorimetric assay calibration curve, the percentage conversion of the lithium hydroxypyruvate substrate for the three transketolases samples was calculated (Table 4.1).

Sample	Percentage conversion (%)
Recombinant transketolase	72
H26Y variant	51
D469Y variant	56

Table. 4.1 The percentage conversion of hydroxypyruvate by the three transketolase samples after 17hrs incubation. Monitored using the tetrazolium red colorimetric activity assay (Smith *et al.*, 2006).

The tetrazolium red colorimetric assay has allowed characterisation of the recombinant *E. coli* transketolase and the two variant forms H26Y and D469Y, by confirming their transketolase activity and the ability to carry out the transfer of a two-carbon subunit from one substrate to another. The recombinant transketolase produced the most efficient conversion (72 %) compared to the two variants H26Y and D469Y, which have a similar conversion percentage of 51 % and 56 % respectively.

4.3.4 Inhibition studies using fluoropyruvic acid (FPA)

The effect of fluoropyruvic acid with the recombinant *E. coli* transketolase, and the two variants H26Y and D469Y were investigated to decipher the possible inhibition properties that FPA may possess.

4.3.4.1 Time/inhibition studies of TK with FPA

The HPLC analysis recorded the amount of the two substrates, GA and HPA and the product L-erythrulose. The two substrates eluted at 6.5 and 8.5 minutes respectively, whilst the product was eluted after 12 minutes (Fig. 4.3). To gauge the possible inhibition effect of FPA on the transketolase reaction the production of product L-erythrulose was monitored. It was found that the initial rate of production of L-erythrulose decreased with increasing incubation time of the TK sample with FPA (Fig. 4.4), from an initial rate of 0.039 mMs^{-1} from a TK sample with no FPA incubation to an initial rate of 0.021 mMs^{-1} from a TK sample after 60 mins incubation with FPA. It can be seen in figure 4.4 that the rate of production of L-erythrulose begins to subside after approximately 30 minutes. It may also be noted that the actual concentration of L-erythrulose after the experiment had concluded was considerably lower for the TK sample after 60 mins incubation with FPA than the TK sample with no incubation with FPA, with concentrations of 10 mM and 16 mM respectively.

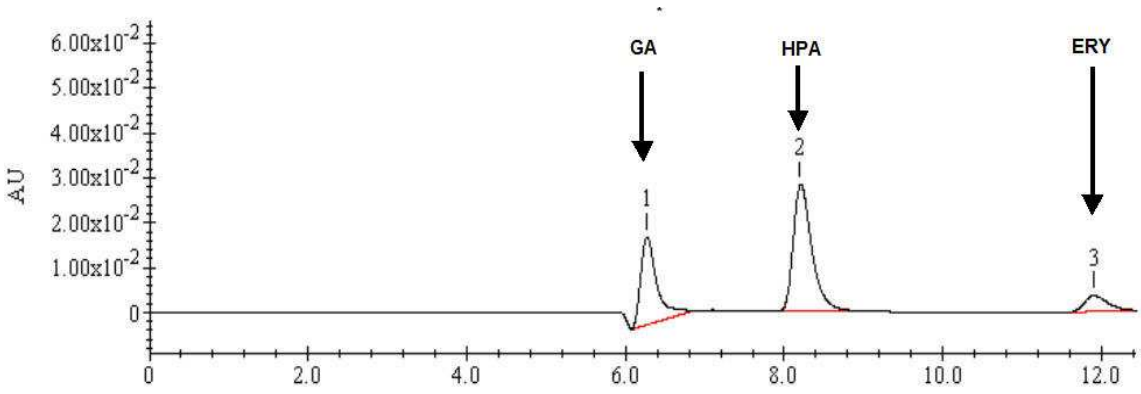


Figure. 4.3 Standard HPLC trace, showing the production of L-erythrulose from substrates glycolaldehyde and hydroxypyruvate for the TK reaction.

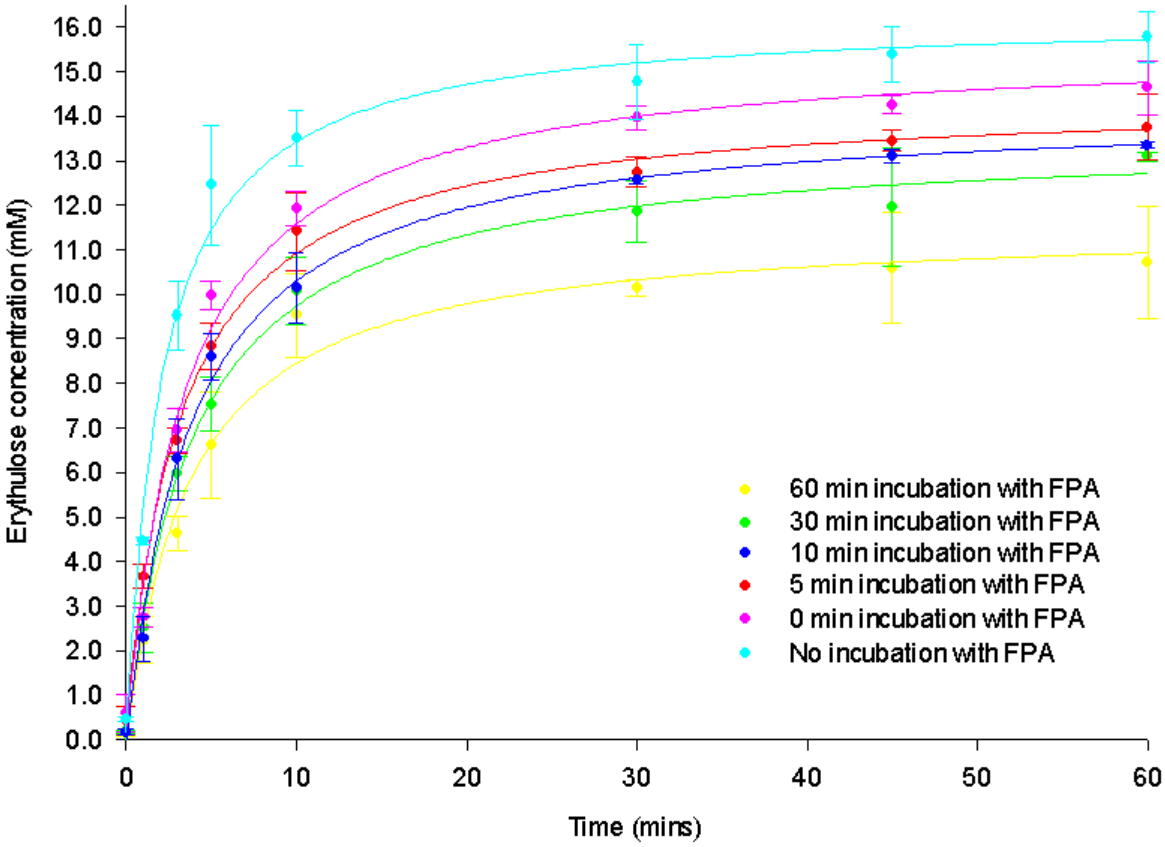


Figure. 4.4 Graph showing the production of L-erythrulose from TK samples, which have been pre-incubated in FPA for 0, 5, 10, 30 and 60 minutes. The error bars show plus and minus one standard deviation of the mean value.

4.3.4.2 Time/Inhibition studies of the variant H26Y with FPA

The variant H26Y was subjected to time/inhibition studies with FPA in the same manner as in section 4.3.4.1. The H26Y variant was of particular interest, due to the effect that the replacement of the histidine with the larger tyrosine residue would have on the rate of the transketolase reaction, as the histidine residue at position 26 is in a very prominent location within the active site.

The H26Y variant showed much lower production of erythrulose compared to that of the recombinant TK (Fig. 4.5). However, as with the recombinant TK the initial rate of production of erythrulose decreased with increasing incubation time of FPA. The initial rate fell from 0.0025 mM s^{-1} to 0.0014 mM s^{-1} for the H26Y samples with no incubation, and 60 mins FPA incubation respectively. Further evidence for the inhibitory properties of FPA can be seen in the final concentration of D-erythrulose after the 60 minutes reaction. With the final D-erythrulose concentrations of un-incubated and 60 min FPA incubated H26Y sample being 9 mM and 5 mM respectively.

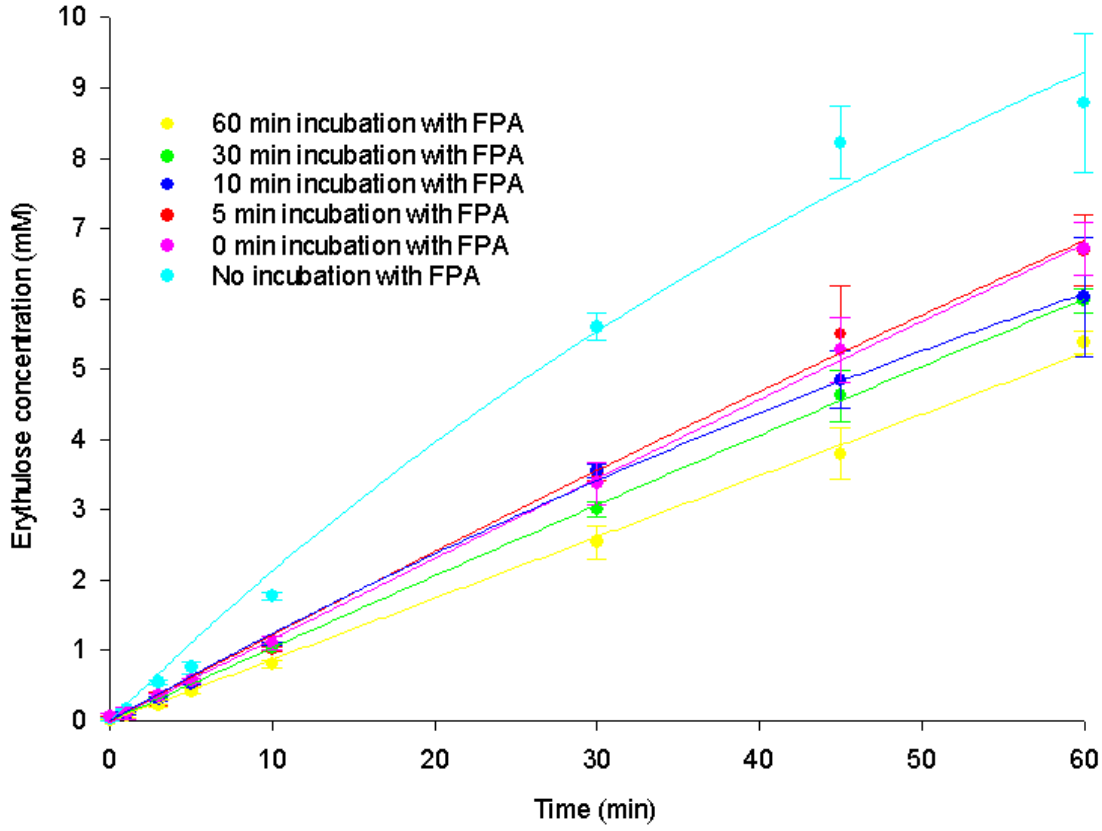


Figure. 4.5 Graph showing the production of D-erythrulose from the H26Y variant samples, which have been pre-incubated in FPA for 0, 5, 10, 30 and 60 minutes. The error bars show plus and minus one standard deviation of the mean value.

4.3.4.3 Time/Inhibition studies of the variant D469Y with FPA

The variant D469Y was subjected to time/inhibition studies with FPA in the same manner as conducted on the H26Y variant in section 4.3.4.2. In the case of the D469Y variant the substitution of the glutamate residue with the tyrosine residue occurred on the opposite side of the active site to the mutation made in the H26Y variant.

As with the recombinant TK, the data for the D469Y showed a decrease in the initial rate of production of D-erythrulose with increasing FPA incubation time (Fig. 4.6). There was an overall decrease in the initial rate from 0.0331 mM s^{-1} for the un-incubated D469Y sample, to 0.0091 mM s^{-1} for the D469Y sample after 60 minutes FPA incubation. After the 60-minute reaction, the concentrations of D-erythrulose for the un-incubated and 60 minutes incubated FPA-D469Y samples were 15.0 mM and 13.7 mM respectively.

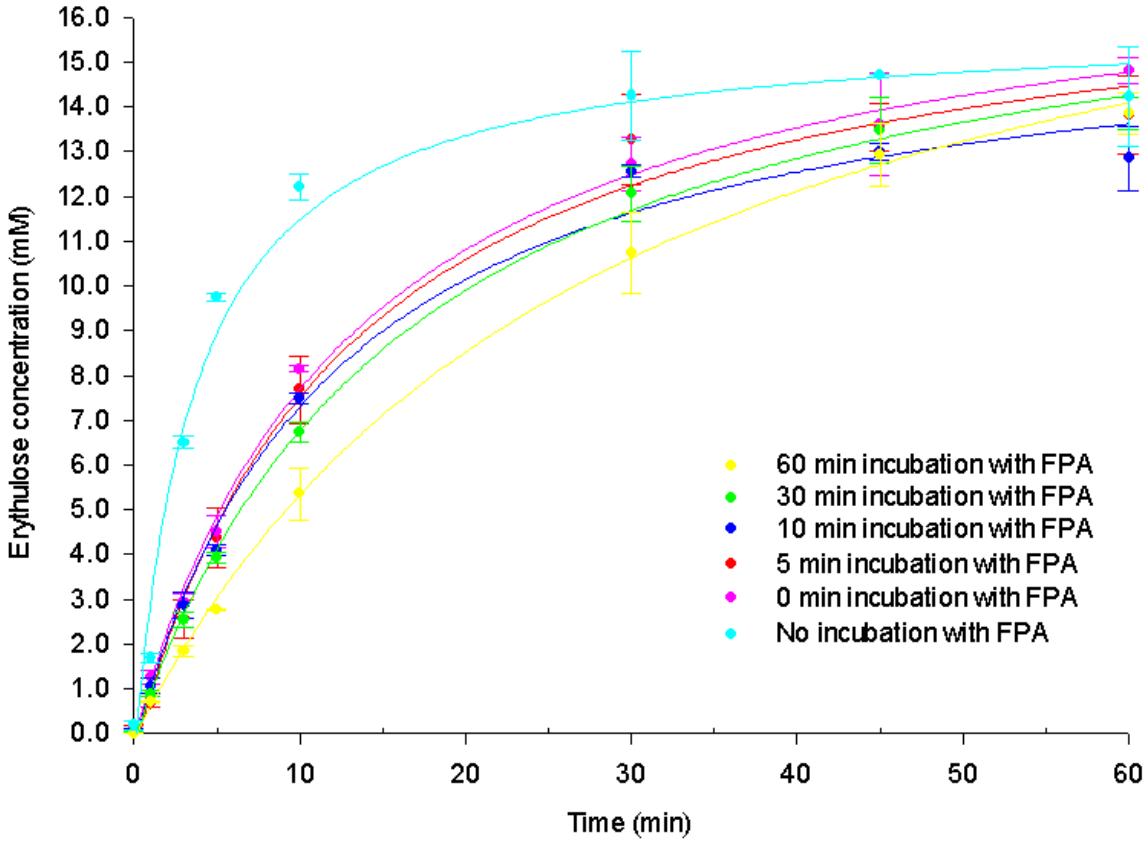


Figure. 4.6 Graph showing the production of **D-erythrulose** from variant D469Y samples, which have been pre-incubated in FPA for 0, 5, 10, 30 and 60 minutes. The error bars show plus and minus one standard deviation of the mean value.

Investigation of the possible inhibition of FPA on the TK reaction scheme proved evident in both the recombinant TK and both variant forms. It is expected that if an analogue of a substrate is introduced into a system then a form of competition will occur between the two substrates. In the case of the addition of FPA to the TK reaction, the competition between the two substrates came to a point where the FPA was seen to inhibit the conversion of HPA to the product L-erythrulose. The result from the recombinant TK studies show that samples that have been incubated with FPA exhibit a significant decrease in the rate of production and final concentration of product, than with samples that have not been incubated, as can be seen clearly in figure 4.4. Studies with the two variants showed more evidence of inhibition by FPA, with both variants displaying a reduction in the initial rate and final concentration of D-erythrulose. The H26Y variant, showed a lesser initial rate of erythrulose compared to that of recombinant TK and D469Y variant. The H26Y variant data indicates that the insertion of the tyrosine residue within this area of the active site decrease the physical size of the channel. This could change the specificity of the active site for either HPA or FPA, and hence the inhibition properties of the FPA. However, this data does not answer the question as to the mechanism in which the FPA inhibits the TK reaction.

4.3.5 Kinetic studies of TK with FPA

HPLC analysis was used to determine the mode of inhibition that the FPA exhibited upon the TK enzyme. In order to establish the rate equation for the transketolase mechanism, analysis of the initial rate must be conducted and fitted against established model kinetic data.

The methodology used was to fix the concentration of one substrate whilst varying the other, with and without the set concentration of inhibitor (and vice-versa). In this case the GA (100 mM) concentration was fixed, whilst varying HPA concentration (5, 10, 25, 50, and 75 mM), with and without FPA (50 mM). Then visa-versa, the HPA (100 mM) concentration was fixed, whilst varying GA concentration (5, 10, 25, 50, and 75 mM), with and without FPA (50 mM).

4.3.5.1 Constant GA, varying HPA, no FPA

The data collected was plotted in a concentration-time graph to obtain the rate of product formation (Fig. 4.7).

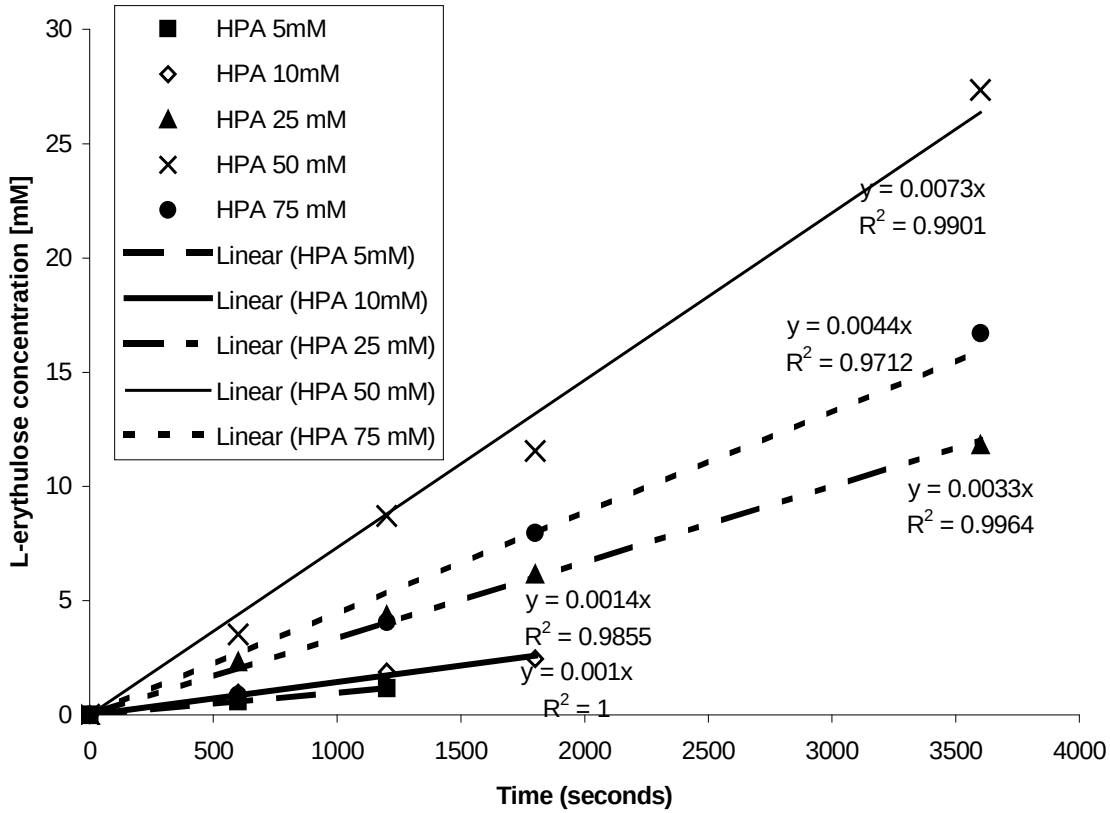


Figure. 4.7 Graph plotting L-erythrulose concentration against time for constant GA, varying HPA, no FPA.

It may be observed that the greatest concentration of HPA (75 mM) (Fig. 4.7, Table. 4.2) had started to inhibit the formation of L-erythrulose. Hence during the calculation of K_m and V_{max} this result was excluded.

HPA concentration [mM]	Initial rate Vo (mmoles/L s)	1/HPA [mM^{-1}]	1/Vo (mmoles/L s)$^{-1}$
5	0.001	0.2	1000
10	0.0014	0.1	714.28
25	0.0033	0.04	303.03
50	0.0073	0.02	136.98
75	0.0044	0.013	227.27

Table. 4.2 The data for the constant GA, varying HPA, no FPA experiment.

A double-reciprocal plot (1/Vo against 1/HPA) shown in figure 4.8 was prepared from the above data (excluding 75 mM) to calculate the K_m and V_{max}.

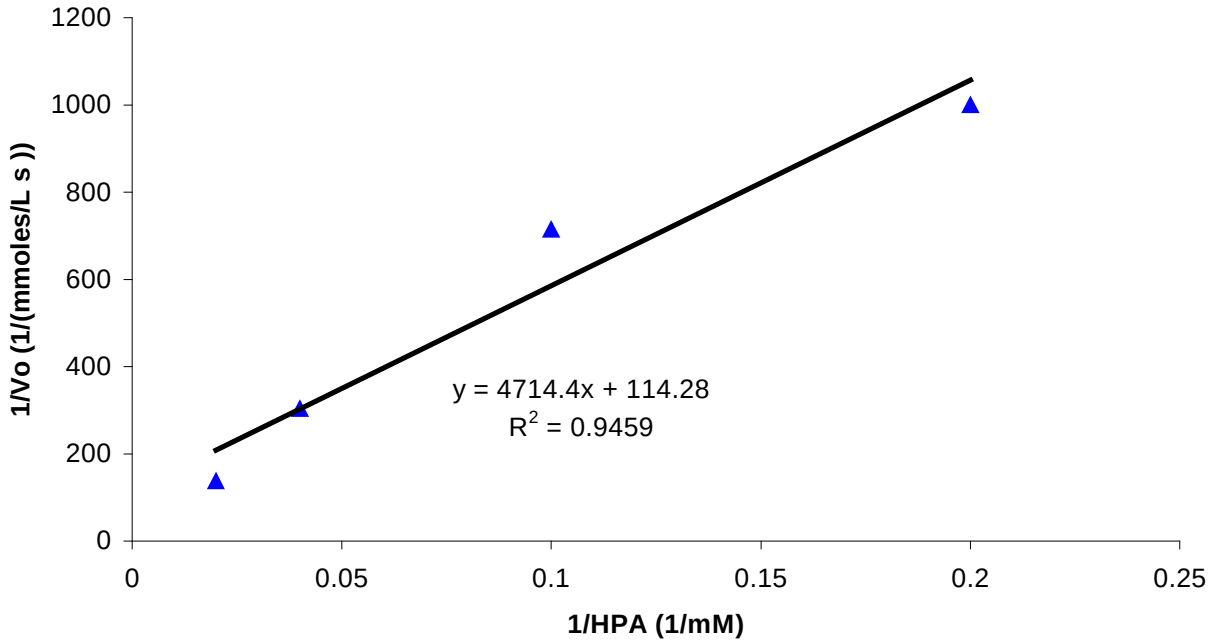


Figure. 4.8 Graph plotting 1/Vo against 1/HPA for constant GA, varying HPA, no FPA.

From figure 4.8 the V_{max} can be calculated from the inverse of the intercept to be 0.00875 mM/s. The K_m calculated from the gradient of the line multiplied by the V_{max} to be 41.25 mM for constant GA, varying HPA, no FPA.

4.3.5.2 Constant GA, varying HPA, with FPA

The data collected was plotted in a concentration-time graph to obtain the rate of product formation (FPA concentration 50 mM) (Fig. 4.9).

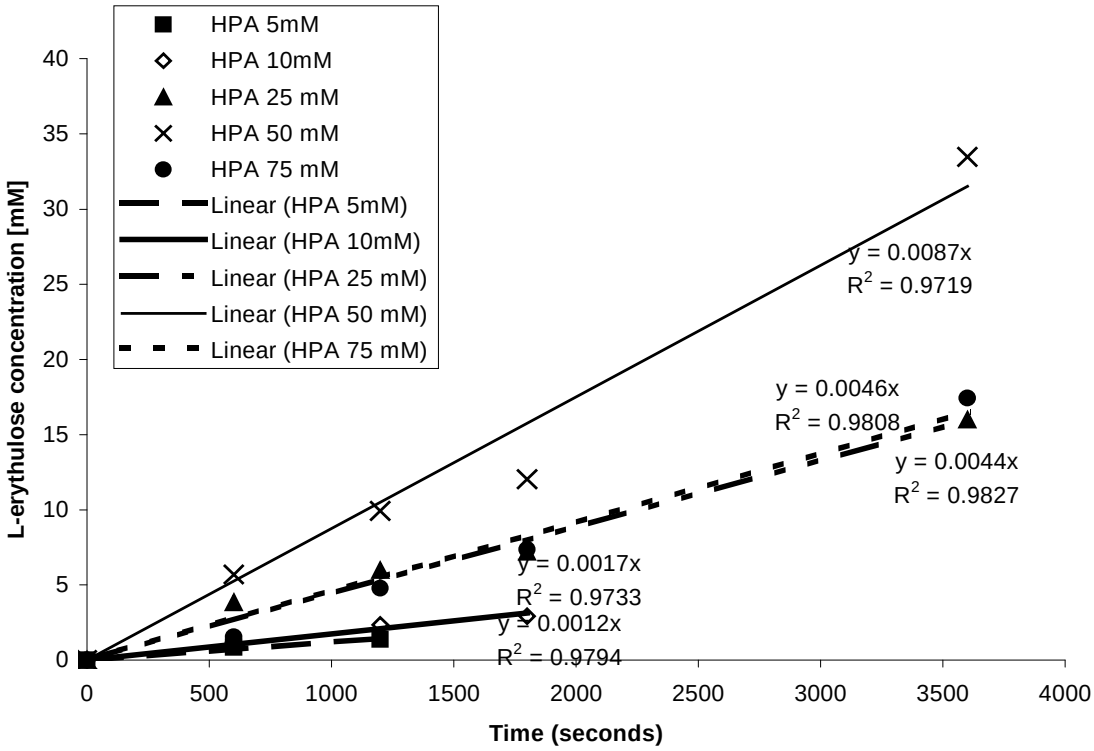


Figure. 4.9 Graph plotting L-erythrulose concentration against time for constant GA, varying HPA, with FPA.

It may be again noted that the greatest concentration of HPA (75 mM) (Fig. 4.9, Table. 4.3) had started to inhibit the formation of L-erythrulose. Hence during calculation of K_m and V_{max} the result was excluded.

HPA concentration [mM]	Initial rate Vo [mmoles/L s]	1/HPA [mM^{-1}]	1/Vo $(mmoles/L\ s)^{-1}$
5	0.0012	0.2	833.33
10	0.0017	0.1	588.23
25	0.0044	0.04	227.27
50	0.0087	0.02	114.94
75	0.0046	0.013	217.39

Table. 4.3 The data for the constant GA, varying HPA, with FPA experiment.

A double-reciprocal plot (1/Vo against 1/HPA) shown in figure 4.10 was prepared from the above data (excluding 75 mM) to calculate the K_m and V_{max}.

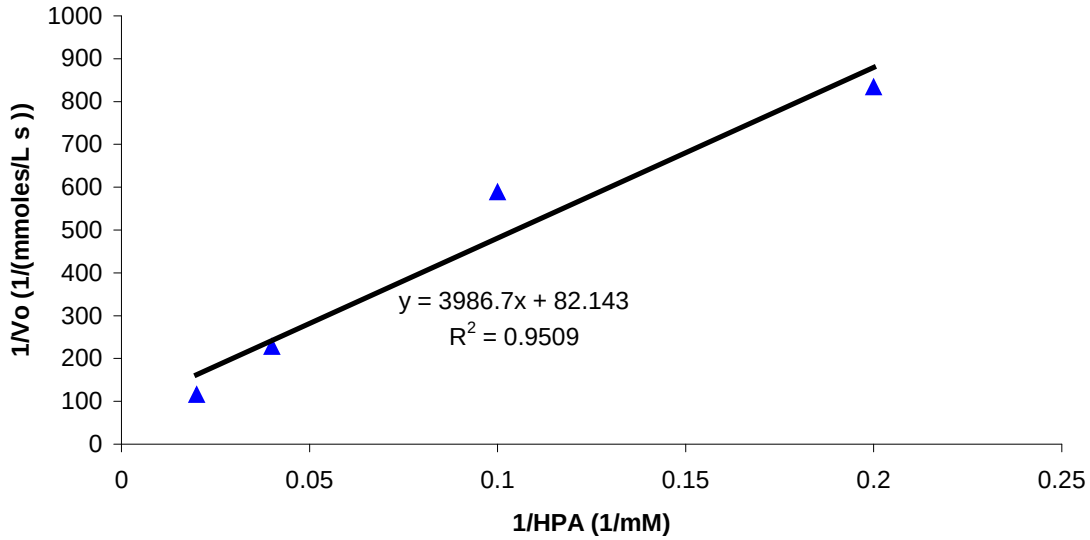


Figure. 4.10 Graph plotting 1/Vo against 1/HPA for constant GA, varying HPA, with FPA.

From figure 4.10 the V_{max} can be calculated from the inverse of the intercept to be 0.012 mM/s. The K_m calculated from the gradient of the line multiplied by the V_{max} to be 48.53 mM for constant GA, varying HPA, with FPA.

For the calculation of K_i for FPA against HPA, we rearrange the Michaelis-Menten equation for competitive inhibition to make K_i the subject (Copeland, 2000).

$$\mathbf{K_{m\ (with\ inhibitor)} = K_{m\ (without\ inhibitor)} \cdot (1+[inhibitor] / K_{inhibitor})}$$

Therefore rearrangement making K_i the subject.

$$\mathbf{K_i = [FPA] / ((K_{m\ (with\ FPA)} / K_{m\ (without\ FPA)}) - 1)}$$

$$\mathbf{K_i = 50 / ((48.53 / 41.25) - 1))}$$

$$\mathbf{K_i = 283.3\ mM}$$

4.3.5.3 Constant HPA, varying GA, no FPA

The data collected was plotted in a concentration-time graph to obtain the rate of product formation (Fig. 4.11).

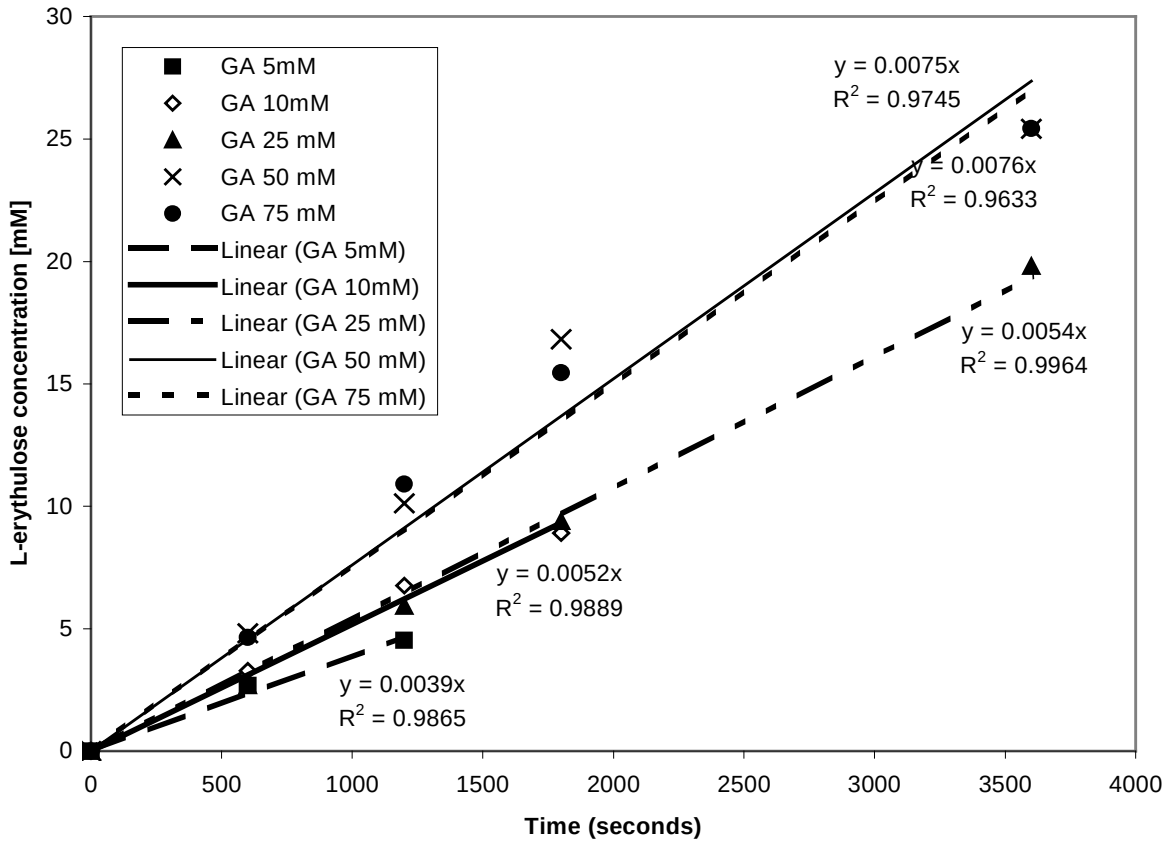


Figure. 4.11 Graph plotting L-erythrulose concentration against time for constant HPA, varying GA, no FPA.

In order to obtain some degree of error, a sensitivity analysis was performed by single point removals (Table 4.4) during the calculation of K_m and V_{max}.

GA concentration [mM]	Initial rate Vo [mmoles/L s]	1/GA [mM^{-1}]	1/Vo (mmoles/L s)$^{-1}$	Sensitivity analysis (single point removal)	
5	0.0039	0.2	256.41	0.00764	5.259
10	0.0052	0.1	192.30	0.00748	4.723
25	0.0054	0.04	185.18	0.00820	5.558
50	0.0076	0.02	131.57	0.00711	4.124
75	0.0075	0.013	133.33	0.00725	4.315

Table. 4.4 The data for the constant HPA, varying GA, no FPA experiment.

A double-reciprocal plot (1/Vo against 1/GA) shown in figure 4.12 was prepared from the above data to calculate the K_m and V_{max}.

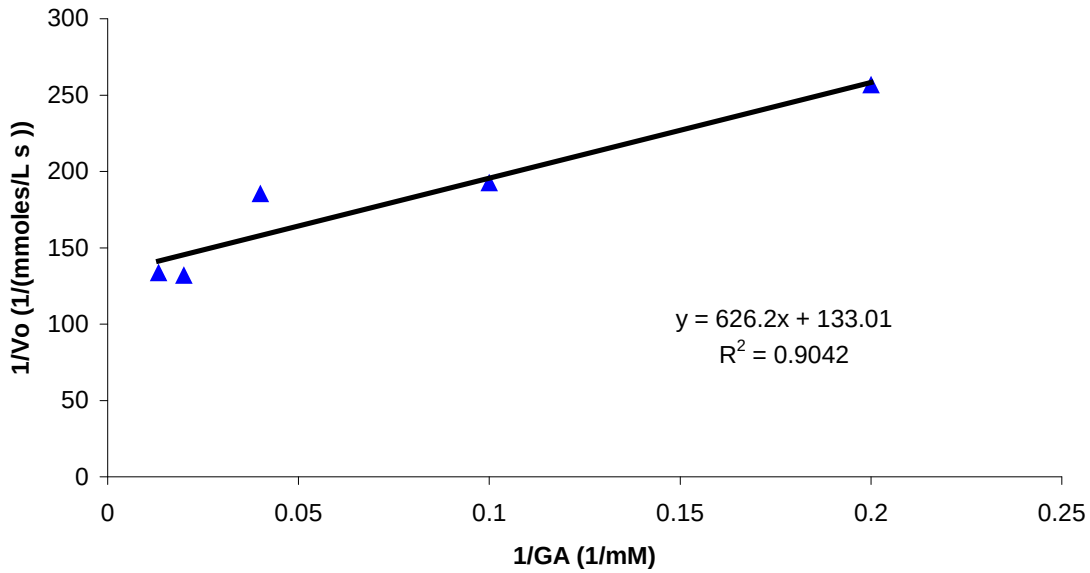


Figure. 4.12 Graph plotting 1/Vo against 1/GA for constant HPA, varying GA, no FPA.

From figure 4.12 the V_{max} can be calculated from the inverse of the intercept to be 0.0075 mM/s. The K_m calculated from the gradient of the line multiplied by the V_{max} to be 4.70 mM for Constant HPA, varying GA, no FPA.

4.3.5.4 Constant HPA, varying GA, with FPA

The data collected was plotted in a concentration-time graph to obtain the rate of product formation (Fig. 4.13).

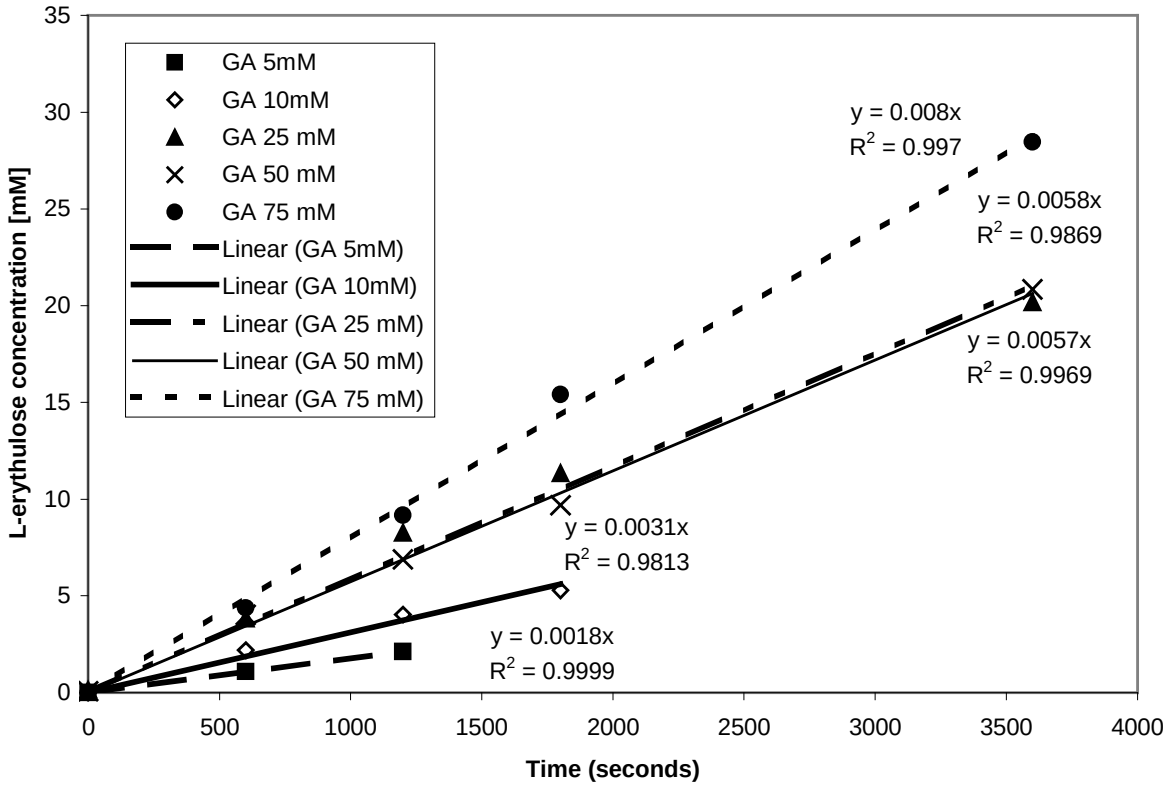


Figure. 4.13 Graph plotting L-erythrulose concentration against time for constant HPA, varying GA, with FPA.

In order to obtain some degree of error, a sensitivity analysis was performed by single point removals (Table 4.5) during the calculation of K_m and V_{ma}

GA concentration [mM]	Initial rate Vo [mmoles/L s]	1/GA [mM^{-1}]	1/Vo (mmoles/L s)$^{-1}$	Sensitivity analysis (single point removal)	
5	0.0018	0.2	555.55	0.00932	19.695
10	0.0031	0.1	322.58	0.00968	21.804
25	0.0058	0.04	172.41	0.00903	19.942
50	0.0057	0.02	175.43	0.01145	26.775
75	0.008	0.013	125.0	0.00934	20.701

Table. 4.5 The data for the constant HPA, varying GA, with FPA experiment.

A double-reciprocal plot (1/Vo against 1/GA) in figure 4.14 was prepared from the above data to calculate the K_m and V_{max}.

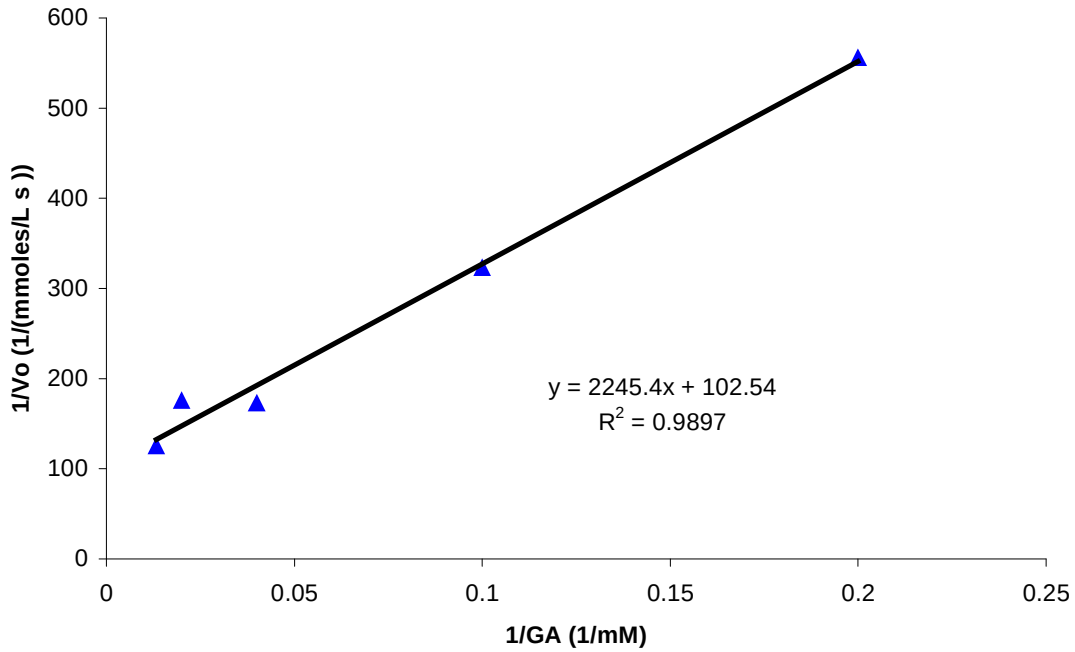


Figure. 4.14 Graph plotting 1/Vo against 1/GA for constant HPA, varying GA, with FPA.

From figure 4.14 the V_{max} can be calculated from the inverse of the intercept to be 0.00976 mM/s. The K_m calculated from the gradient of the line multiplied by the V_{max} to be 21.78 mM for constant HPA, varying GA, with FPA.

For the calculation of K_i for FPA against GA,

$$\mathbf{K_i = [FPA] / ((K_{m\ (with\ FPA)} / K_{m\ (without\ FPA)}) - 1)}$$

$$\mathbf{K_i = 50 / ((21.78 / 4.79) - 1))}$$

$$\mathbf{K_i = 14.12\ mM}$$

4.3.5.5 Summary and discussion

A summary of the kinetics results from the inhibition studies of FPA upon the *E. coli* transketolase enzyme are shown in table 4.6 and 4.7.

	no FPA	with FPA	Ratio	Ki (mM) against HPA
Km HPA (100 mM GA)	41.25	48.53	1.18	283.30
Vm (100 mM GA)	0.00875	0.0121	1.39	

Table 4.6 Summary of data for kinetics at a constant GA concentration (100mM)

	no FPA	Error from SA	with FPA	Error from SA	Ratio	Ki (mM) against GA
Km GA(100 mM HPA)	4.70	0.763	21.78	4.99	4.54	14.11
Vm (100 mM HPA)	0.00754	0.0006	0.0097	0.0017	1.29	

Table 4.7 Summary of data for kinetics at a constant HPA concentration (100mM)

It must be noted that at high concentrations the HPA inhibits the TK, however by removing these data the V_{max} remains unaffected. The HPA inhibition observed may have a systematic effect upon the K_m for GA between FPA and no FPA, however for this study it was assumed not to. The sensitivity analysis performed at constant GA concentration gave a much greater error, due to there only being 4 data points (removal of 75 mM [HPA]). However, the V_{max} and K_m are both essentially the same within large error at constant GA.

Two key observations may be made from these data. The V_{max} remains invariable (within error) at constant HPA concentration, both with and without FPA. The other observation is the 4.5 fold increase in the K_m for GA, both with and without FPA. Therefore it can be said that FPA affects the K_m of GA significantly more than the K_m of HPA. These data suggest a form of competitive binding mostly against GA, the V_{max} has remained constant and the K_m has increased and both these observations are indicative of competitive binding inhibition.

It may be concluded that GA is less able than HPA to displace FPA under these conditions and a number of possible explanations may be drawn from this information.

One possible conclusion assumes that GA would normally bind more strongly in the presence of the enamine intermediate formed from the addition of the HPA in the donor stage of the TK biotransformation. Therefore the GA is then less able to displace the FPA molecule that prefers to bind in the donor stage of the TK biotransformation, where there is no enamine intermediate formed. An alternative possibility is that FPA prefers to bind after the enamine has been formed by the donation of a two-carbon subunit from HPA. This proposal would fit with the observed large increase in the K_m of GA compared to that of the K_m of HPA. Another possible theory, which would be more apt to the previous observations of (Gyamerah and Willetts, 1997) is that the FPA competes directly with the HPA for the active site, however the competitive binding observed is very weak, as the HPA forms the enamine far more readily than the FPA. The FPA would then be positioned within the active site and during the acceptor stage, where the GA molecule strips the α,β-dihydroxyethyl–TPP complex of the two-carbon subunits maybe inhibited by the position of the FPA. This theory would fit with the large increase in the K_m of GA.

Chapter 5

Crystallisation

5.1 Introduction

In order to determine the three-dimensional structure of a protein, the powerful technique of X-ray crystallography is used. This requires crystals to be grown from proteins of interest that diffract using X-ray irradiation to yield good quality diffraction data.

The initial stage is to grow crystals suitable for use in X-ray and crystallography and determine crystallisation conditions. There is a long list of factors that are key to predicting whether a protein may crystallise, which has caused protein crystallographers an array of problems through the ages in attaining reproducible results.

Initially, proteins were crystallised in the 1900s as little more than scientific curiosity, or as a means of 'proof' of protein purity (McPherson, 1991). As time moved on into the 1930's scientists discovered the phenomena of using X-ray diffraction to study macromolecular crystals. Bernal, Crowfoot, Perutz and others were pioneers in utilising the art of protein crystallography in combination with X-ray diffraction (Bernal and Crowfoot, 1934; Perutz, 1964). With technological advances in recombinant DNA, protein purification, robotics, X-ray source, X-ray detectors, and computational devices/software, major 'bottlenecks' for protein crystallographers were removed, which allowed them to delve deeper into the 'art' of crystal growth (Pusey *et al.*, 2005).

Even with these impressive technological advances the science behind protein crystal growth continues to puzzle crystallographers. At its simplest, protein crystallisation consists of two stages, nucleation and growth. During nucleation molecules must overcome an energy barrier to form periodically ordered aggregates of critical mass to 'survive' in a thermodynamic sense. Growth is achieved by making the solid state more attractive to individual molecules than the free state. To promote either stage a condition of supersaturation must be created in the mother liquor (Bergfors, 1999) (Fig. 5.1).

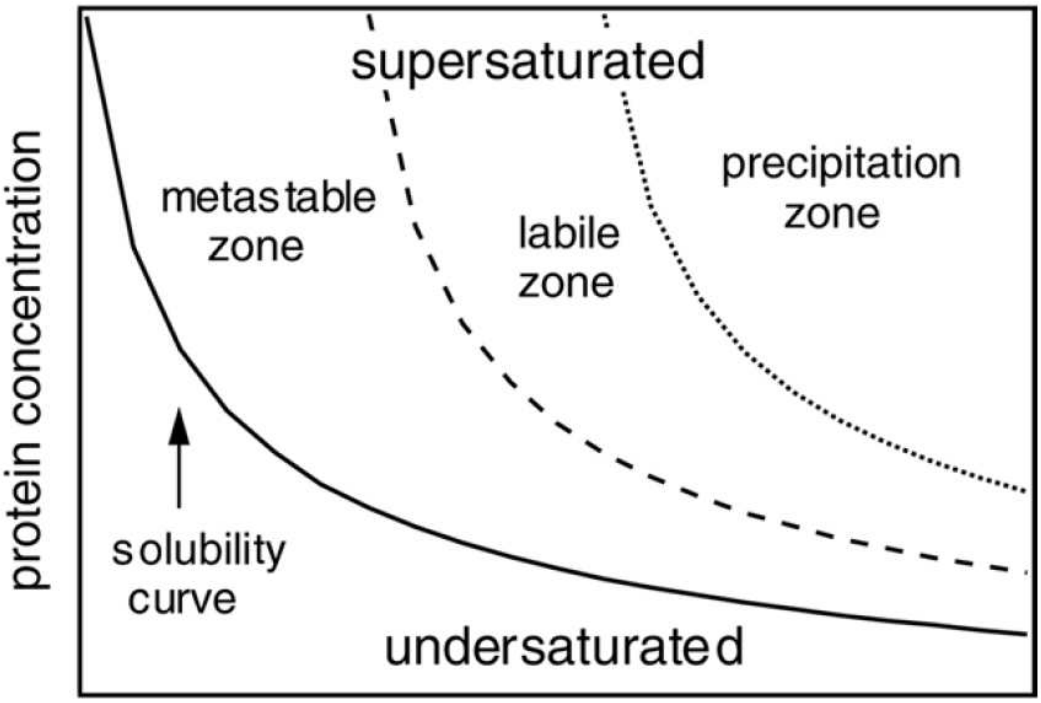


Figure 5.1 A schematic phase diagram showing the solubility of a protein in solution as a function of the concentration of the precipitant present. Taken from McPherson (1999).

A long list of factors has been understood to influence crystal formation, and these can be divided into three parameters, physical, chemical and biochemical, (Table. 5.1). Upon crystallisation of a protein these factors (some more so than others) must be taken into account. In general, crystallisation is usually achieved by varying the physical parameters that affect the solubility of a protein in order to obtain a state of supersaturation. As previously stated, the initial stage of crystallisation is nucleation in which chance collisions between macromolecules occur forming molecular aggregates. It is therefore understandable that the probability of more collisions occurring increases with increasing saturation of solution. So, by systematically altering these physical parameters a supersaturated solution may be achieved, however, this does not guarantee crystal growth. It cannot be stressed enough that the homogeneity of a protein is the single most important factor in producing diffraction quality protein crystals. Hence, during protein purification a protocol is in place in order to gain reproducible homogenous protein. Crystallisation with a homogenous protein sample allows the identical subunits of the protein to arrange themselves into an ordered fashion producing the crystal lattice. If the sample contains impurities the protein may still crystallize, but an undesirable crystal form may be obtained.

Physical
1. Temperature/temperature variation
2. Surfaces
3. Methodology/approach to equilibrium
4. Gravity
5. Pressure
6. Time
7. Vibrations/sounds/mechanical perturbations
8. Electrostatic/magnetic fields
9. Dielectric properties of medium
10. Viscosity of medium
11. Rate of equilibration
12. Homogenous or heterogeneous nucleants
Chemical
1. pH
2. Precipitant type
3. Precipitant concentration
4. Ionic strength
5. Specific ions
6. Degree of supersaturation
7. Reductive/oxidative environment
8.Concentration of macromolecules
9. Metal ions
10. Crosslinkers/polyions
11. Detergents/surfactants/amphophiles
12. Non-macromolecular impurities
Biochemical
1. Purity of macromolecule/impurities
2. Ligands, inhibitors, effectors
3. Aggregation state of macromolecule
4. Post-translational modifications
5. Source of macromolecule
6. Proteolysis/hydrolysis
7. Chemical modifications
8. Genetic modifications
9. Inherent symmetry of the macromolecule
10. Stability of macromolecule
11. Isoelectric point
12. History of sample

Table 5.1 Factors effecting crystallisation, taken from (McPherson, 2004).

5.2 Methods of crystallisation

There are numerous methods of protein crystallisation ranging from simple batch methods to more complex dialysis methods, but the most commonly used method is the vapour diffusion method.

The vapour diffusion method allows the separate protein and reservoir solution to attain equilibrium by diffusion through air in a sealed system. Nucleation occurs when the protein concentration (in the protein solution) increases due to the equilibrium movement of the water to the more hygroscopic reservoir solution. In this method nucleation occurs through macromolecular collisions forming aggregates.

The batch method is the most basic; it combines all the crystallisation components in one solution, which is sealed and allowed to nucleate. Initially this method was performed on a millilitre scale, but with the recent advances of robotics a microlitre scale using an inert oil cap to prevent evaporation has been derived. This microbatch method allows many conditions to be screened using minimal amounts of protein.

The dialysis method allows the diffusion of low molecular weight molecules to pass in and out of the protein solution through a semi-permeable membrane. Unlike vapour diffusion the protein concentration is kept constant throughout the crystallisation and is the most effective crystallisation method for slight change in ionic strengths.

The two techniques used during these studies were the sitting drop vapour diffusion and microbatch methods

5.2.1 Sitting drop vapour diffusion method

Vapour diffusion techniques are the most popular methods used in crystallisation (Bergfors, 1999). Sitting drop vapour diffusion is primarily used to accommodate larger volumes of protein (10-20 µl), leading to growth of larger crystals. This is due to the increased amount of available protein and the slower equilibration rates associated with larger sample volumes (Fox and Karplus, 1993).

In setting up a sitting drop vapour diffusion system, the precipitant of choice (500 µl) is pipetted into the well in the Linbro plate, which has been cleaned previously

with compressed air to remove dust particles. A microbridge is then lowered into the well and the protein sample mixed with a low concentration of the precipitant (10-20 µl) is pipetted onto the microbridge. The Linbro plate well is then sealed using vacuum grease and a coverslip (Fig. 5.2), the system is now closed and equilibrates as the water vapour is drawn from the protein drop towards the surrounding vapour and reservoir below. Supersaturation is gradually reached over time as the concentration of the precipitant in the protein drop reaches that of the concentration of the reservoir. The solubility of the protein then reaches a point where the stable crystalline lattice is formed.

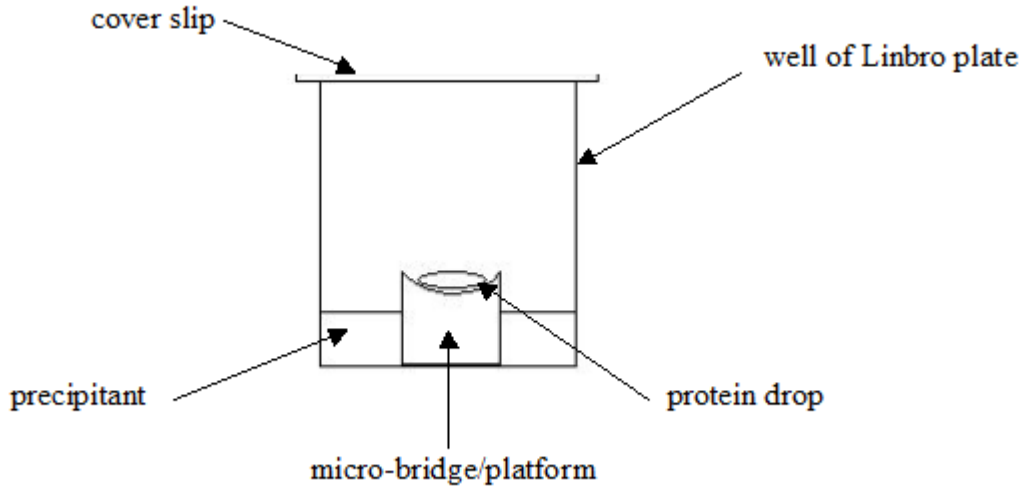


Figure 5.2 Diagram showing the set-up of a sitting drop vapour diffusion experiment.

5.2.2 Microbatch method

The microbatch method is a relatively new method, derived from the batch method but on a smaller (microlitre) scale, hence microbatch. Microbatch allows high-throughput screening of a range of crystallisation condition within a small time frame (D'Arcy *et al.*, 2003). Microbatch eliminates the need for addition of reservoir solution or any form of sealing as the protein/precipitant drops are immediately placed under a layer of oil, protecting the protein from both oxidation and airborne contaminants (D'Arcy *et al.*, 2004).

In setting up a microbatch plate, the robot used for screening and optimisation was the Oryx6 system crystallisation robot (Douglas Instruments, Cambridge, UK). The robot pipettes the desired amount of protein solution for the experiment through one of the channels on the dual channel pipette. The robotic arm with the dual channel pipette then proceeds to the well containing the screening precipitant. The pipette is then lowered into the well and draws up the set amount of precipitant through the second channel. The pipette then deposits both the protein and precipitant onto the microbatch plate, which is then covered with a layer of oil through a secondary robotic pipette arm (Fig. 5.3). The system is now sealed and incubated in differing environments to allow crystal growth.

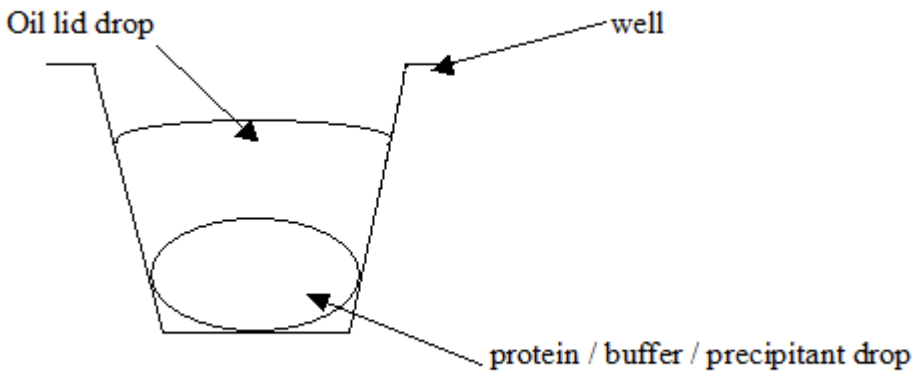


Figure 5.3. Diagram showing the set up of a microbatch experiment.

5.3 Crystallisation of Escherichia *coli* transketolase

In order to determine the conditions for which a protein will crystallise it is common practice to screen a wide range of conditions, knowledge of the protein's history is also beneficial in attaining diffraction quality crystals. The *E. coli* enzyme transketolase had been previously crystallised using sitting drop vapour diffusion over a reservoir of 48 % ammonium sulfate in 50 mM PIPES buffer pH 6.4 (Littlechild *et al.*, 1995). It was initially decided to recreate these conditions and attempt screening for other crystallisation conditions.

5.3.1 Microbatch method crystallisation of *E. coli* transketolase

5.3.1.1 Materials and methods

5.3.1.1.1 Sample preparation

The purified TK protein was prepared as described in section 2.3.1 using a protein concentration of 20 mg/ml in crystallisation buffer E (50 mM PIPES buffer pH 6.4 containing 2 mM Thiamine pyrophosphate and 9 mM $CaCl_2$).

5.3.1.1.2 Crystallisation trials

The prepared protein solution was screened against NeXtal 'The pHclear™' (Qiagen, UK) (See details in appendix II) using the Oryx6 system crystallisation robot (Douglas Instruments, Cambridge, UK) using the method described in section 2.3.2.2.

5.3.1.2 Results

Initial crystallisation trials produced multiple results within a range of alternative conditions. All of the crystals grown using the NeXtal 'The pHclear™' screen grew within a 10 day period, and varied in shape and form. The results obtained from these trials are summarised in table 5.2.

Crystal clusters were grown in a high molecular weight PEG precipitant containing 0.1 M MES, 20 % [w/v] PEG 6000, pH 6.0. These crystal clusters had

typical dimensions of 0.2 mm x 0.1 mm x 0.1 mm (Fig. 5.4), but their use was limited because of their inability to be separated. Even so, evidence of the crystallisation of *E. coli* transketolase in with PEG precipitant would prove advantageous for substrate soaks. The crystal was found to diffract to a low resolution using an 'in-house' Cu/K α rotating anode radiation source, confirming protein diffraction pattern and not that of a salt.

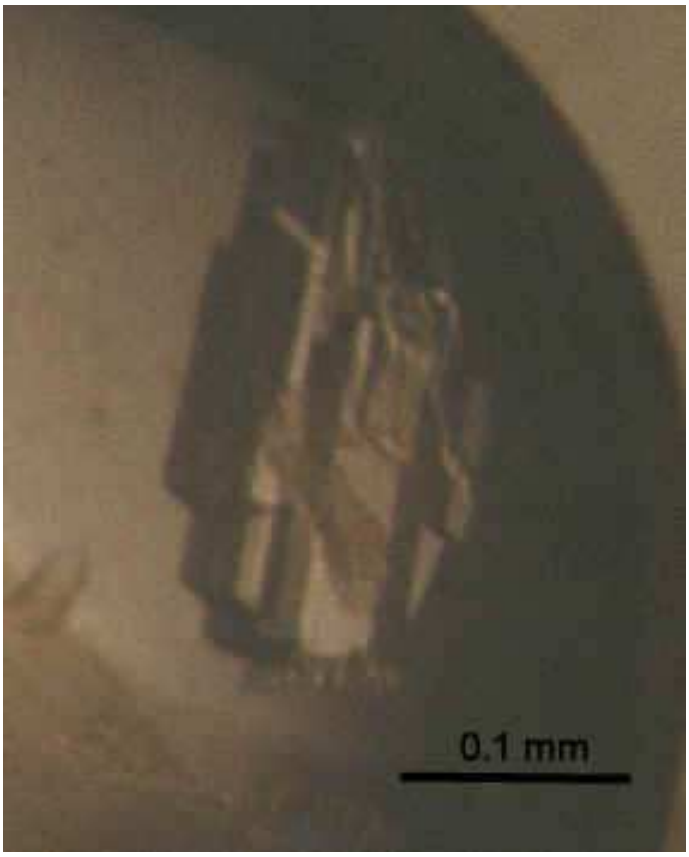


Figure 5.4 Photograph of microbatch *E. coli* transketolase crystals obtained using microbatch method from high molecular weight PEG.

Two crystals were also grown in similar ammonium sulfate precipitant conditions, but within a pH unit each. The first in 0.1 M HEPES, 3.2 M $(NH_4)_2SO_4$, pH 7.0 (Fig. 5.5), and the other in 0.1 M MES, 3.2 M $(NH_4)_2SO_4$, pH 6.0 (Fig. 5.6). These crystal forms had similar dimensions to each other of 0.25 mm x 0.05 mm x 0.01 mm.

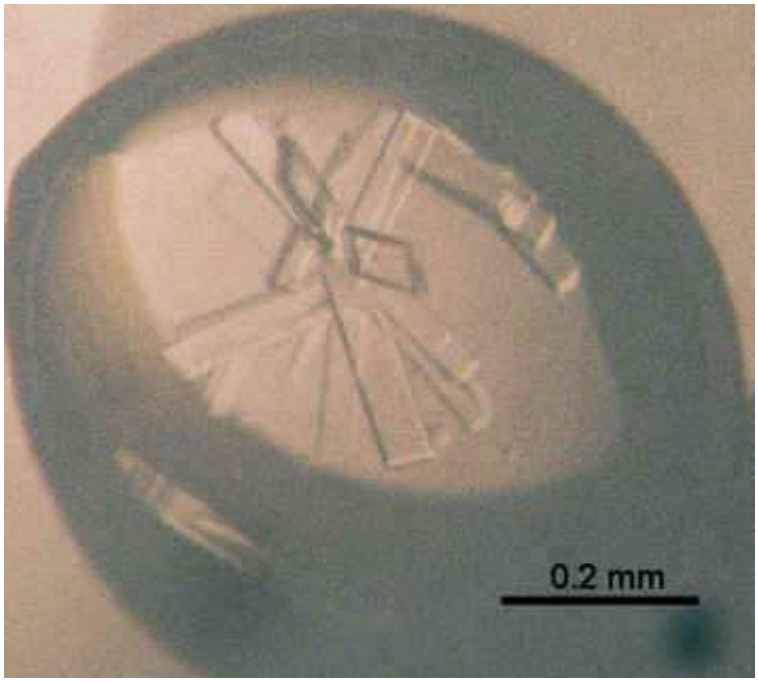


Figure 5.5 Photograph of *E. coli* transketolase crystals grown using the microbatch method in HEPES and ammonium sulfate conditions.

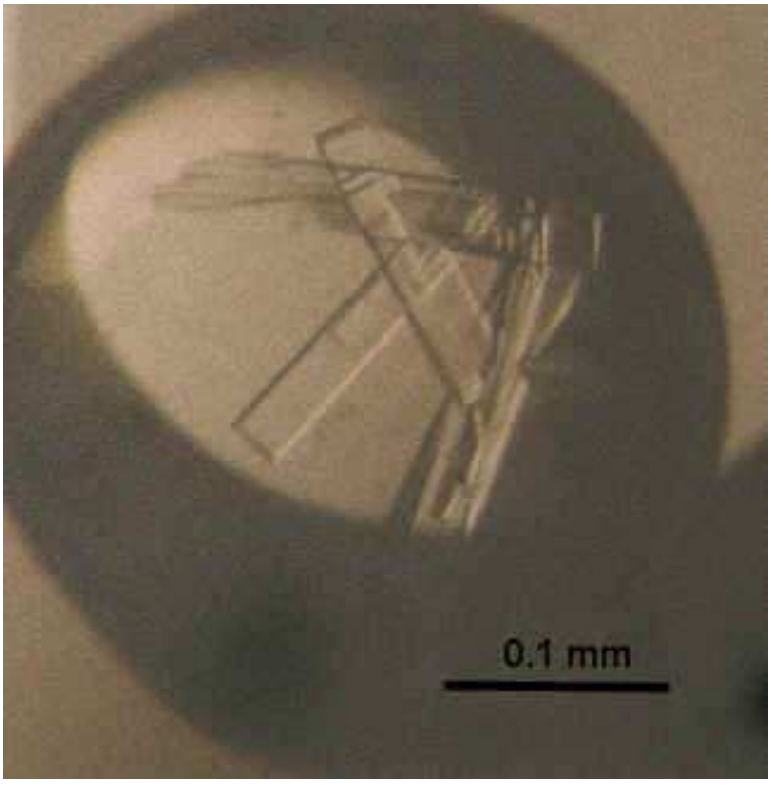


Figure 5.6 Photograph of *E. coli* transketolase crystals grown using the microbatch method in MES and ammonium sulfate conditions**.**

The crystals, although superimposed, were possible to separate from each other. Both types of crystal diffracted to a resolution of approximately 2.5 Å on the 'in-house' radiation source, but were deemed unsuitable for data collection at this diffraction quality, since a structure of the protein at this resolution had already been solved.

Conditions from NeXtal'The pHclear™ screen	Results obtained	Diffraction
0.1 M MES, 20 % [w/v] PEG 6000, pH 6.0	Untidy crystal cluster	Yes (poor)
0.1 M HEPES, 3.2 M $(NH_4)_2SO_4$, pH 7.0	Superimposed crystal rectangular plates	Yes (2.5 Å)
0.1 M MES, 3.2 M $(NH_4)_2SO_4$, pH 6.0	Superimposed crystal rectangular plates	Yes (2.5 Å)

Table 5.2 Summary of the results obtained from the microbatch crystallisation trial for *E. coli* transketolase (20 mg/ml) in crystallisation buffer E.

5.3.2 Vapour diffusion method for crystallisation of *E. coli* transketolase

5.3.2.1 Materials and methods

5.3.2.1.1 Sample preparation

The purified TK protein was prepared as described in section 2.3.1 using a protein concentration of 20 mg/ml in three different crystallisation buffers C, D and E.

Buffer C 50 mM PIPES buffer pH 6.4 containing 20 mM Thiamine pyrophosphate and 90 mM $CaCl_2$.

Buffer D 50 mM PIPES buffer pH 6.4 containing 10 mM Thiamine pyrophosphate and 45 mM $CaCl_2$.

Buffer E 50 mM PIPES buffer pH 6.4 containing 2 mM Thiamine pyrophosphate and 9 mM $CaCl_2$.

5.3.2.1.2 Crystallisation trials

The prepared protein solutions were screened using variations on the ammonium sulfate conditions used by Professor Jenny Littlechild in 1995 (Littlechild *et al.*, 1995). The reservoir solutions contained a range of ammonium sulfate concentrations (40 % - 60 % saturation) in 50 mM PIPES, pH 6.4, 9 µl of protein sample in the three different crystallisation buffers, see appendix II for details.

5.3.2.3 Results

Initial crystallisation trials using buffer C as the crystallisation buffer yielded a series of precipitates between 40 % - 60 % saturation of ammonium sulfate. Results using crystallisation buffers D and E, which had lower concentrations of cofactors TPP and $CaCl_2$, produced crystal clusters after 16 days.

Crystals clusters from crystallisation buffer D grown in 50 % ammonium sulfate saturation had typical dimension 0.1 mm x 0.01 mm x 0.01 mm (Fig. 5.7), but again their use was limited because of their inability to be separated.

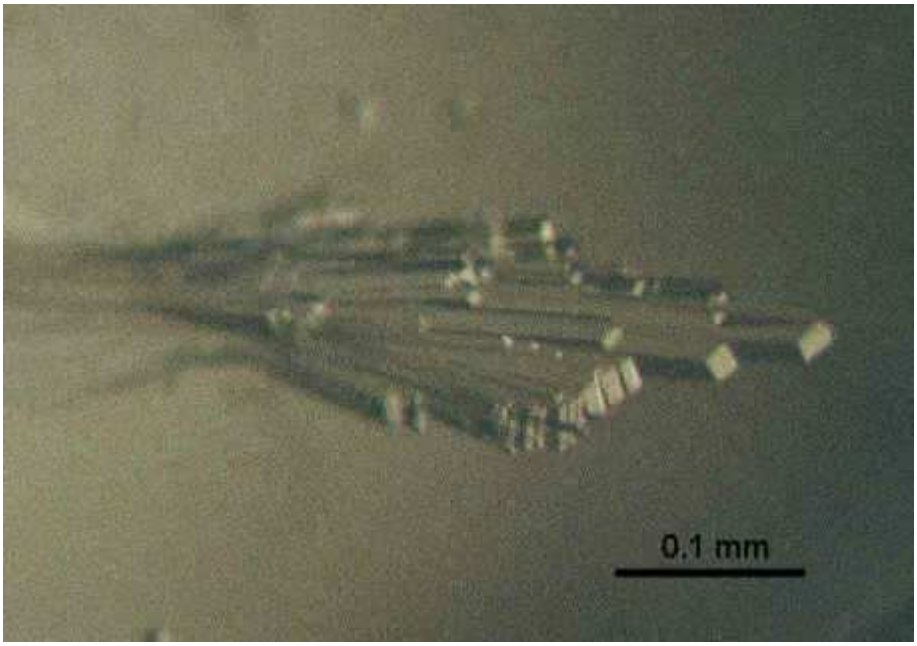


Figure 5.7 Photograph of *E. coli* transketolase crystals grown in buffer D using the sitting drop vapour diffusion method with a 50 % $(NH_4)_2SO_4$, 50 mM PIPES, pH 6.4 precipitant.

The larger more ordered crystals clusters from crystallisation buffer E grown in 50 % ammonium sulfate saturation had typical dimension 0.4 mm x 0.15 mm x 0.05 mm (Fig. 5.8), even though attempts were made to separate these larger crystal clusters, it was deemed beneficial to further optimise the conditions to obtain single crystals.

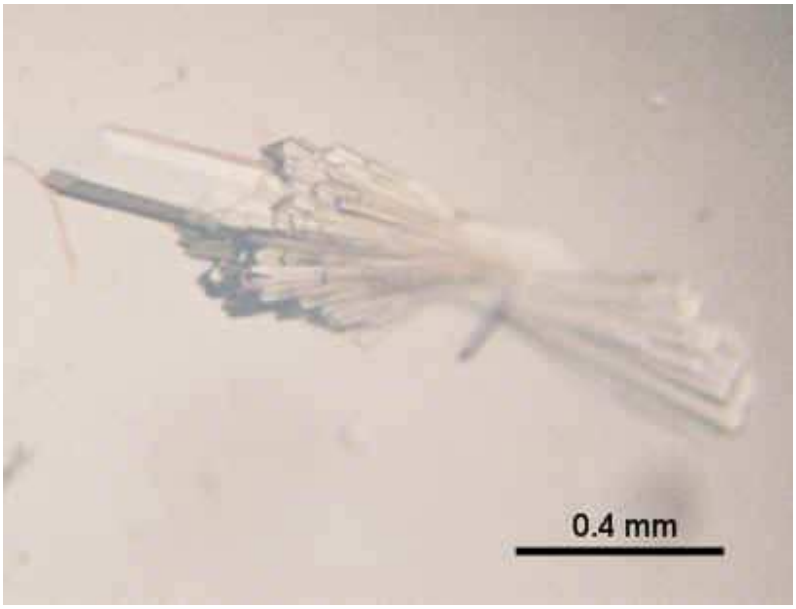


Figure 5.8 Photograph of *E. coli* transketolase crystals grown in buffer E using the sitting drop vapour diffusion method with a 50 % $(NH_4)_2SO_4$, 50 mM PIPES, pH 6.4 precipitant.

The crystallisation trial results are summarised in table 5.3.

% $(NH_4)_2SO_4$	Buffer + pH in Solvent Well	Sitting Drop Composition	Initial protein concentration + Buffer	Results
40-60	50 mM PIPES pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C	Precipitates
40-60	50 mM PIPES pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D	Small crystal clusters
40-60	50 mM PIPES pH 6.4	9 µl +1 µl	20 mg/ml + Buffer E	Crystal clusters

Table 5.3 Results of the sitting drop vapour diffusion crystallisation trials varying the ammonium sulfate concentration and crystallisation buffer.

Crystals with the greatest size and order from the crystallisation trials were grown from 45 %-50 % ammonium sulfate saturation with crystallisation buffer E. Optimisation of these conditions yielded a large single crystal from 48 % ammonium sulfate in 50 mM PIPES, pH 6.4 of typical dimension 0.8 x 0.3 x 0.1 mm (Fig. 5.9).

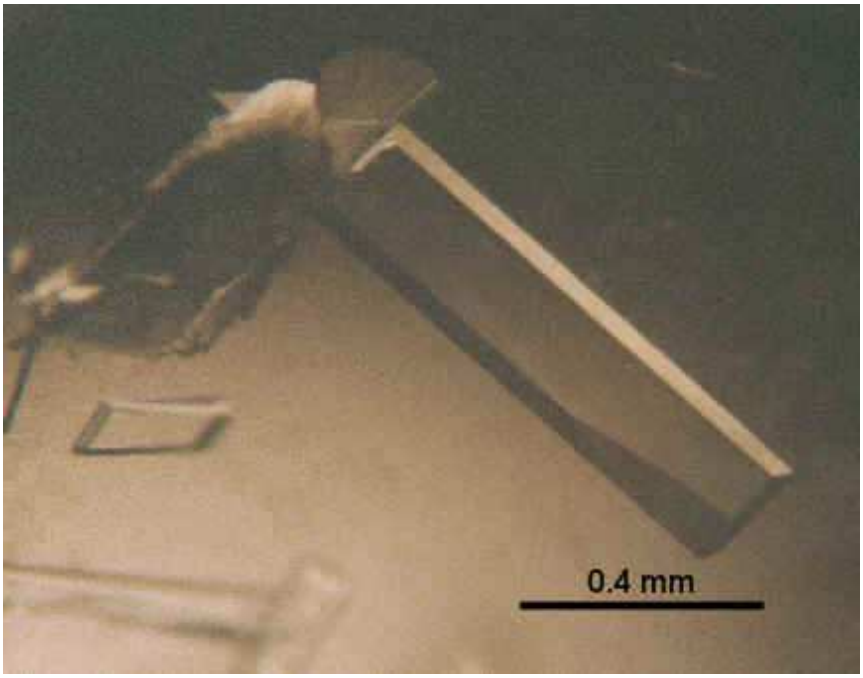


Figure 5.9 Photograph of *E. coli* transketolase crystals grown with buffer E using the sitting drop vapour diffusion method with a 48 % $(NH_4)_2SO_4$, 50 mM PIPES, pH 6.4 precipitant.

The crystal shown in figure 5.9 was of excellent diffraction quality, and found to diffract to 1.55 Å using the 'in-house' radiation source.

5.3.3 Co-crystallisation of *E. coli* transketolase with hydroxypyruvate and fluoropyruvate

5.3.3.1 Materials and methods

5.3.3.1.1 Sample preparation

The purified TK protein was prepared as described in section 2.3.3.1.3 using a protein concentration of 20 mg/ml in crystallisation buffers E with ranging β-hydroxypyruvic acid (Fluka, UK) concentrations (0.3 mM – 40 mM), or β-fluoropyruvic acid (sodium salt) (Fluka, UK) concentrations (0.3 mM – 40 mM).

5.3.3.1.2 Crystallisation trials

Co-crystallisation experiments were conducted using the optimised vapour diffusion conditions used in section 5.3.2.1.2.

5.3.3.2 Results

Large single crystals of typical dimension 0.6 mm x 0.3 mm x 0.05 mm which were separated from larger clusters grew from 48 % ammonium sulfate in 50 mM PIPES, pH 6.4, for co-crystallisation using fluoropyruvate (Fig. 5.10). The crystals grew in all concentrations from 0.3 mM to 40 mM, but only the 40 mM crystals were deemed suited for X-ray diffraction due to increased likelihood of trapping the fluoropyruvate in the active site.

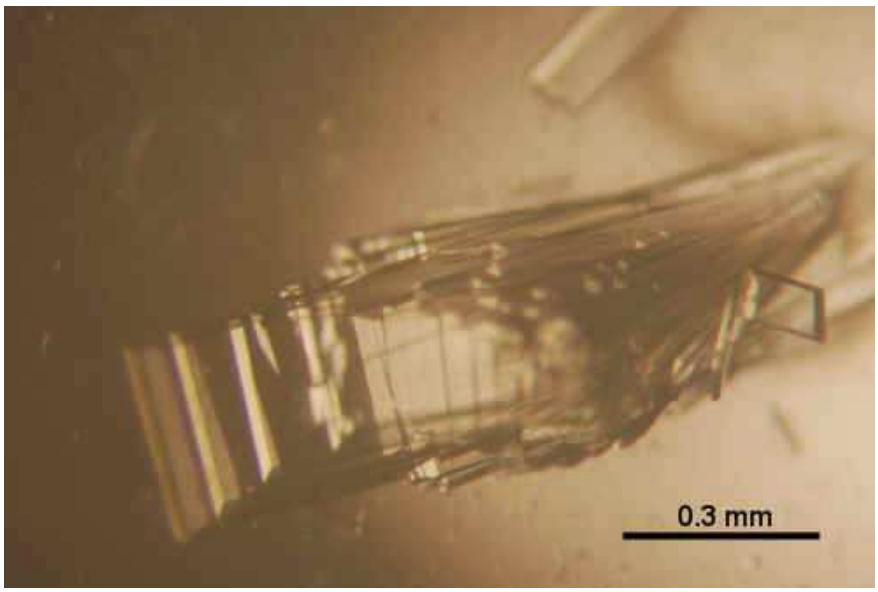


Figure 5.10 Photograph of *E. coli* transketolase crystals grown with buffer E using the sitting drop vapour diffusion method with 40 mM fluoropyruvate and 48 % $(NH_4)_2SO_4$, 50 mM PIPES, pH 6.4 precipitant.

A large single crystal of typical dimension 0.5 mm x 0.2 mm x 0.01 mm grew from 48 % ammonium sulfate in 50 mM PIPES, pH 6.4, for co-crystallisation using hydroxypyruvate (Fig. 5.11). The crystal grew separate from the larger crystal clusters within the 40 mM hydroxypyruvate concentration. Again crystals were present in all concentrations varying from 0.3 mM to 40 mM but for the same reasons as above only the 40 mM crystals were deemed suited for X-ray beam diffraction.

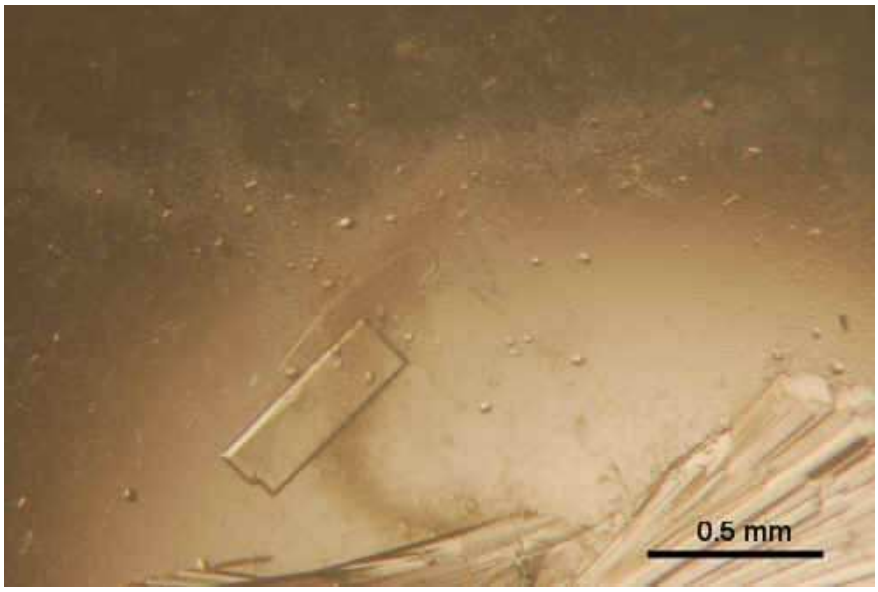


Figure 5.11 Photograph of *E. coli* transketolase crystals grown with buffer E using the sitting drop vapour diffusion method with 40 mM hydroxypyruvate and 48 % $(NH_4)_2SO_4$, 50 mM PIPES, pH 6.4 precipitant.

5.3.4 Crystallisation of *E. coli* transketolase mutants H26Y and D469Y

In order to gain further mechanistic knowledge, crystallisation of active site mutants of TK was deemed necessary. It was assumed that both of the mutants were stable, as they produced very similar purification results to the native transketolase. Hence, very similar conditions were used in crystallisation.

5.3.4.1 Materials and methods

5.3.4.1.1 Sample preparation

The purified TK mutant proteins H26Y and D469Y were prepared as described in section 2.3.3.1.3 using a protein concentration of 20 mg/ml in crystallisation buffers E.

5.3.4.1.2 Crystallisation trials

Crystallisation trials for H26Y and D469Y were carried out by vapour diffusion over reservoirs containing 48-58 % ammonium sulfate in 50 mM PIPES, pH 6.4. The protein droplets contained 9 µl of 20 mg/ml H26Y or D469Y enzyme in 50 mM PIPES buffer pH 6.4 containing 2mM TPP and 9 mM $CaCl_2$ (Buffer E) and 1 µl 100 % (saturation) ammonium sulfate solution.

5.3.4.2 Results

Subsequent crystals for the H26Y mutant grew in crystal clusters with single crystals from 50 % ammonium sulfate in 50 mM PIPES, pH 6.4 of typical dimension 0.4 mm x 0.15 mm x 0.01 mm (Fig. 5.12). The crystals that grew from 48 % - 58 % for D469Y mutant were fine needle clusters (Fig. 5.13), which were unsuitable for any form of X-ray diffraction. Changes in protein concentration and condition were experimented with, but to no avail.

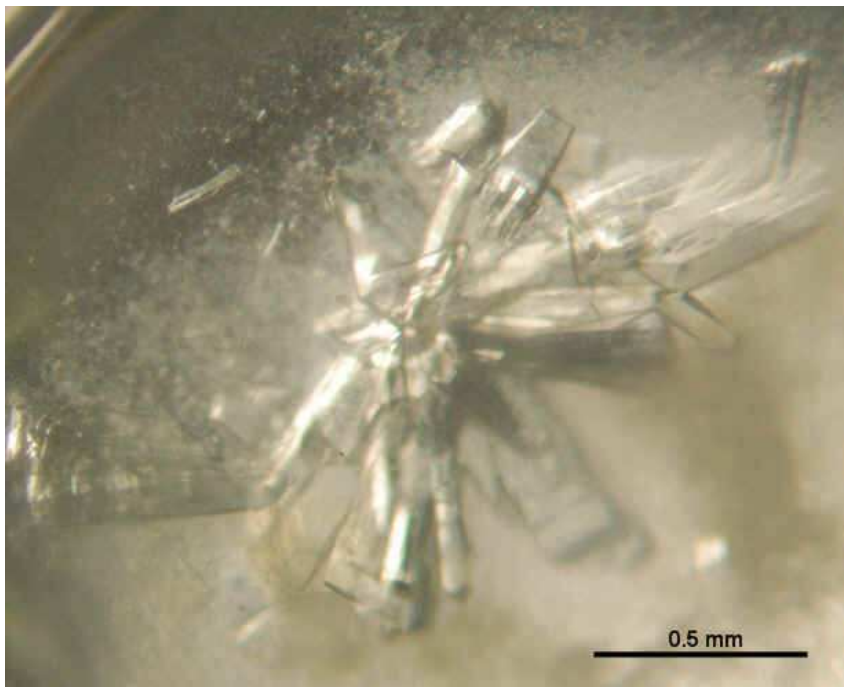


Figure 5.12 Photograph of H26Y mutant *E. coli* transketolase grown with buffer E using the sitting drop vapour diffusion method with 50 % $(NH_4)_2SO_4$, 50 mM PIPES, pH 6.4 precipitant.

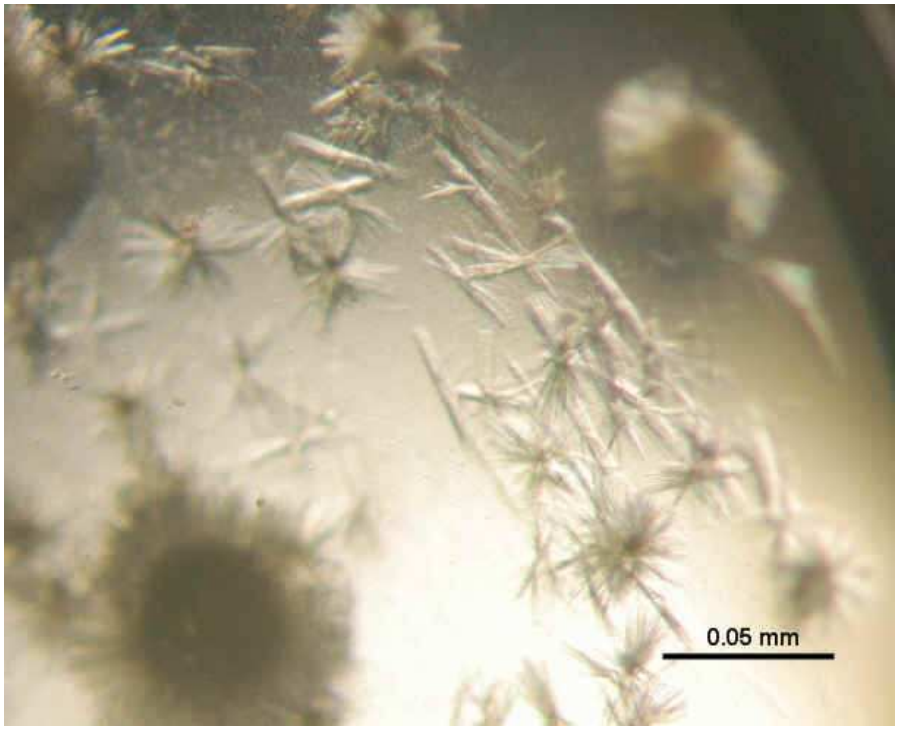


Fig. 5.13 Photograph of D469Y mutant *E. coli* transketolase grown with buffer E using the sitting drop vapour diffusion method with 55 % $(NH_4)_2SO_4$, 50 mM PIPES, pH 6.4 precipitant.

5.4 Discussion

Recreation of the crystallisation conditions of past work on the *E. coli* transketolase (Littlechild *et al.*, 1995) proved to stand the test of time, although the concentrations of co-factors was decreased 10-fold in order to gain the large crystals.

Attempts to crystallise the *E. coli* transketolase using the microbatch method gave excellent results in a shorter period of time to that of the vapour diffusion method, but yielded smaller crystals. The crystals that grew in ammonium sulfate were of similar conditions to that of the vapour diffusion experiments, but the crystals that grew in PEG provided another alternative condition. The PEG conditions were similar to the PEG conditions that were used to crystallise the yeast transketolase which was the first transketolase structure reported in 1993 (Nilsson *et al.*, 1993).

The co-crystallisations trials using the inhibitor fluoropyruvate and substrate hydroxypyruvate provided a range of crystals that were suitable for X-ray diffraction experiments. The crystals grew in all of the experimental concentrations of fluoropyruvate and hydroxypyruvate. The 40 mM concentration grown crystals were chosen as the most suitable candidates for diffraction due to likelihood of substrate in the active site.

Crystallisation of the variant H26Y yielded crystals similar to that of any other *E. coli* transketolase previously crystallised, the crystals were of good size and diffraction quality. Crystallisation of the variant D469Y however, proved more testing, as only small needle crystals could be grown.

Chapter 6

X-ray Diffraction Studies on E. *coli* transketolase

The studies conducted during this chapter were aided and performed under the supervision of Dr. Michail Isupov (University of Exeter, UK) and Dr. Tsutomu Nakamura (National Institute of Advanced Industrial Science and Technology, Japan)

6.1 Introduction

In order to construct the three-dimensional structure of a protein, the protein crystals are assessed for their suitability to yield good quality diffraction data. Protein crystals that produce valid X-ray diffraction data are subjected to full X-ray diffraction data studies. In collection of X-ray diffraction data the crystal is mounted on a goniometer head and exposed to an X-ray source, the X-ray beam is scattered by the electrons within the crystal generating a two-dimensional diffraction pattern of routinely arranged spots called a reflection frame. The data is then collected by the 'rotation method' (Minor *et al.*, 1974). The crystal data set is then transformed into a three-dimensional electron density map using complex mathematics. The positions of the atoms within the crystal can then be deduced.

6.1.1 History of the x-ray source and early protein crystallography

The phenomena of the X-ray was initially discovered by the physicist Wilhelm Röntgen in 1885 (Rontgen, 1896), and followed by the polarisation and characterisation of X-rays by Charles Barkla in 1905 and 1909 respectively (Barkla, 1904; Barkla, 1909). The nature of this new phenomenon, particle or a wave was disputed for the next decade by William Henry Bragg and Charles Barkla, respectively, until 1922 when Arthur Compton confirmed William Henry Bragg's photon model theory through studying the scattering of X-rays from electrons (Compton, 1923). Max von Laue and Paul Peter Ewald proposed that the use of electromagnetic radiation in the form of X-rays would have a shorter wavelength comparable to the unit-cell spacings in crystals (Ewald, 1979). Laue was later awarded a Nobel prize for physics in 1914, for his copper sulfate crystal diffraction studies where he developed a law relating the scattering angle

and the size of orientation of the unit cell spacings in the crystal (Friedrich *et al.*, 1912). William Lawrence Bragg later developed on the pioneering work by Laue, creating Braggs's Law (section 6.1.2) connecting the observed scattering with reflections from evenly spaced planes within the crystal (Bragg, 1912). In 1914, structures of simple molecules and minerals such as sodium chloride and diamond were determined using trial and error by changing atomic coordinates until they matched that of the diffraction pattern intensities. Later the mathematical Fourier transform method was introduced, aiding the calculation of electron density from the diffractions patterns. However, secondary scattering compromised the electron density calculations, giving rise to the phase determination problem (section 6.1.2). During this time protein crystals were starting to be used in structural studies, the first X-ray diffraction pattern of a protein was collected in 1934 (Bernal and Crowfoot, 1934). In 1950, early structures of the α-helix and the β-sheet were proposed by Pauling and Corey (Pauling and Corey, 1951a; Pauling and Corey, 1951b) before the ground-breaking publication of the DNA helix by Watson and Crick in 1953 (Watson and Crick, 1953). Mid-way through the 20th century advances were made in computational instruments, with computer software assisting in the analysis of X-ray diffraction data. The other major advance in crystallography at this time was the birth of the principles of synchrotron radiation, where John Blewett of General Electric (NY, USA) urged experimental tests to be carried out on their 70 MeV synchrotron after being inspired by the work of Iwanenko and Pomeranchuk (Iwanenko and Pomeranchuk, 1944; Blewett, 1998). On the 24th April 1947, technician Floyd Haber observed sparkling in the transparent synchrotron tube that had been assembled by Robert Languir and Henry Pollock in October 1946 (Pollock, 1983), the sparkling was proven to be the first 'eye-balled' sight of synchrotron radiation (Elder *et al.*, 1947). A decade later Parratt proposed in principle the use of synchrotron radiation for crystallography (Parratt, 1959) after the first protein X-ray crystal structure had been solved by John Kendrew a year earlier (Kendrew *et al.*, 1958). In 1976, the first X-ray protein crystallography experiments using synchrotron radiation took place in Stanford, USA (Phillips *et al.*, 1976), since then X-ray crystallographers around the world have benefited from the increased flux and variable wavelength that the synchrotron radiation provides. Crystallographers had always found that data being collected from protein crystals deteriorated with time, this was believed to be due to radiation damage caused by the X-ray flux. This prevented crystallographers obtaining a complete data set, as the protein crystal gradually loss the ability to diffract. This was due to the X-ray radiation disordering the crystal lattice, before a complete data set could be obtained. It was

observed that collection at sub-zero temperatures reduced radiation damage and so a method of cryofreezing was conceived. Protein crystals consist of a high percentage of water that forms ice when frozen, so a cryoprotectant mother liquor was designed to reduce the formation of ice during freezing. A cryoprotectant typically contains an anti-freezing agent of the likes of glycerol, but all forms of small molecules have been used such as polyethylene glycol, glucose and other organic solvents. The frozen crystal in cryoprotectant mother liquor was then placed in a liquid nitrogen cryostream during data collection. Full data sets and increased quality of diffraction meant that the cryofreezing technique was a huge step forward in X-ray crystallography (Petsko, 1975).

1.2 Crystal symmetry

Protein crystals are the visible product of a series of regularly ordered repeating protein molecules, from which a unit cell can be obtained. The unit cell is made up of six parameters, the lengths of the unit cell edges (a,b,c) and the angles adjacent to them (α,β,γ). These lengths and angles describe the overall shape of the unit cell (Fig. 6.1).

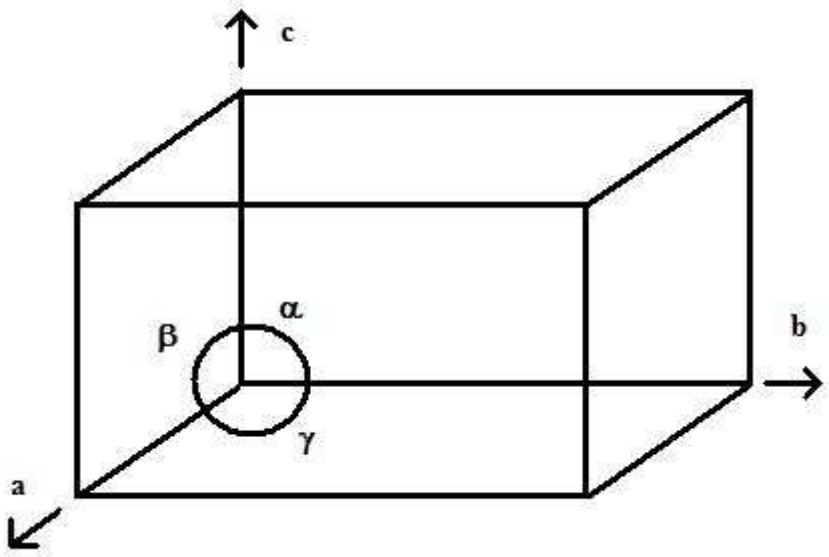


Figure 6.1 The Unit Cell, lengths (a,b,c), angles (α,β,γ).

Unit cells can consist of just the one or many asymmetric units arranged in a pattern, which maps the symmetry of the crystal. This is known as the space group of the crystal, in which an asymmetric unit can be translated to another by a series of symmetry operations forming a set of equivalent, parallel planes of atoms throughout the crystal lattice. Indices h, k, and l are used to describe the crystal lattice planes.

There are seven crystal systems known, all of which describe the most minimal symmetry the crystal can consist of, even though it is possible for a crystal to contain more than one of the crystal systems (Table 6.1). The three forms of symmetry present in crystals are pure rotation axes, pure translations and screw axes. A pure rotation axis is simply a rotation of 1,2,3,4, or 6 –fold, with pure translations being simply that of a translation, and a screw axis a combination of the previous two along an axis.

Crystal System	Rotational symmetry	Cell dimension constraints	Lattice
Triclinic	1-fold	None	**P**
Monoclinic	2-fold	$\alpha=\gamma=90°$	**P,C**
Orthorhombic	Three perpendicular 2-folds	$\alpha=\beta=\gamma=90°$	**P,I,F,C**
Tetragonal	4-fold	a=b, $\alpha=\beta=\gamma=90°$	**P,I**
Trigonal	3-fold	a=b, $\alpha=\beta=\gamma=120°$	**P,R**
Hexagonal	6-fold	a=b, $\alpha=\beta=\gamma=120°$	**P**
Cubic	3 and 2-folds	a=b=c, $\alpha=\beta=\gamma=90°$	**P,F,I**

Table 6.1 The seven crystal system.

6.1.3 X-ray scattering by crystals

When X-rays are directed towards a crystal, the X-rays are scattered by the electrons of each atom within the crystal forming a diffraction pattern. The reflections from the diffraction pattern observed correspond to the dimensions of the crystal. The reflections can only be seen if the path difference between the waves scattered by the adjacent lattice planes are multiples of the wavelength of the radiation. If this is not the case then destructive interference between the waves occur and no reflections are observed. Hence, Bragg showed that for diffraction to occur then the angle of incident

of radiation must be equal to that of the angle of reflection (Bragg, 1912) (Fig. 6.2) (Eqn. 6.1).

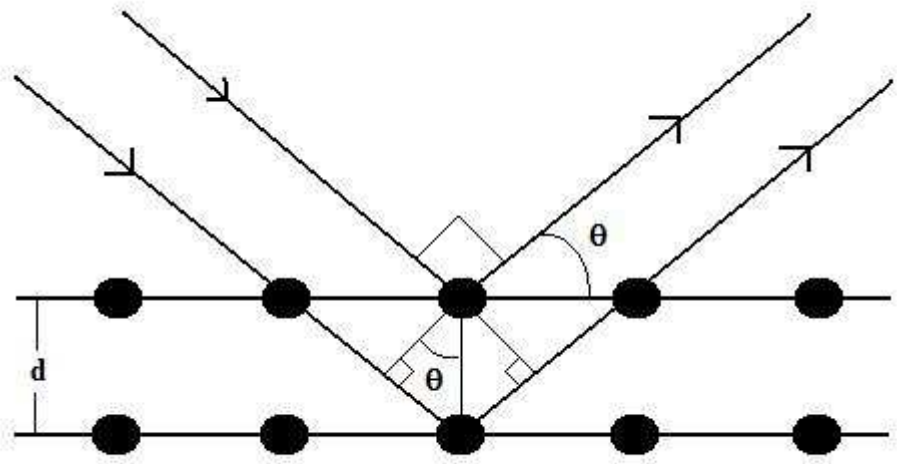


Figure 6.2 Conditions of Bragg's law, allowing diffraction to be observed.

$$2d_{hkl}\sin\theta = n\lambda$$

Equation 6.1 Bragg's Law. Where λ wavelength of radiation, d_{hkl} is the spacing between lattice planes, n is an integer and θ is the angle of incident.

The relationship between the observed diffraction pattern and the crystal is a reciprocal one, where the lattice of the diffraction pattern is the reciprocal of the lattice from the actual crystal causing the scattering. Hence, all measurements obtained from the diffraction pattern are inversely proportional to the actual crystal lattice, i.e. the lengths of the translations in the reciprocal lattice are inversely proportional in the crystal lattice. Therefore a small reciprocal lattice and close diffraction pattern would give a large unit cell and crystal lattice. The angle of scattering determines the size and shape of the unit cell, with the intensities of the reflections being calculated from the positions of the atoms in the unit cell. These reflections are given in Miller indices h, k, and l, where each indices represents a plane that divides each unit cell vector a, b, and c, respectively. Therefore, both the intensity and the angle of the reflection are collected during the X-ray diffraction experiments. Once the diffraction data is collected the 'real' space is calculated by deciphering the structural factors; from this Fourier

transformation mathematics is then used to calculate the electron density. Here now lies the phase problem, in which the exact location of the electrons cannot be inferred, as the information of the phase of the diffracted beam from a single diffraction pattern cannot be established.

6.1.4 Protein structure determination

In determining a protein structure the observed intensities are matched to that calculated from the model of the hypothesized protein of the crystal lattice. This brings us on to the phase problem described earlier, in which the phases from the X-ray diffraction experiments must be determined. There are a number of methods suitable for solving this issue, but all are based around the calculated difference of Patterson maps (Patterson, 1934). The following section 6.1.4 has been well reviewed and referenced by Jan Drenth in 'Principles of protein X-ray crystallography' (Drenth, 1999).

6.1.4.1 Isomorphous replacement

Isomorphous replacement was the preferred method for protein structure determination by the early protein crystallographers of the 1950's and was used by Kendrew in the determination of the sperm whale myoglobin crystal structure (Kendrew *et al.*, 1958), but was later superseded by alternative approaches. The method is based on the comparison of two isomorphous crystals, in which one of them contains one or more strong scattering centres. As both crystals will have the same unit cell, addition of a heavy atom to one of the crystals via soaking or co-crystallisation, will alter the reflections and solve the phase problem.

6.1.4.2 Anomalous diffraction

The anomalous diffraction technique arose with the introduction of synchrotron radiation, where it was possible to alter the wavelength of the radiation source. The X-ray wavelengths are scanned past the edge of an atom causing anomalous differences, which can be used to locate the atoms position. The synchrotron beam is tuned to different wavelength allowing anomalous scatter. The two types of data collection are, multiple anomalous diffraction (MAD), in which three data sets are collected, one at the

maxima absorption edge, and one above and below. The other is single anomalous diffraction (SAD) where only one data set is collected, typically at the absorption peak.

6.1.4.3 Molecular replacement

Molecular replacement is a method that can be utilised to solve the structure when there is a model that shows a large similarity to the structure within the crystal. As a general rule the similarity shown between two protein structures correspond with the likeness of their amino acid sequence. The Protein Data Bank (www.pdb.org, 2008) consists of over fifty thousands protein structures, thus if a protein structure exists that relates to the crystallised protein, then it is possible to use the molecular replacement technique. With the use of modern software programs molecular replacement can be a relatively straightforward process, when the sequence identity between the crystallised protein and model protein is 30 % or greater. Although molecular replacement is possible for lesser sequence identities, the task is significantly harder.

Molecular replacement in principle utilise rotation matrices and translation vectors to superimpose the model onto the target structure (Fig. 6.3).

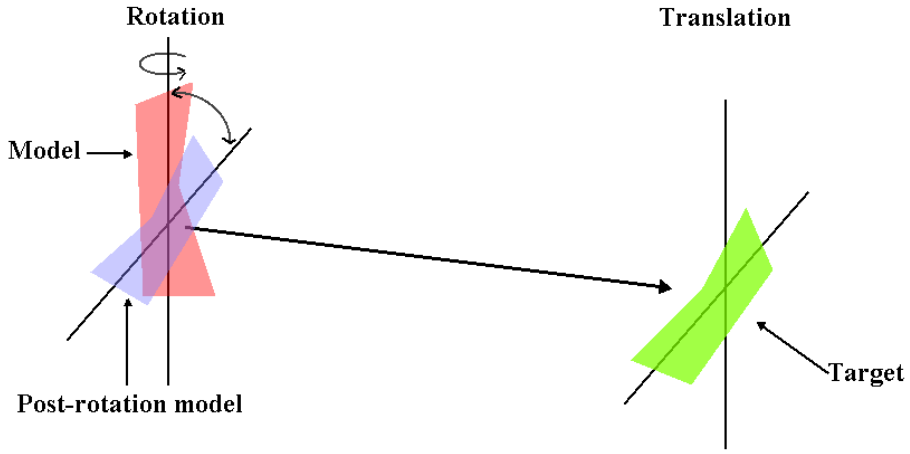


Figure 6.3 Representation of the main principles of molecular replacement method. The Model molecule (red) goes through a rotation operation to reach Post-rotation model (blue), then through the translation operation to reach the Target (green).

Molecular replacement relied upon the Patterson function that calculates an interatomic vector map (Patterson map) created by squaring the structure factor amplitudes and setting all phases to zero (Patterson, 1934). Patterson maps can be calculated for the model and compared to the observed Patterson map. The model Patterson map will look similar to the observed Patterson map when it is in the correct orientation and position

within the unit cell. This approach gives rise to a search in six dimensions for each molecule in the asymmetric unit, three angles to specify the molecule orientation and the three coordinates to specify the position of its centre of mass. Fortunately as the shape of most protein molecules does not deviate much from spherical most intramolecular vectors are located closer to origin of the Patterson synthesis while most intermolecular vectors are further away from the origin. Therefore the latter can be excluded from the rotation function calculations by a choice of appropriate integration radius. Intramolecular vectors relate only to the orientation of the molecule within the unit cell, and not the position, thus can be exploited using the rotation function. The intermolecular vectors depend on both the orientation and the position of the molecule, therefore once the orientation is known these can be exploited in the translation function.

The rotation function allows for the determination of the orientation of the target molecule in the absence of any phases of the unknown structure. This is achieved by comparing the Patterson map intramolecular vectors of the target molecule, with that of the known structure (search model) at different orientations (Fig. 6.4).

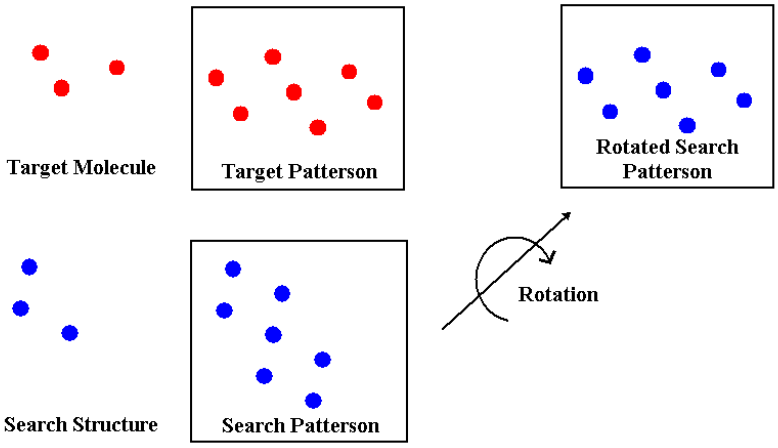


Figure 6.4 Representation of the rotation function. The search Patterson map (blue) goes through a rotation operation to reach Post-rotation model (blue), which is similar to the target Patterson (red).

The agreement between the search and target Patterson maps was quantified by a function proposed by Rossmann and Blow (Rossmann and Blow, 1962), however the integration calculation in real space required a large degree of computing (Eqn.6.2a). Rossmann and Blow again proposed a reciprocal space method for calculating R that was faster, with the integral over volume being simplified to spherical model to equation 6.2b.

$$R = \int P_T(\underline{u})\, P_s([C]\underline{u})\, d\underline{u} \quad \text{(a)}$$

$$R = \Sigma_{\underline{h}} \Sigma_{\underline{p}} |F_{\underline{h}}|^2 \; |F_{\underline{p}}|^2 \, G_{\underline{h},\underline{p}} \quad \text{(b)}$$

Equation 6.2 (a) Rotation function proposed by Rossmann and Blow, where $P_T(\underline{u})$ is the Patterson of the target structure, $P_S([C]\underline{u})$ is the Patterson of the search molecule which has been rotated by matrix [C]. (b) Rotation function simplified to the spherical model, where $F_{\underline{h}}$ is structure factor for target structure, and $F_{\underline{p}}$ is the structure factor for the search structure, and $G_{\underline{h},\underline{p}}$ is an interference function.

The fast rotation function was developed by Crowther (Crowther, 1972), and works by approximating spherical harmonics from the Patterson, with the rotation function being calculated using a fast fourier transform from the spherical harmonics. Programs like AMORE (Navaza, 1994) and MOLREP (Vagin and Teplyakov, 1997) use this method, and calculation of structures even when there is more than one molecule in the asymmetric unit. When calculating the rotation function, consideration must be given to the radius of integration, the resolution limits, and the angular variables. These considerations along with the concept of molecular replacement have been well review by Rossmann (Rossmann, 1990).

As discussed previously the translation function relies on the intermolecular vectors from both the orientation and position of the molecule. Having calculated the three dimensions (α, β, γ) to specify the orientation, the rotation matrix [C] may be calculated and applied to the co-ordinates of the search model. The translation function measures the overlap of the target Patterson map with the Patterson map generated by rotation of a search model by an angle found from the rotation function and positioned at a given translation in the target unit cell. (Fig. 6.5). The agreement between the two Patterson maps is quantified by the translation function and is defined in equation 6.3. This equation was expanded in reciprocal space by Crowther and Blow (Crowther and Blow, 1967).

$$T(\underline{t}) = \int P_c(\underline{u},\underline{t})\, P_o(\underline{u})\, d\underline{u}$$

Equation 6.3 The translation function, where $P_c(\underline{u},\underline{t})$ is the search molecule cross-Patterson and $P_o(\underline{u})$ is the observed Patterson function of the target structure.

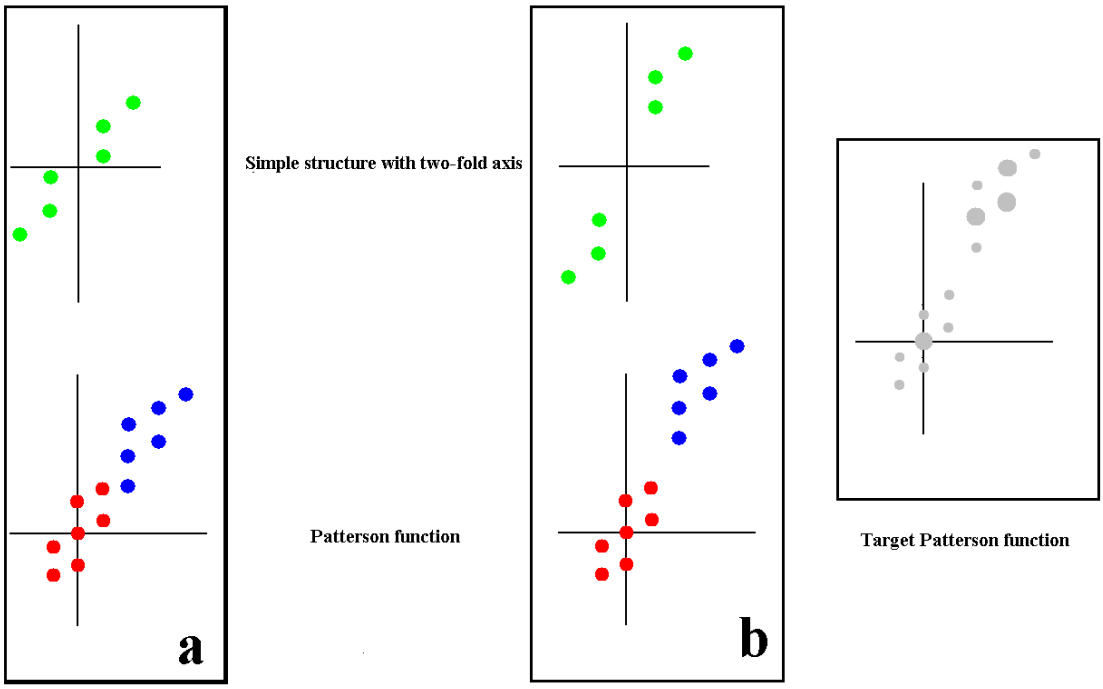


Figure 6.5 Representation of the translation function for a simple structure with two-fold axis. Two correctly rotated model structures (green), with their respective Patterson functions (self-vectors (red), cross-vectors (blue)). Only the Patterson function (b) matches the target Patterson function (grey).

6.1.4.4 Protein structure refinement

Once the phases have been determined by one of the methods described previously, an electron density map can be formed into which the model structure can start to be assembled. A computational model-building program typically produces this initial model structure. The model the software produces holds a premature resemblance to the final protein structure, but contains limited content and structural detail. Comparisons of the electron density map of the model and the initial electron density map allows a difference map to be produced. The difference map allows identification of areas of negative density (areas where the model is incorrect), and positive density (areas where there is non-representation by the model). Once the structural model has been built, refinement of the model can take place until the model best represents the initial X-ray diffraction data. During the process of refinement a minimal agreement must be reached between the proposed structure factors and the measured structure factors, this is commonly performed using an automated refinement program and a visualisation interface. Refinement is carried out using the automated refinement

program, the model is then checked manually using the visualisation interface where the user corrects, builds and repositions atoms before resubmitting the model for further refinement. This process is continually repeated until all structural details of the model have been determined. The difference between the two structure factors (proposed and measured) is given by the *R*-factor, this is a representation of 'agreement' between the proposed electron density (from the model structure) and the actual electron density (from the X-ray diffraction data) if they are identical the *R*-factor is zero, as there is complete equivalence between the two. The other measure of agreement is Free *R*-factor (R_{free}), this appraisal uses a small percentage of the rejected reflections from the refinement, and this gives an unbiased overview of the model (Brünger, 1997). Further to the final refinement of the model other computer based validation programs are used to verify the models suitability in comparison with other structures previously solved at a given resolution. These programs check parameters like bond angle and steric orientation to confirm the model is chemically acceptable. A Ramachandran plot should also be obtained for the model to ensure the conformational angles of the polypeptide main chain are of good concordance with in the Ramachandran plot regions (Ramakrishnan and Ramachandran, 1965).

6.2 Materials and methods

6.2.1 Mother liquor and cryogenic liquor

A mother liquor containing 50 mM PIPES, 2mM TPP, 9 mM $CaCl_2$, pH 6.4 55 % saturated ammonium sulfate was prepared to stabilise the crystal once it was moved from the crystallisation drop. A cryo liquor containing 50 mM PIPES, 2mM TPP, 9mM $CaCl_2$, pH 6.4 55 % saturated ammonium sulfate and 30 % (v/v) glycerol was prepared to protect the crystal whilst being frozen and stored in a liquid nitrogen dewar prior to data collection.

6.2.2 Crystal soaks

Crystals were soaked in mother liquor (prepared in section 6.2.1), containing varied concentrations of fluoropyruvate (750 mM) or hydroxypyruvate (40 mM) for time intervals between 30 sec and 24 hrs. The conditions of the crystal were monitored using a microscope. After the soaking experiments the crystals were collected using a

mounted nylon loop and placed in cryogenic liquor and frozen using liquid nitrogen prior to X-ray diffraction experiments.

6.2.3 'In-house' data collection

The frozen mounted crystals in cryogenic liquor were removed from the dewar and placed upon the X-ray goniometer head (in the case of a soaking experiment, 'soaks' were conducted prior to being mounted upon the X-ray). The 'in-house' data was collected using a rotating anode generator (Bruker AXS) operated at 100 mA and 35 KV (3.5 kW) producing copper Kα radiation with XENOCS FOX2D CU25_25P mirrors. Data was collected under cryocooled conditions (100 K in a stream of gaseous nitrogen) at varying distance and rotation angle (dependant on data set) using a MAR Research 345 Image Plate.

6.2.4 High resolution data collection

The frozen mounted crystals in cryogenic liquor were removed from the dewar and placed on the X-ray goniometer head (in the case of a soaking experiment, 'soaks' were conducted prior to being mounted upon the X-ray). Data was collected at the Daresbury synchrotron radiation source (Daresbury SRS, UK) on beamline 14.2 under cryocooled conditions (100 K in a stream of liquid nitrogen) using a ADSC Q4R CCD detector. Data was also collected on beamline 14.1 under cryocooled conditions (100 K in a stream of gaseous nitrogen) using a Quantum 4 ADSC detector.

6.2.5 Data processing

Data frames were indexed and scaled using DENZO (Otwinowski and Minor, 2007), and SCALEPACK (Otwinowski and Minor, 2007) respectively.

6.3 Results and discussion

6.3.1 *E. coli* transketolase

The TK crystals were found to behave in a stable manner in both the mother and the cryogenic liquors, and they diffracted well without any signs of radiation damage. A TK crystal was soaked prior to data collection in mother liquor containing 40 mM hydroxypyruvate for 1 min. The soaked crystal did not crack, and was subjected to data collection. Another TK crystal was soaked in 750 mM fluoropyruvate for 10 mins. The crystal started to crack after 4 mins but retained clean edges making the crystal suitable for data collection.

6.3.1.1 'In house' data

The TK crystal soaked in 750 mM fluoropyruvate (TK-FPA) diffracted to 1.60 Å resolution 'in house'. Indexing and scaling revealed the crystal belonged to the orthorhombic space group $P2_12_12_1$ with unit cell parameters a = 73.5 Å, b = 124.7 Å, c= 149.9 Å, $\alpha=\beta=\gamma=90.0°$. 180433 unique reflections were observed and data set had 99.1 % completeness. The statistics for this data is shown in table 6.2

6.3.1.2 High resolution data

Two high resolution data set were collected, the native TK crystal and the TK crystal soaked in 40 mM hydroxypyruvate (TK-HPA). The native TK crystal diffracted to 1.18 Å resolution (Fig. 6.7), indexing and scaling of this data suggested that the crystal belonged to orthorhombic space group $P2_12_12_1$ with unit cell parameters a = 73.5 Å, b = 124.5 Å, c= 149.6 Å, $\alpha=\beta=\gamma=90.0°$. 438271 unique reflections were observed and data set had 97.8 % completeness. The TK crystal soaked in 40 mM hydroxypyruvate diffracted to 1.05 Å resolution, and was found to belong to the same with unit cell parameters a = 73.2 Å, b = 124.3 Å, c= 149.5 Å, $\alpha=\beta=\gamma=90.0°$. 603123 unique reflections were observed and data set had 95.8 % completeness. The statistics for these data sets are shown in Table 6.2. In all three data sets the asymmetric unit contained two transketolase monomers. The solvent content of the crystals has been estimated at 49.9 % (v/v), V_M = 2.45 $Å^3$/Da (Matthews, 1968)

	TK	TK-HPA	TK-FPA
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit Cell (a,b,c Å)	73.5, 124.5, 149.6	73.2, 124.3, 149.5	73.5, 124.7, 149.9
Wavelength	0.978	0.978	1.542
Resolution range (Å)	25 - 1.18 (1.20 - 1.18)	25 - 1.05 (1.07 - 1.05)	15 - 1.60 (1.63 - 1.60)
Number of unique reflections	438271	603123	180433
Completeness (%)	97.8 (95.9)	95.8 (68.6)	99.1 (99.2)
Redundancy	5.2 (4.9)	4.1 (1.7)	3.8 (3.5)
B-Factor of data from Wilson Plot ($Å^2$)	12	12	30
$I > 3\sigma(I)$ (%)	75.8 (52.8)	70.5 (13.7)	77.2 (38.7)
$(I)/\sigma(I)$	20.2 (4.37)	16.8 (1.6)	15.8 (2.8)
R_{sym} (%)	10.5 (54.4)	7.1 (44.5)	7.6 (52.2)

Table 6.2 Data processing statistics for the three TK data sets ($R_{sym} = \Sigma_h\Sigma_J|<I_h>-I_J(h)|/\Sigma_h\Sigma_J I(h)$. Where (h) is the intensity of the reflection h. Σ_h is the sum over all reflections and Σ_J is the sum over J measurements of the reflection. Values given in the parentheses are given for the highest shell resolution.

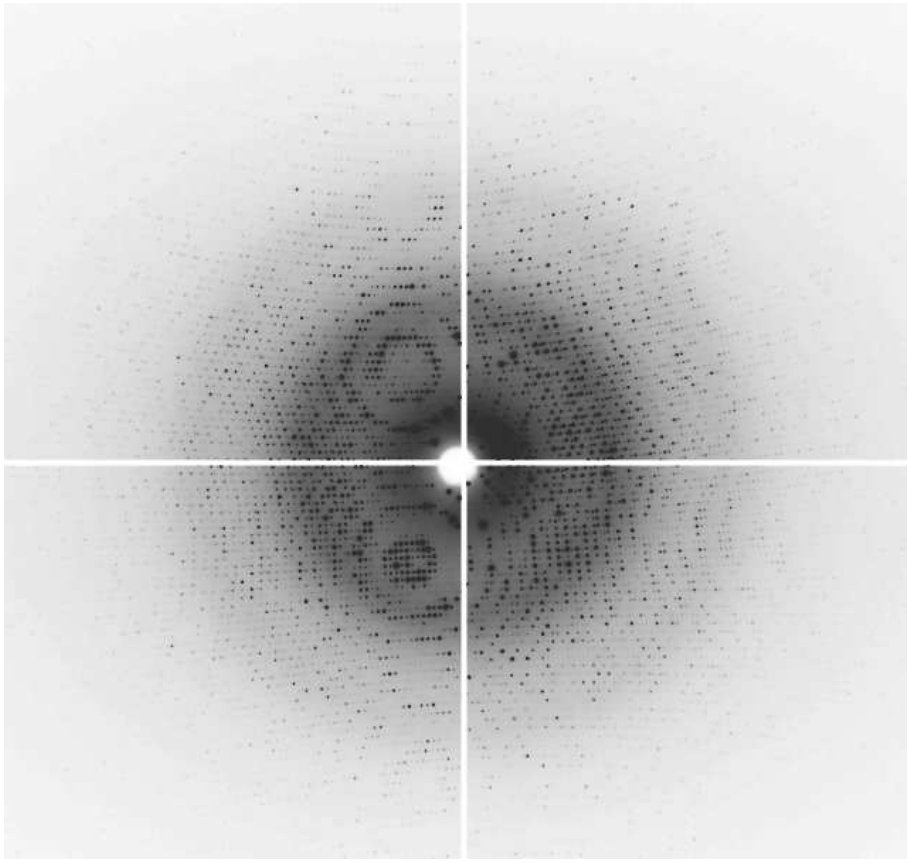


Figure 6.7 A frame of data collected at the Daresbury synchrotron beamline 14.1 from the native TK crystal, which diffracted to 1.18 Å.

6.3.2 E. *coli* transketolase mutant H26Y

The H26Y crystals, as with the native TK were found to behave in a stable manner in both the mother and the cryogenic liquors and diffracted well without any signs of radiation damage. A H26Y crystal was soaked prior to data collection in mother liquor containing 750 mM fluoropyruvate for 10 mins. The soaked crystal did crack but retained clean edges and was subjected to data collection

6.3.2.1 'In house' data

The H26Y crystal soaked in 750 mM fluoropyruvate (TK-FPA) diffracted to 1.66 Å resolution 'in house'. Indexing and scaling revealed the crystal belonged to the same orthorhombic space group $P2_12_12_1$ as the native TK, with unit cell parameters a = 73.2 Å, b = 124.4 Å, c= 149.6 Å, $\alpha=\beta=\gamma=90.0°$. 158322 unique reflections were observed and data set had 98.1 % completeness. The statistics for this data set is shown in Table 6.3. The asymmetric unit as with the native TK contained two transketolase

H26Y mutant monomers. The solvent content of the crystals has been estimated at 49.5 % (v/v), V_M = 2.43 $Å^3$/Da (Matthews, 1968)

	H26Y-FPA
Space group	$P2_12_12_1$
Unit Cell (a,b,c Å)	73.2, 124.4, 149.6
Wavelength	1.542
Resolution range (Å)	15 – 1.66 (1.69 - 1.66
Number of unique reflections	158322
Completeness	98.1 (95.4)
Redundancy	2.7 (2.2)
B-Factor of data from Wilson Plot ($Å^2$)	26
I > 3σ(I) (%)	81.0 (52.8)
(I)/σ (I)	19.4 (4.5)
R_{sym} (%)	4.2 (22.9)

Table 6.3 Data processing statistics for the H26Y-FPA data set ($R_{sym} = \Sigma_h\Sigma_J|<I_h>-I_J(h)|/\Sigma_h\Sigma_J I(h)$. Where (h) is the intensity of the reflection h. Σ_h is the sum over all reflections and Σ_J is the sum over J measurements of the reflection. Values given in the parentheses are given for the highest shell resolution.

6.4 Structure solution and refinement

The four data sets collected, three TK data sets (TK, TK-HPA and TK-FPA) and the variant H26Y data sets (H26Y-FPA), were then processed and refined as described below.

6.4.1 Phase determination

Molecular replacement was performed using the programme MOLREP (Vagin and Teplyakov, 1997) for phase determination. The structure of transketolase from *E. coli* (PDB code 1QGD) was used as a model.

6.4.2 Model building and refinement

The electron density map produced from the calculated phases was used to build the transketolase models. The programme REFMAC 5.2 (Murshudov *et al.*, 1997) and SHELX (Sheldrick, 2008) was used for refinement and manual model building was performed in COOT (Emsley and Cowtan, 2004). After each session of introducing changes by model building, each refinement was 10 rounds. Following the model building and refinement the positioning of amino acid side chains was possible. During the final stages of model building any co-factors and substrate present were built in the sketcher interface of the program LIBCHECK (CCP4, 1994), and modelled into the electron density map. Also any observed amino acid secondary positions were built. Solvent molecules were added and checked manually.

6.4.3 Model building of the inhibitor and substrate

The two additional ligands (FPA and TPP-HPA) were built using the CCP4i SKETCHER program (CCP4, 1994), and positioned into the additional electron density map using COOT (Emsley and Cowtan, 2004). The models were then further refined using REFMAC 5.2 (Murshudov *et al.*, 1997) for the TK-FPA and H26Y-FPA structures, and SHELX (Sheldrick, 2008) for the TK-HPA structure as in section 6.3.2. Refinement was continued until the substrate complexes found coordinates in agreement with the surrounding B-factors of neighbouring residues and atoms.

6.4.4 Structure validation

Once refinement was complete, the program PROCHECK (Laskowski *et al.*, 1993) was used to check the quality of the models. The product of this program allowed identification of possible error in the structure, which were corrected and subjected to further refinement. Final validation of the structure was also carried out using PROCHECK (Laskowski *et al.*, 1993)

6.5 Results and discussion

The data collected from both TK and H26Y crystals diffracted to high-resolution limits, and in the case of the native TK and TK-HPA to resolution limits of 1.18 and 1.05Å respectively. The results of TK and H26Y with their complexes are discussed separately in this section.

6.5.1 E. *coli* transketolase

The native TK structure was solved using molecular replacement with the *E*. coli transketolase 1QGD as the model structure. Automated model building was used to build the majority of the amino acid structure, but a four amino acid region from 138 - 141 on subunit B was required to be built manually due to the disorder in the electron density. This region of disorder had be previously seen in the first *E. coli* transketolase structure by (Littlechild *et al.*, 1995). Subunits A and B comprised of 663 amino acids each, which were all fitted into the electron density map. A succinic acid residue was fitted at position 157 on both subunits, due to the evidence from the electron density map. The final model consisted of 1326 amino acids residues (26 additional amino acid residue secondary positions), 2 TPP co-factors, 2 calcium ions, 6 sulfate molecules, 17 glycerol molecules and 2324 water molecules. The final refined structure had an R-factor of 12.9 and R_{FREE} of 15.7. A full list of the data statistics can be found in Table 6.4. PROCHECK (Laskowski *et al.*, 1993) calculated the G-factor of the model to be 0.0, which is better than expected for a structure refined at this resolution. The Ramachandran plot (Ramakrishnan and Ramachandran, 1965) for the structure is viewable in Appendix IIIa.

The TK structure in complex with HPA was solved using the native TK model. The TPP- HPA molecule was positioned using the vacant density omitted from the electron density map. Again no changes in active site amino acid residues were observed. The final model consisted of 1346 IIIb amino acid residues (20 additional amino acid residue secondary positions), 2 TPP-HPA co-factors, 2 calcium ions, 5 sulfate molecules, 17 glycerol molecules and 1825 water molecules. The final refined structure had an R-factor of 12.4 and R_{FREE} of 14.9 (Table 6.4). The G factor calculated by PROCHECK was 0.0, which is better than expected for a structure refined at this resolution. The Ramachandran plot (Ramakrishnan and Ramachandran, 1965) for the structure is viewable in Appendix IIIb.

The TK structure in complex with the inhibitor FPA was solved using the model from the native TK structure. Refinement of the TK-FPA model allowed the positioning of the FPA molecule within the vacant density omitted from the electron density map. There were no structural alternations of the amino acid residues within the active site. The final model consisted of 1326 amino acid residues (24 additional amino acid residue secondary positions), 2 FPA molecules, 2 TPP co-factors, 2 calcium ions, 4 sulfate molecules, 16 glycerol molecules and 896 water molecules. The final refined structure had an R-factor of 17.2 and R_{FREE} of 19.1 (Table 6.4). The G factor calculated by PROCHECK was 0.1, which is better than expected for a structure refined at this resolution. The Ramachandran plot (Ramakrishnan and Ramachandran, 1965) for the structure is viewable in Appendix IIIc.

	TK	TK-HPA	TK-FPA
Final R_{cryst} (%)	12.9	12.4	17.2
R_{FREE}	15.7	14.9	19.1
Number of Protein Residues	1326	1326	1326
Number of water molecules	2324	1825	896
Average B-factor (whole pdb Å^2)	14.6	14.2	19.4
Average B-factor (Protein Å^2)	11.0	11.9	18.3
Average B-factor (Solvent Å^2)	31.2	29.4	29.2
Average B-factor (Ligand Å^2)	24.2	21.0	37.6
Rms derivations from ideality			
Bond lengths (Å)	0.026 (0.022)	0.021 (0.022)	0.013 (0.022)
Bond angles (°)	2.318 (1.973)	2.378 (1.971)	1.413 (1.970)
Planes (Å)	0.015 (0.021)	0.019 (0.020)	0.006 (0.020)

Table 6.4 Refinement statistics for the three TK structure models. B-factor = $8\pi^2U^2$, where U^2 is the mean square displacement of any given atom. The target values are given in parentheses.

6.5.2 E. *coli* transketolase mutant H26Y

The TK mutant H26Y structure in complex with FPA was solved in the same manner as the TK structure using molecular replacement with the *E. coli* transketolase 1QGD as the model structure. The electron density allowed the positioning of the 663 amino acids residues in each of the two monomers, including the replacement of the histidine with a tyrosine at position 26. The FPA molecule was positioned using the vacant density omitted from the electron density map. Similar disorder was observed as in the TK structures with residues 138 to 141 in subunit B. However unlike the TK structures there was some disorder in the residues 257 to 261. These had to be built manually using COOT (Emsley and Cowtan, 2004), giving possible reasoning for the alternative configuration of the H26Y mutants reaction products. The final model consisted of 1326 amino acid residues (28 additional amino acid residue secondary positions), 2 FPA molecules, 2 TPP co-factors, 2 calcium ions, 6 sulfate molecules, 19 glycerol molecules and 1159 water molecules. The final refined structure had an R-

factor of 14.6 and R_{FREE} of 17.5, (Table 6.5). The G factor calculated by PROCHECK was 0.1, which is better than expected for a structure refined at this resolution. The Ramachandran plot (Ramakrishnan and Ramachandran, 1965) for the structure is viewable in Appendix IIId

	H26Y-FPA
Final R_{cryst} (%)	14.6
R_{FREE}	17.5
Number of Protein Residues	1326
Number of water molecules	1159
Average B-factor (whole pdb Å^2)	16.1
Average B-factor (Protein Å^2)	14.5
Average B-factor (Solvent Å^2)	28.3
Average B-factor (Ligand Å^2)	31.1
Rms derivations from ideality	
Bond lengths (Å)	0.010 (0.022)
Bond angles (°)	1.22 (1.97)
Planes (Å)	0.005 (0.02)

Table 6.5 Refinement statistics for the H26Y-FPA structure model B-factor = $8\pi^2U^2$, where U^2 is the mean square displacement of any given atom. The target values are given in parentheses.

Chapter 7

Structure of the *E. coli* transketolase

7.1 Monomer structure

The *E. coli* transketolase monomer subunit is 'bean' shaped with approximate dimensions of 95 x 55 x 45 Å. The monomer is split into 3 domains of α/β type, the PP-domain consists of residues 1-317, the Pyr-domain consists of residues 320-520, and the, C-terminal domain consists of residues 530-663 (Fig. 7.4b).

The secondary structure of the PP domain has a five-stranded parallel β-sheet (β1, β2, β3, β6 and β7) with -1x, -1x, -2x, +1x topology (Richardson, 1981). The PP domain β-sheet separates the 3 α-helices interface with the second domain (Pyr-domain), whilst the other side of the β-sheet is covered by 8 α-helices with a β-hairpin formed by β4 and β5 with residues 187-188 and 191-192 respectively (Fig. 7.4a). The secondary structure of the Pyr-domain has a six-stranded parallel β-sheet (β9, β10, β11, β12, β13 and β14) with -1x, +2x, +1x, +2x, -1x topology (Richardson, 1981), that dissects a 4 α-helices adjacent to the PP domain and the 5 α-helices on the opposite side. A small antiparallel β-sheet is formed by β8 and β15 strands (Fig. 7.4a). The PP and Pyr domains are involved in the binding of the TPP cofactor. The TPP binding fold is consistent with other related TPP-dependant enzyme structures (Muller *et al.*, 1993). The C-terminal domain has a five-stranded β-sheet (β16, β17, β18, β19 and β20) with β17, β18, β19 and β20 running in parallel and β16 antiparallel forming a -2, +1x, -2x, -1x topology. The β-sheet divides the 4 α-helices on the interface of the Pyr domain with 3 α-helices closes to the C-terminus (Fig. 7.4a). The overall fold of the monomer can be seen in figure 7.1, and the overall secondary structure of the monomer displayed in figure 7.2, 7.3 and 7.4. The monomers present in subunit A and B have a solvent accessible area of 24683 and 24476 $Å^2$ respectively.

Figure 7.1 The overall fold of the *E. coli* transketolase monomer displayed as a Cα trace, with every tenth residue numbered. This divergent stereo picture was created using PYMOL (DeLano, 2006)

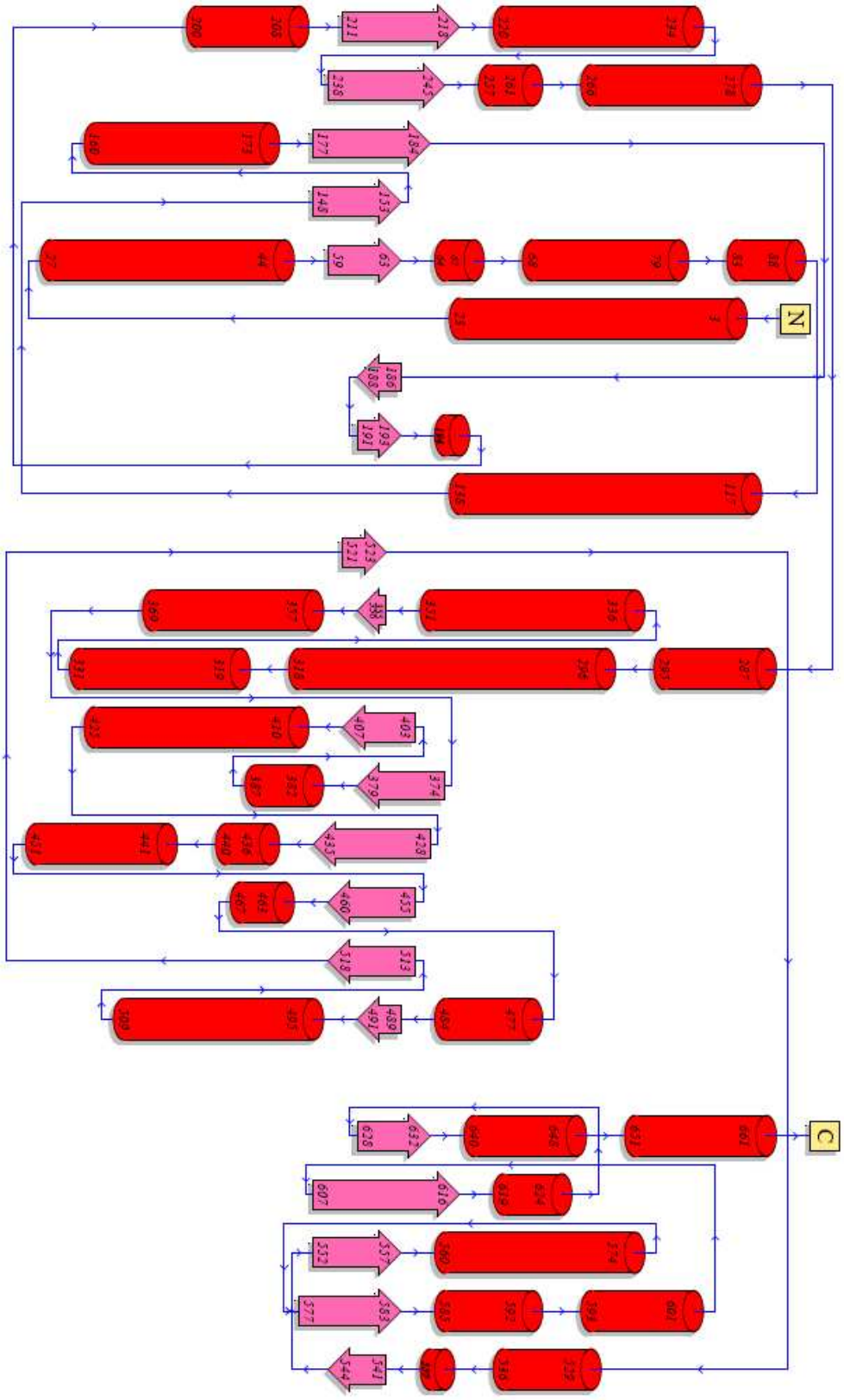


Figure 7.2 Topology diagram showing the secondary structure assignment. Created using pdbsum on the EBI server (www.ebi.ac.uk/pdbsum/, 2008).

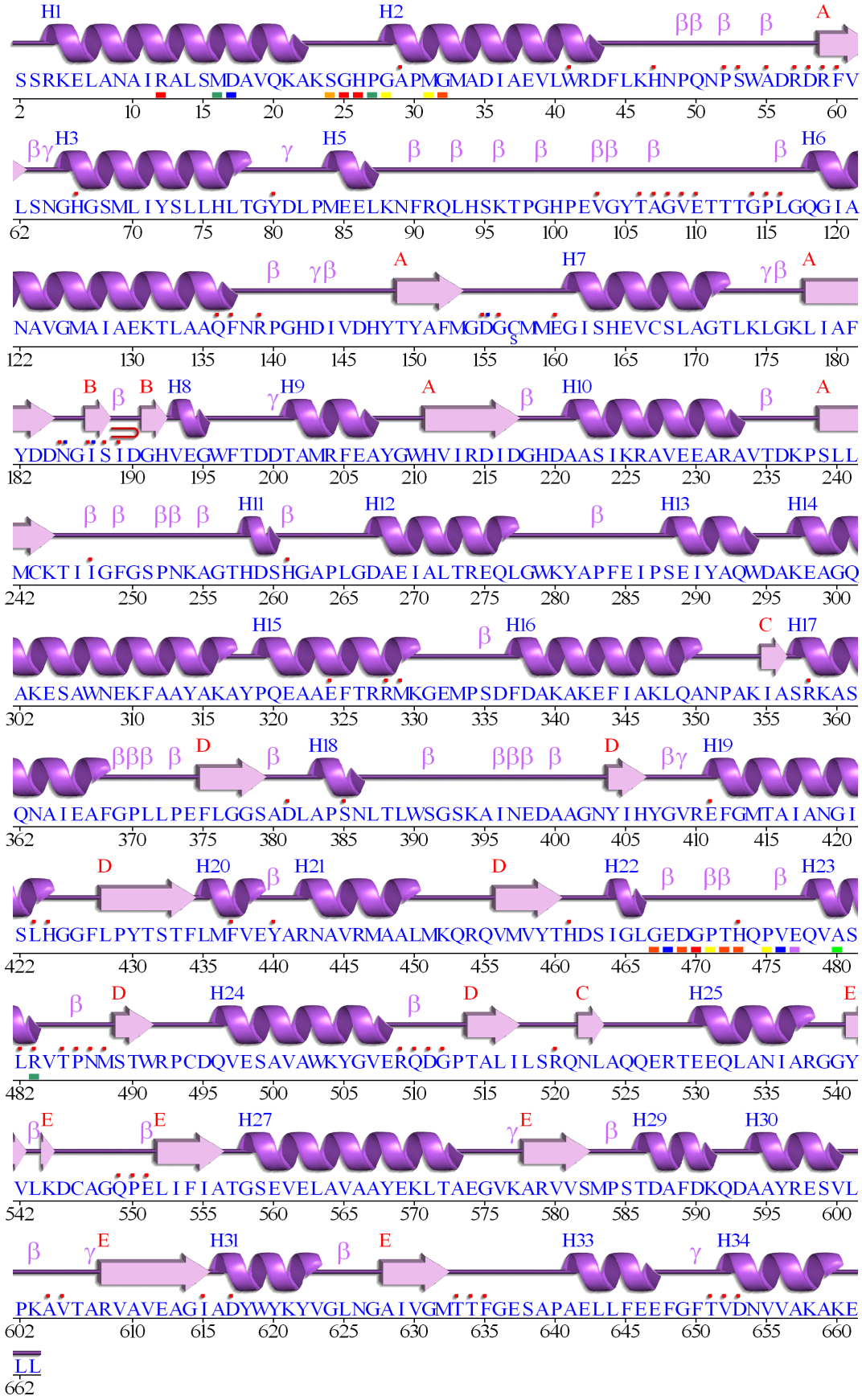
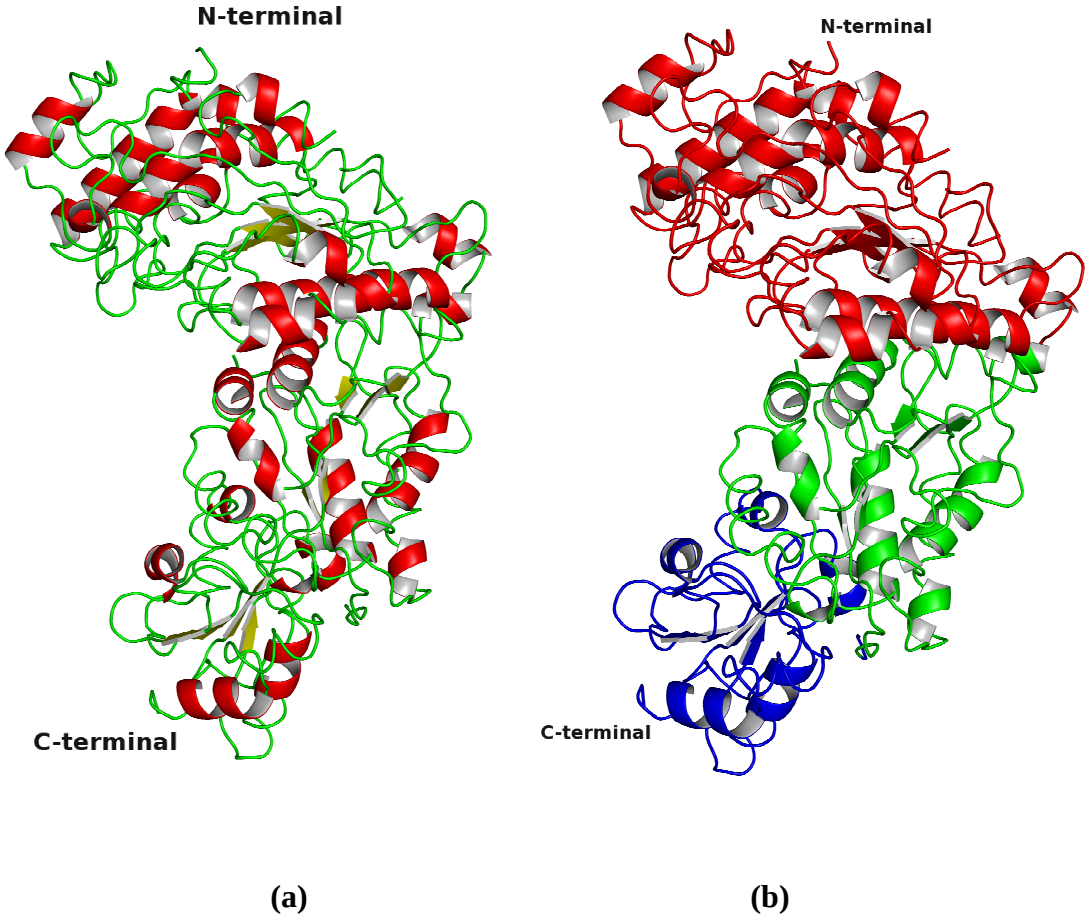


Figure 7.3 Assignment of secondary structure of the monomer. Generated using procheck EBI on the server (www.ebi.ac.uk/pdbsum/, 2008).



(a) **(b)**

Figure 7.4 The secondary structure diagram of the *E. coli* transketolase monomer. The secondary structure elements are displayed in (a) by α-helices (red), β-sheets (yellow) and loops (green). The secondary structure domains are displayed in (b) by PP-domain (red), Pyr-domain (green) and C-domain (blue). Figures were created using PYMOL (DeLano, 2006)

The three α/β folded domains that compose the monomer subunit of the *E. coli* transketolase structure, are consistent with that of the general transketolase family. Both the PP and Pyr domains participate in the formation of the dimer through binding at the dimer interface. Both of these domains are also involved in the binding of the TPP cofactor. The biological function of the C-domain is still unresolved, as no residue within the domain is involved in formation of the active dimer or binding of the TPP (Schneider and Lindqvist, 1998). To date there are five transketolase enzyme structures, Yeast, *E. coli*, Maize, *L. mexicana*, and *T. thermophilus*. (Nikkola *et al.*, 1994; Littlechild *et al.*, 1995; Gerhardt *et al.*, 2003; Veitch *et al.*, 2004; Yoshida *et al.*, 2007). All of which exhibit the three α/β folded domains.

The yeast transketolase structure of 680 amino acids was solved at a 2.5 Å resolution, and has a 42.4 % sequence identity with the *E. coli* enzyme. The PP-domain of the yeast structure (1TRK) contains a five-stranded β-sheet with six α –helices above and seven α –helices below. The Pyr-domain contains a six-stranded β-sheet with very similar α/β-units to the PP-domain. The C-domain contains a five-stranded β-sheet with one anti-parallel strand.

The maize transketolase structure of 666 amino acids was solved at a 2.3 Å resolution, and has a 46.7 % sequence identity with the *E. coli* enzyme. The maize structure (1ITZ) comprises of exactly the same α/β-units in all three domains as the yeast and *E. coli* structures.

The *L. mexicana* transketolase structure of 669 amino acids was solved at a 2.22 Å resolution, and has a 43.9 % sequence identity with the *E. coli* enzyme. The L. *mexicana* structure (1R9J) again has the same α/β-units in all three domains as the yeast, *E. coli* and maize structures.

The *T. thermophilus* transketolase structure of 651 amino acids was solved at a 2.09 Å resolution, and has a 46.7 % sequence identity with the E. *coli* enzyme. The *T. thermophilus* structure (2E6K) as before has the same α/β-units in all three domains as the yeast, E. *coli*, maize, and L. *mexicana* structures.

The four transketolase structures were superimposed with the Cα atoms of the *E. coli* transketolase structure using PYMOL (DeLano, 2006). Superimposition of the 1TRK (yeast) calculated an RMS deviation value of 0.809 Å over 574 Cα atoms (Fig. 7.6). Superimposition of the 1ITZ (Maize) calculated an RMS of 0.770 Å over 583 Cα atoms (Fig. 7.6). Superimposition of the 1R9J (L. *mexicana*) calculated an RMS of 0.769 Å over 520 Cα atoms (Fig. 7.6). Superimposition of the 2E6K (T. *thermophilus*)

calculated an RMS of 0.715 Å over 567 Cα atoms (Fig. 7.6). The five enzymes are very similar in structure, which is shown by their very small RMS Figures. This implies that the transketolase structure is very well conserved throughout plants, bacteria, protozoa, and yeast. Sequence alignment of the five transketolases maybe seen in figure. 7.5 The survival of the C-domain (that has no obvious biological importance) through evolution, must give some evidence to the case of the domains involvement with other binding sites for regulatory molecules (Schneider and Lindqvist, 1998). The only slightly noticeable differences between the four structures can be seen on study of the loops connecting the α/β-units, and even then these differences are very slight. It would be expected that the transketolases isolated from *T. thermophilus* would have shorter loops to increase the proteins stability, due to the nature of its environment. In comparison between the *E. coli* TK and the *T. thermophilus* TK it maybe noted that the later contains shorter loops in some regions, the most notable loop in the *T. thermophilus* TK structure is from Ala-517 to Lys-529, with the loop consisting of an increased number of proline residues.

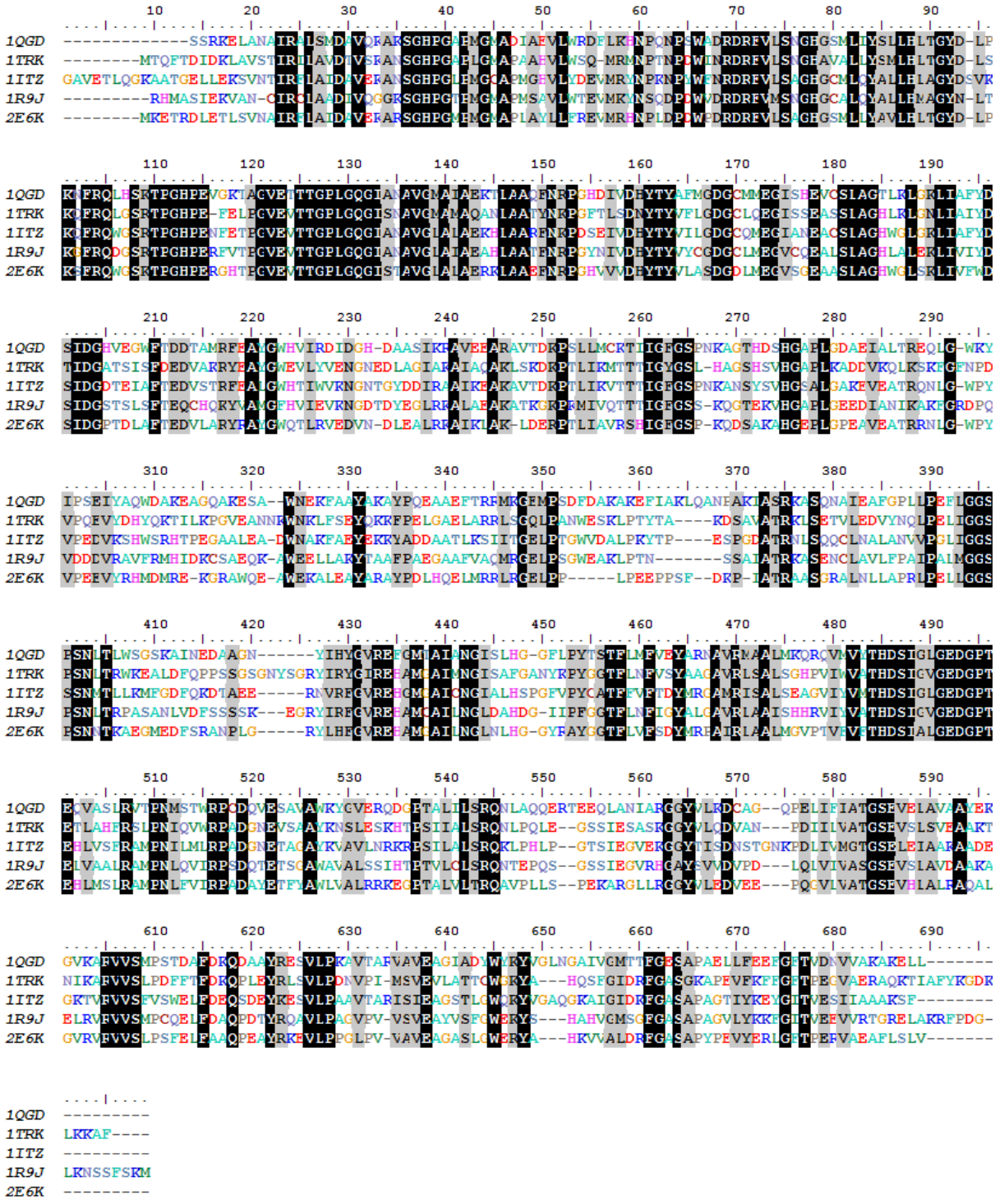


Figure 7.5 Sequence alignment of Yeast (1TRK), *E. coli* (1QGD), Maize (1ITZ), *L. mexicana* (1R9J), and *T. thermophilus* (2E6K)

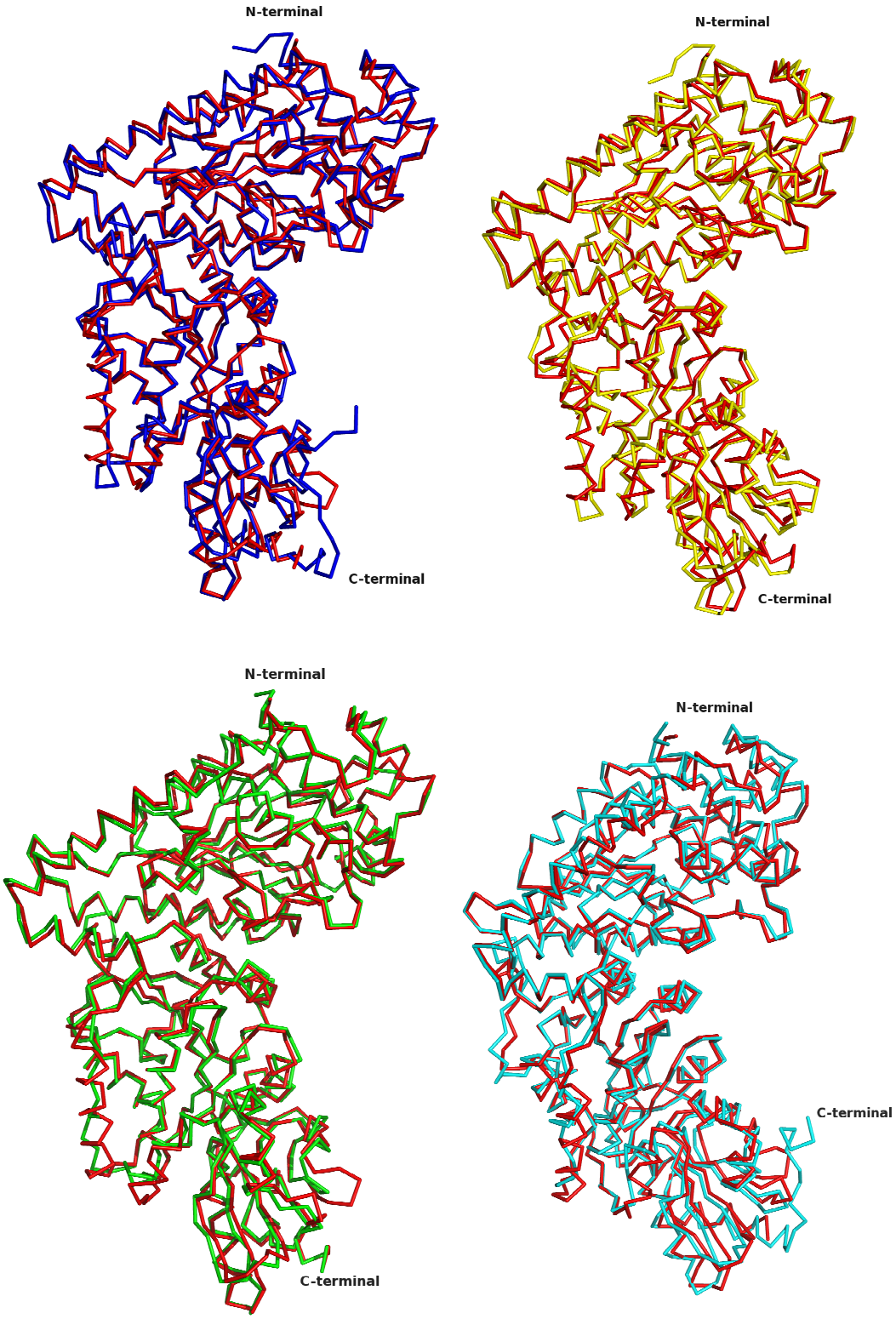


Figure 7.6 Superimposition of the 1TRK (yeast) (blue), 1ITZ (maize) (yellow), 1R9J (*L. mexicana*) (cyan) and 2E6K (*T. thermophilus*) (green) with the *E. coli* transketolase structure (red) using PYMOL (DeLano, 2006).

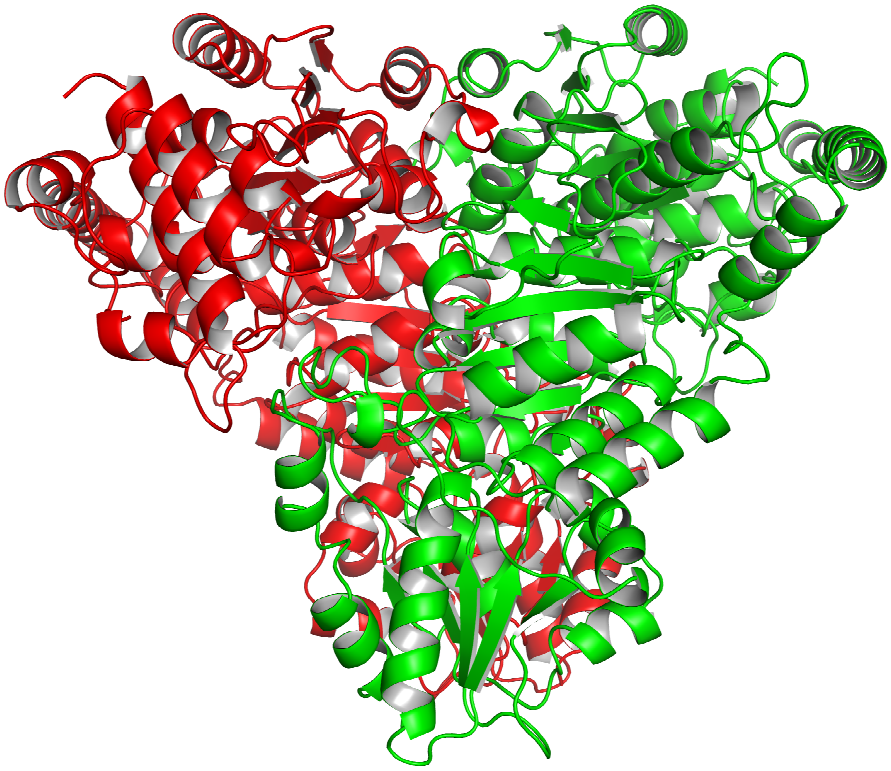
7.2 Oligomeric structure

The *E. coli* transketolase exists as a homodimer (Fig. 7.7). The two subunits show a two-fold symmetry axis, and superimposition of one subunit with the other gives an RMS deviation of 0.188 Å for 553 Cα atoms. The homodimer present in the *E. coli* structure is of the same oligomeric state present in the other four transketolase structures, with RMS values for superimposition of their subunits similar to that of the *E. coli* structure.

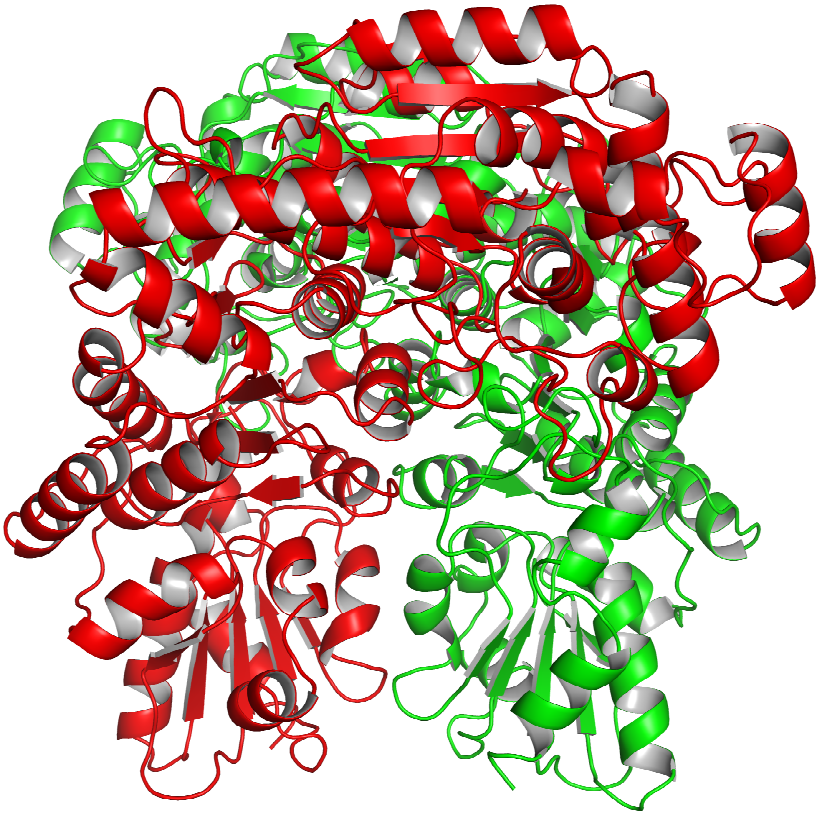
The overall shape of the *E. coli* transketolase dimer can be represented as a trigonal prism, with approximate dimensions of 95 x 90 x 85 Å. The dimer has a solvent accessible surface area 40675.8 $Å^2$. Formation of the dimer buries 8482.8 $Å^2$ equating to 17.1 % of the solvent accessible surface area of each monomer.

The *E. coli* transketolase homodimer is formed through a series of interactions between the PP and Pyr domains (Fig. 7.8). The hydrogen bonding interactions between the PP-domain of both subunits occurs between the residues Asp-200 --- Ala-207 of helix 8, and through residues Gly-161 --- Glu-165, Met-158 --- His-164 of helix 7. Residues Ser-168 and Thr-172 of helix 7 also form hydrogen bonds with Asp-199 and Thr-198 respectively. The hydrogen bonding interactions between the PP and Pyr domain are focused on Asp-190, which forms hydrogen bonds with Ala-383, His-406, and Asp-381. The latter Asp-381 also forms hydrogen bonds with Ser-188, Ile-189, and Asp-381. Asp-469 interacts with Arg-91 and His-100, whilst Leu-116 forms a hydrogen bond with His-164. Another interaction between the PP and Pyr domains of note is between Trp-196 and Arg-410. The key contacts within the Pyr-domain involve the interactions of residues Glu-439 --- Asn-443, and Glu477 --- Arg446 of the two subunits. Contacts made between the subunits of the C-domain are non-existent (Fig. 7.8), although it may be seen that hydrogen bonding does occur with the PP-domain, involving residues 638-Ser --- Gly-99 and Glu-647 --- His-94. The TPP cofactor forms salt-bridges and additional hydrogen bonding between the two subunits, thereby increasing the stability of the dimer.

Superimposition of the *E. coli* transketolase homodimer with the other four known transketolase structures, yeast, maize, *L. mexicana* and *T. thermophilus* show a series of similar interaction between the two monomer subunits, yielding a calculated RMS of 1.021, 0.994, 0.885 and 0.815 Å of the Cα backbone respectively.



(a)



(b)

Figure 7.7 The *E. coli* Transketolase dimer show in secondary structure ribbon, subunit A (red) and B (green). (a) View face on (b) 90° rotation. Figures produced in PYMOL (DeLano, 2006).

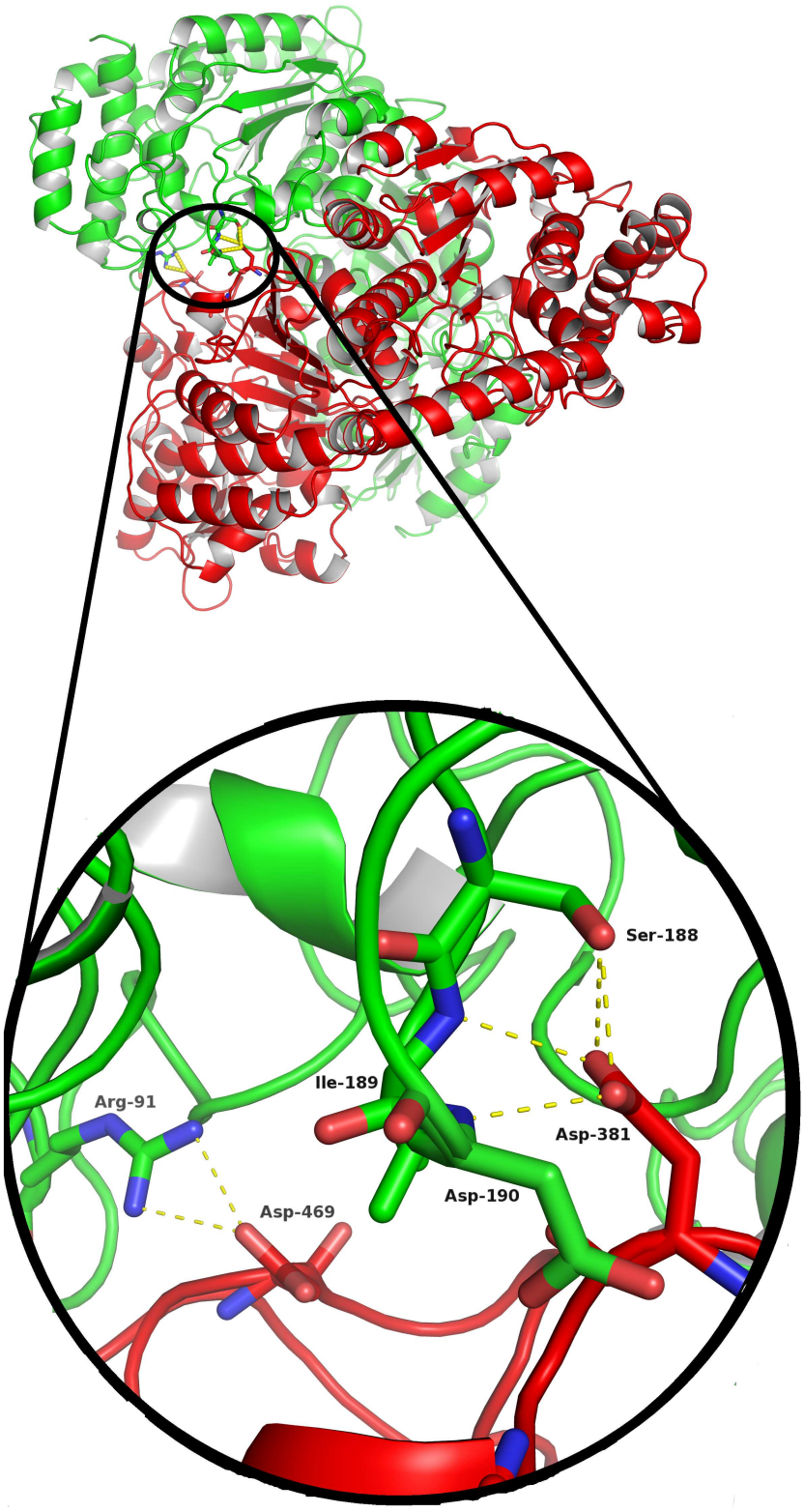
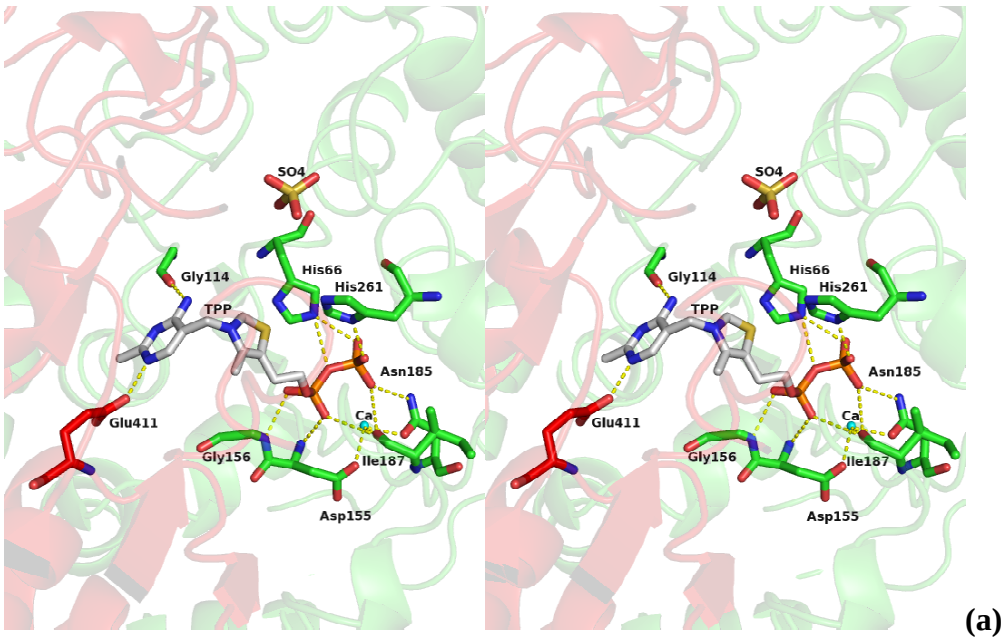


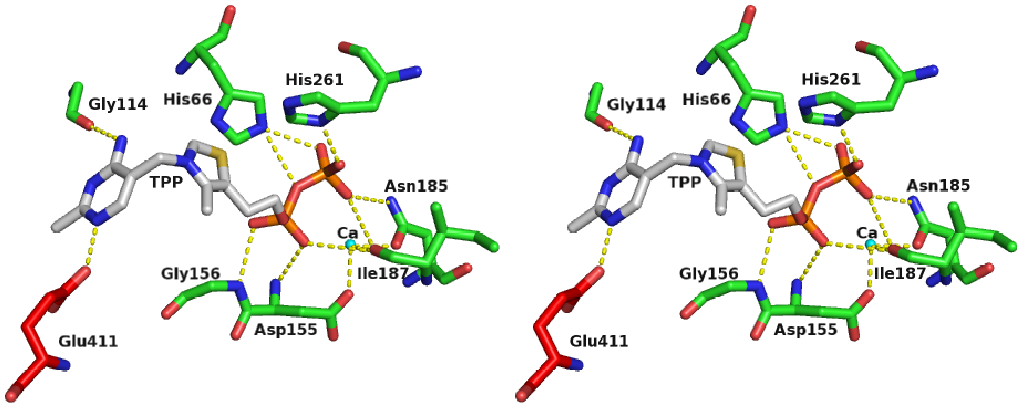
Figure 7.8 The hydrogen-bonding network (yellow) between subunits A (red) and B (green), depicting the holo-dimer interface interactions in the PP and Pyr domains. Enlarged image to focus on the interactions of Asp-381 and Asp469 within the PP and Pyr domains. Figures produced in PYMOL (DeLano, 2006).

7.3 Active site

The two identical active sites of the enzyme are formed by a cleft at the interface of the dimer between the PP and Pyr domains. The thiamine pyrophosphate molecule and calcium ion are bound at the base of cleft and held in place by a series of hydrophobic interactions, hydrogen bonds and salt bridges formed with the active site residues. The TPP molecule adopts the higher energy V-conformation (Fig. 7.9b, Fig. 7.11), which is present in most TTP-dependant protein structures (Hubner *et al.*, 1998; Leeper *et al.*, 2005). The conformation has the 4'-amino group on the pyrimidine ring in close contact (3.0 Å) to the hydrogen of the C2 in the thiazolium ring. The TPP molecule itself has three moieties, amino-pyrimidine ring, thiazolium ring, and the phosphate group. The amino-pyrimidine ring is positioned in a largely hydrophobic area made up of the Phe-437, Val-438, Tyr-440, Ala-441, and Ala-444 residues. Hydrogen bonding interactions of the amino-pyrimidine ring are between 4'NH_2 with the Gly114 residue, and 1'N with Glu-411 residue. With the later amino acid protonating the 1'N, causing deprotonation of 4'NH_2 allowing the formation of the C2 thiazolium ring ylid. Although no hydrogen bonding occurs between the thiazolium ring and active site, a quartet of histidine residues His-26, His-66, His-100, and His-261 are positioned towards the entrance of the active site adding stability to the complex formed during catalysis. The diphosphate group acts like an anchor within the deep cleft, forming interaction through hydrogen bonds and salt bridges. The hydrogen bonds are formed with the His-66, Asp-155, Gly-156, Asn-185, Ile-187, and His-261 residues. The Calcium ion is bound at the base of the diphosphate group by the Asp-155, Asn-185, Ile-187 residues, and the two diphosphate oxygen atoms. The metal ion has an important role in the homo-dimerisation (Wang *et al.*, 1997) and catalytic activity (Heinrich *et al.*, 1972). The sulfate molecule is located at the entrance of the active site (Fig. 7.9a, 7.10a), close to the surface. The sulfate molecule interacts with two arginine residues Arg-358, Arg-520 and the His-461 residue. The C2 of TPP is situated at the base of the substrate channel formed by the two subunits. The substrate channel is narrow; making it implausible that the donor and acceptor substrate can bind simultaneously. This observation is evidence for the Bi Bi Ping Pong mechanism that has been proposed (Kremer *et al.*, 1980; Gyamerah and Willetts, 1997).

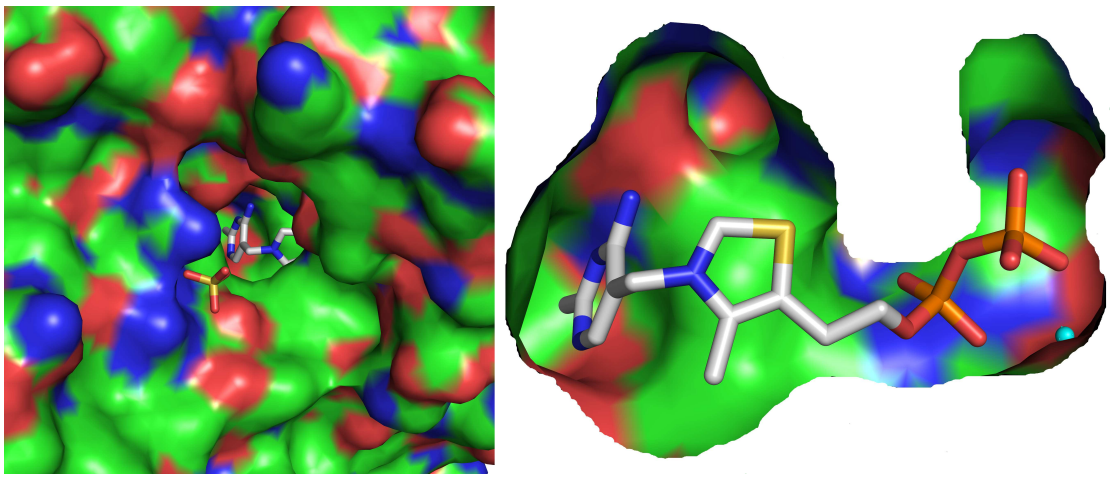


(a)



(b)

Figure 7.9 (a) The active site shown in secondary structure, with TPP at centre and key residues highlighted. (b) Key residues highlighted showing hydrogen bond (yellow) interaction with TPP (silver). Subunits A and B residues coloured in red and green, respectively. Divergent Stereo figures produced in PYMOL (DeLano, 2006).



(a) **(b)**

Figure 7.10 The active site surface, with hydrophobic areas (red) hydrophilic areas (blue), and neutral (green). (a) Shows the entrance to the active site with only the sulfate ion, the aminopyrimdine ring and C2 and N1 of the thiazolium ring viewable. (b) Shows a 'slice through' presentation of the active site with TPP and calcium ion viewable. Figures produced in PYMOL (DeLano, 2006).

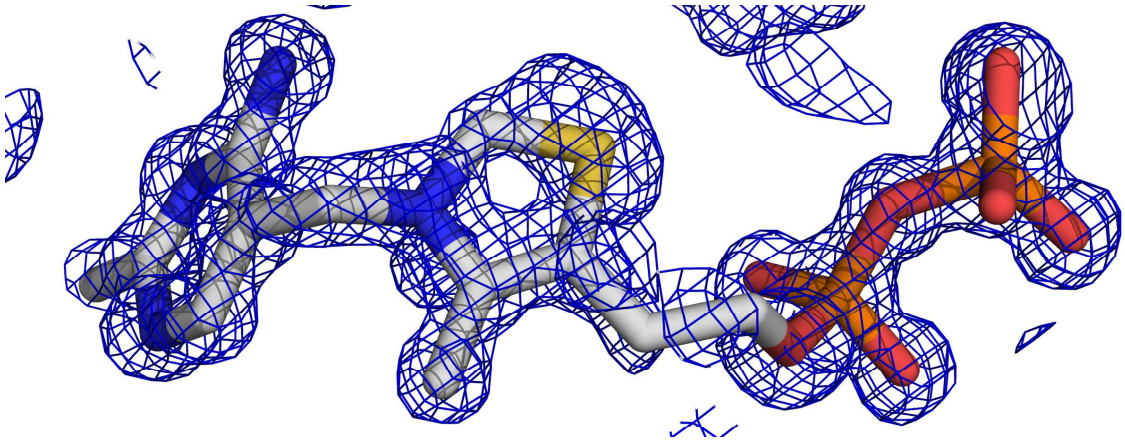


Figure 7.11 $2F_o$-F_c electron density map (contoured to 1.0σ) of the thiamine pyrophosphate cofactor within the active site of the transketolase. Figure produced in PYMOL (DeLano, 2006).

7.3.1 Residues and cofactors involved in catalytic activity

A number of structures of thiamine-diphosphate dependant enzymes with reaction intermediates have been solved, leading to the characterisation of key residues involved in catalysis (Muller *et al.*, 1993; Esakova *et al.*, 2004). The catalytic residues within the active site are all conserved throughout all known transketolase structures. Superimposition of the yeast, maize, *L. mexicana* and *T. thermophilus* with the E. *coli* transketolase, show complete conservation of the residues that are required during catalysis (Table 7.1). Arguably the most important residue involved in transketolase catalysis is glutamic acid 411. The Glu-411 facilitates the protonation of the N1' atom of the amino-pyrimidine ring, which effectively allows the formation of the C2 carbanion in the thiazolium ring (Wikner *et al.*, 1994). The His-473 residue has been proposed to stabilise the charged 4'imino group of the TPP in transition state by deprotonation (Wikner *et al.*, 1997). The His-66 and His-100 residues form a small hydrophilic pocket at the base of the substrate channel close to TPP. The His-66 residue forms a hydrogen bond to the diphosphate tail of TPP (Fig. 7.9b). The His-66 residue may also hydrogen bond with the Cα of the donor substrate, due to its close proximity to the C2 of the thiazolium ring. The role of the His-100 residue focuses around the binding of the donor substrate, by maintaining the substrates orientation during catalysis (Wikner *et al.*, 1994). The interactions of the His-261 and His-26 residues are also concerned with the donor substrate. The location of the two residues allows them to hydrogen bond with Cγ hydroxyl group of the substrate. The natural substrates for transketolase are phosphorylated sugars. Residues His-461, Arg-358, Ser-385 and Asp-469 interact with the phosphate tail of the natural substrates at the entrance of the active cleft during catalysis. The latter residue Asp-469 allows selection of the correct stereoisomer (Nilsson *et al.*, 1997).

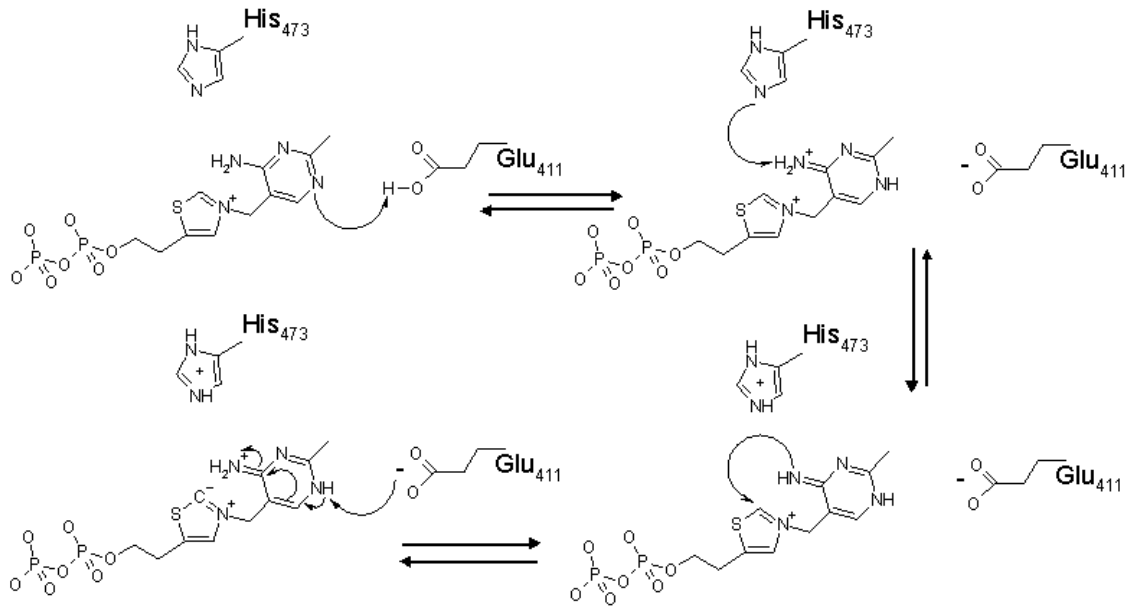


Figure 7.12 The E. *coli* transketolase residues involved in the formation and stabilisation of the production of the 'active' thiazolium ylide.

E. coli **Active site residues**	**Yeast Active site residues**	**Maize Active site residues**	***L. mexicana*** **Active site residues**	***T. thermophilus*** **Active site residues**
Glu-411	Glu-418	Glu-423	Glu-409	Glu-406
His-100	His-103	His-113	His-100	His-104
His-473	His-481	His-486	His-471	His-468
Asp-469	Asp-477	Asp-482	Asp-467	Asp-464
Ser-385	Ser-386	Ser-396	Ser-380	Ser-379
Arg-358	Arg-359	Arg-369	Arg-353	Arg-352
His-461	His-469	His-474	His-459	His-456
His-26	His-30	His-38	His-26	His-30
His-66	His-69	His-78	His-66	His-70
His-261	His-263	His-275	His-261	His-263
Asp-381	Asp-382	Asp-392	Asp-376	Asp-375
Arg-520	Arg-528	Arg-533	Arg-518	Arg-515

Table 7.1 The conserved active site residues, of the five known TK structures, with variations in their amino acid sequence position.

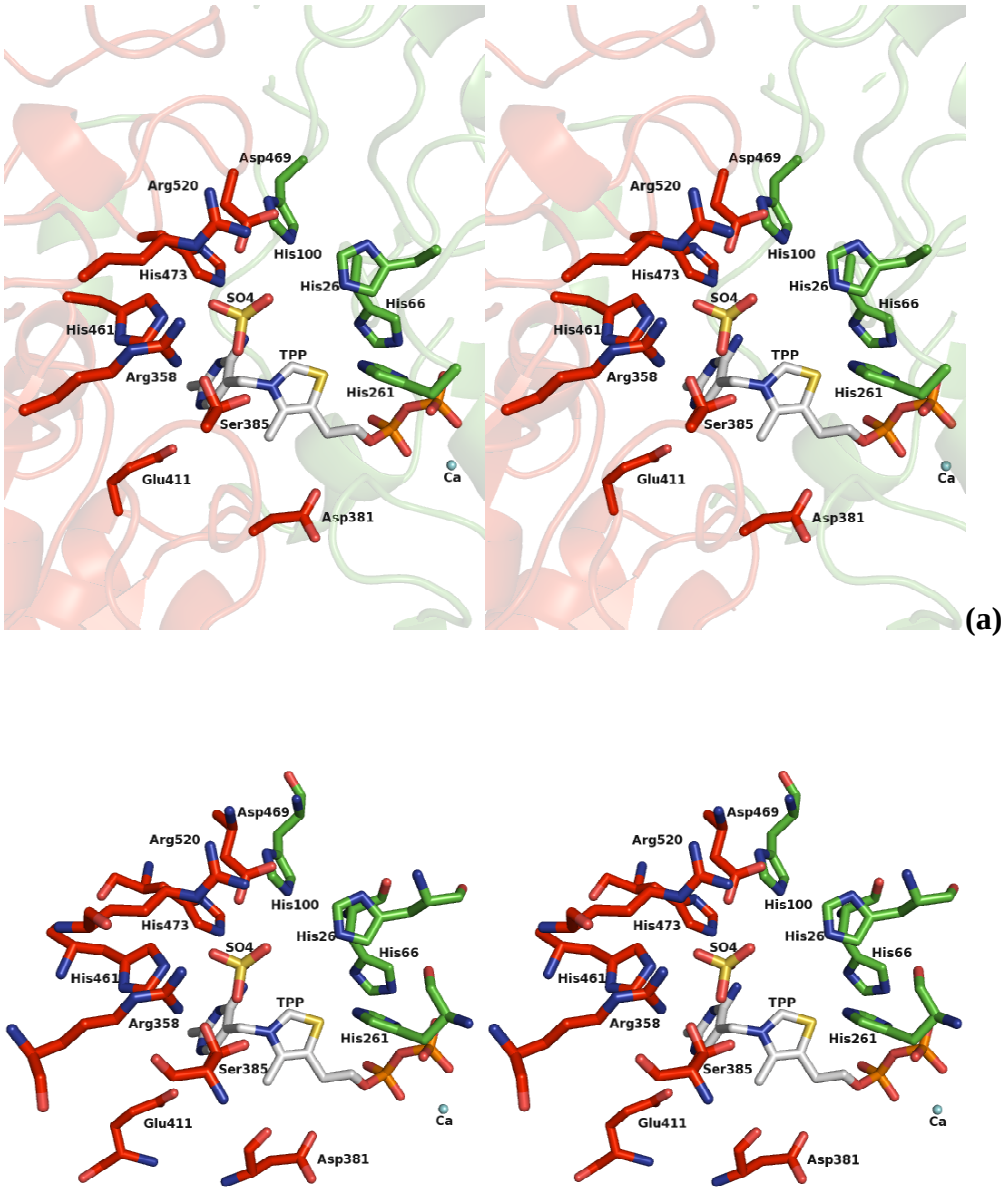


Figure 7.13 (a) The active site shown in secondary structure, with the TPP at centre and key catalytic residues highlighted. Subunits A and B residues coloured in red and green respectively. (b) Key catalytic residues highlighted. Divergent Stereo figures produced in PYMOL (DeLano, 2006).

7.3.2 Catalytic reaction pathway

The reaction mechanism of thiamine proposed by Ronald Breslow in the mid 1950's holds true more than half a century later (Breslow, 1958; Breslow and McNelis, 1959; Breslow, 1962). The catalytic residues surrounding the TPP molecule within the active site were further unravelled by (Kremer *et al.*, 1980), leading to the accepted transketolase catalytic reaction pathway.

The reaction happens in two stages. Firstly, the cleavage of the substrate donor with the formation of the covalent intermediate α-carbanion and an aldose by-product. The second stage involves nucleophilic attack of the α-carbanion by the substrate acceptor, forming a final ketose product that is two-carbon units longer that the substrate acceptor molecule. The catalytic reaction is facilitated by a number of conserved amino acids side groups, which interact and stabilise the donor and acceptor substrate during the formation of the intermediate α-carbanion.

The initial step of the transketolase mechanism is the formation of the covalent intermediate α-carbanion. This is facilitated by the protonation of the N1 nitrogen of the amino-pyrimidine ring by the Glu-411 residue, forming the 4-imino group by subsequent deprotonation of the 4-NH_2. The 4-imino group now acts as a base and is able to deprotonate the C2 of the thiazolium ring, due to the TPP being held in the V-conformation with the now deprotonated 4-NH_2 group in close proximity to the C2. The newly formed C2 thiazolium ring ylid may now attack the carbonyl carbon of the donor substrate, which is held in place by a series of histidine residues. During the formation of the bond between the C2 of TPP and the Cβ of the donor substrate, the Oβ gains a negative charge, which is stabilised by His-473. The covalent intermediate α-carbanion is finally formed as both residues His-26 and His-261 allow deprotonation yielding the enamine and the aldose by-product. The by-product then leaves the active site via the channel it entered (Wikner *et al.*, 1997).

The second stage of the transketolase reaction can now take place, the acceptor substrate enters the channel and is held in place in the same manner as the donor substrate. The previously protonated His-26 and His-261 residues facilitate nucleophilic attack of the α-carbanion on the carbonyl carbon of the acceptor substrate. The intermediate is again stabilised by the His-473 residues removing the proton from the hydroxyl group of the α-carbon. This releases the ketose product leaving a negative

charge on the C2 of the thiazolium ring, which is quenched by the previously protonated 4-imino group.

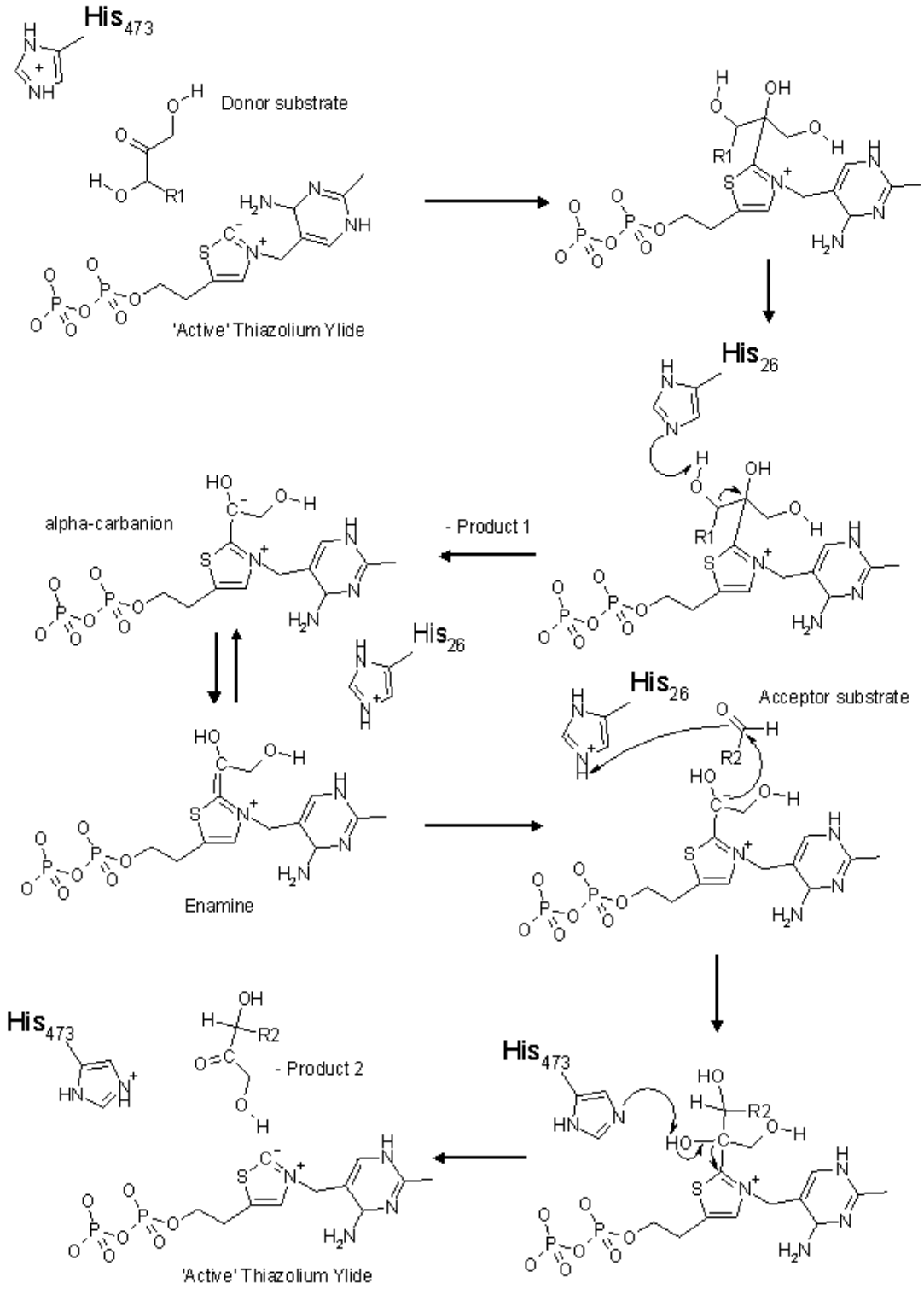


Figure 7.14 The *E. coli* transketolase mechanism, with the donor substrate and acceptor substrate alkyl groups as R1 and R2 respectively. Summarised from proposed mechanism of the yeast transketolase (Wikner *et al.*, 1995).

7.3.3 Water channel ‘The Proton wire’

The transketolase dimer consists of two identical active centres formed on the interface between the two subunits. Each active site containing one molecule of TPP and a divalent metal ion (Kochetov *et al.*, 1976). Even though the two active site crystal structures of the known TK molecules are structurally identical, it has been long hypothesised that there is a form of communication between them (Nikkola *et al.*, 1994). Other evidence supporting communication between the two active sites has been shown by their kinetic non-equivalence. Where observed cooperativity in the TPP binding to TK is associated with the backward reaction of the secondary TPP binding, which indicates conformational instability of one of the active sites in the holo-dimer. This illustrates ‘half-of-the-site’ reactivity (Kovina *et al.*, 1997). Thus the difference between the active sites are only seen on a functional level (Kovina and Kochetov, 1998). This ‘half-of-the-site’ reactivity was recently reviewed by (Frank *et al.*, 2007). Cooperation between the two active sites maybe mediated via a chain of water molecules, which could form a ‘proton wire’ between the two TPP molecules. The chain of waters is observed in the *E. coli* TK structure, where a ‘water channel’ spanning the 20 Å void between the two TPP molecules is formed linking the water molecules through a network of hydrogen bonds. This in theory would facilitate proton transfer between the two TPP molecules, hence allowing communication between the two active sites.

Chapter 8

Structures of the *E. coli* transketolase in complex with substrate analogues

8.1 Building of the hydroxypyruvate and fluoropyruvate complexes

The electron density maps from the two data sets containing the transketolase intermediates were resolved. In separate crystallisation soaking experiments (section 6.2.2) the substrate (hydroxypyruvate) and substrate analogue (fluoropyruvate) were used. In both data sets electron density was observed representing a bound substrate in both clefts of the TK active sites. In the case of TK with the hydroxypyruvate substrate, the most significant feature of the $2F_o$-F_c difference electron density maps was the strong positive electron density extending from the C2 position of the thiazolium ring on the TPP molecule (Fig. 8.2). This indicated that a covalent bond had formed between the hydroxypyruvate and the TPP molecule forming the α-carbanion (enamine) complex that had been alluded to in the mechanism. The α,β-dihydroxyethyl–TPP complex was created in SKETCHER (CCP4, 1994). The model was built into the positive electron density, and refined to feature the entrapment of the hydroxypyruvate substrate. The B-factors of the α,β-dihydroxyethyl moiety were in agreement with that of the rest of the TPP molecule and the surrounding active-site residues.

The $2F_o$-F_c electron density maps of the TK fluoropyruvate complex showed a strong unassigned electron density within the active site (Fig. 8.3). However, unlike the case of TK with the hydroxypyruvate substrate, there was no sign of any electron density extending from the C2 position of the thiazolium ring. This suggested that there was no covalent bond between the TPP and the fluoropyruvate molecule, indicating that the fluoropyruvate molecule was trapped in a 'Michaelis-Menten' like complex. The strong positive electron density showed that the fluoropyruvate molecule has a delocalised double bond between the oxygens of the Cγ, and a single bond connecting the Cγ and Cβ, with the atoms of the Cβ and Cα all within the same plane (Fig. 8.1). The fluoropyruvate dictionary was created in SKETCHER (CCP4, 1994). The molecule was built into the positive electron density, and refinement featured the fluoropyruvate trapped within the active site. After post refinement the B-factors of the fluoropyruvate

were similar to that of the TPP molecule and surrounding residues, showing a good stoichiometric agreement for occupancy of the fluoropyruvate.

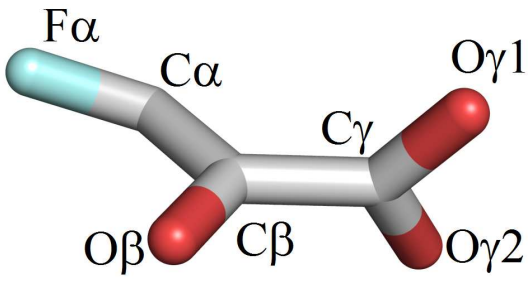


Figure. 8.1 Fluoropyruvate molecule created in SKETCHER (CCP4, 1994), with labelled carbon and fluorine atoms.

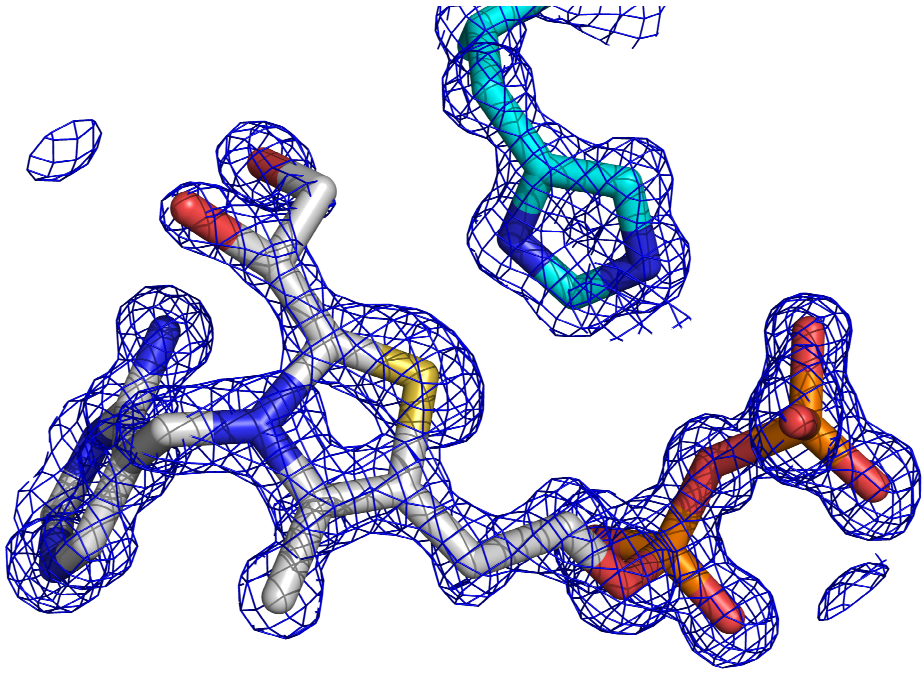
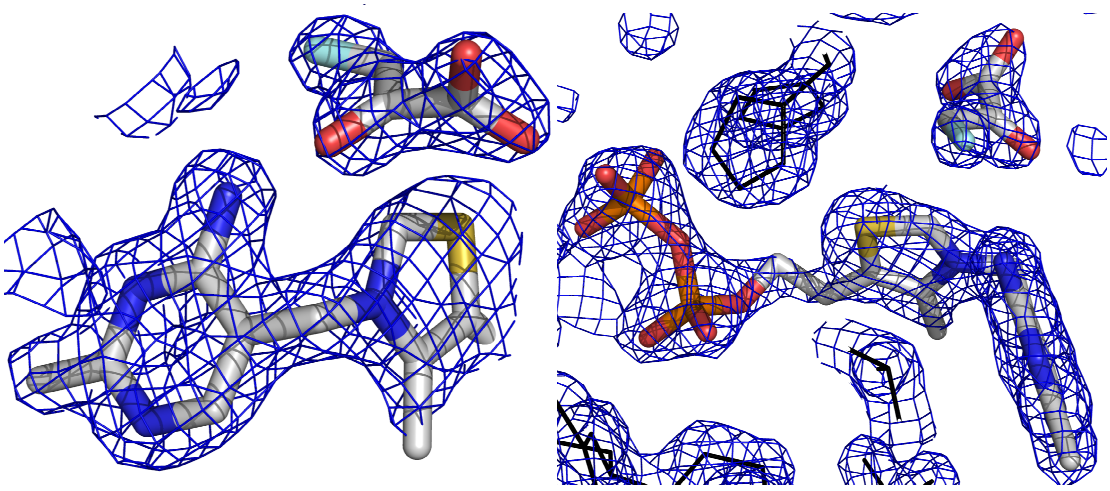


Figure. 8.2 $2F_o$-F_c electron density map (contoured to 1.0σ) around the bound α,β-dihydroxyethyl–TPP (silver) intermediate and His-66 (cyan) within the TK active site.



(a) **(b)**

Figure. 8.3 $2F_o$-F_c electron density maps around the fluoropyruvate complex (silver) within the TK active site, (contoured to 1.0σ). (a) depicts the fluoropyruvate with the aminopyrimidine ring and thiazolium ring of the TPP molecule (silver). (b) depicts fluoropyruvate complex (silver) with His-66, His-261 and Leu-116 residues (black).

8.2 Binding of the hydroxypyruvate molecule within the active site

The hydroxypyruvate molecule forms a covalent bond to the TPP molecule creating the α,β-dihydroxyethyl–TPP complex within the active site. The structure shows the TPP molecule forming hydrogen bonds and salt bridges with the active site residues as previously described in section 7.3 (Fig. 8.4). The high-resolution crystallographic data indicates that the atoms of the thiazolium ring form an angle of 166° (Fig. 8.6) with the plane of the bond between the C2 and Cβ of the α,β-dihydroxyethyl complex. Previous results indicated that the bond between the C2 and Cβ of the α,β-dihydroxyethyl complex existed within the same plane as the thiazolium ring (Fiedler *et al.*, 2002), however results from this structure show a 14° deviation from planarity. It may also be noted that evidence from the electron density map indicated that the C2-N, and C2-S bonds of the thiazolium ring might lie outside the plane. Good agreement was made when the in-plane restraints of the C2-N, and C2-S bonds were lifted.

The α,β-dihydroxyethyl moiety itself forms a total of six hydrogen bonds with the surrounding active site residues and water molecules (Fig. 8.5). The Oβ atom forms two hydrogen bonds, one with the His-473 residue (2.7 Å) and another with a nearby water molecule (3.2 Å) that is held in position by the Asp-469 and His-26 residues. The

Oβ atom lies in the Cα and Cβ plane, with the Oα atom forming an angle of 105 ° angle with the plane of the Cα confirming the sp^3 hybridisation of the Cα. The Oα atom forms four hydrogen bonds, three with residues His-473 (3.3 Å), His-100 (2.7 Å) and Gly-114 (3.4 Å). The other hydrogen bond is formed through a water molecule (2.7 Å) that is conserved through its interactions with the Gly-114 and His-66 residues. The Oα atom is positioned in close proximity to the 4-NH_2 group of the amino-pyrimidine ring. In theory the Oα atom could form hydrogen bonds with either of the hydrogens on the nitrogen, as a distance of 2.9-3.0 Å separates them. However, a hydrogen bond angle of 84° calculates to a weak bond enthalpy, therefore this event is unlikely to occur. The high-resolution crystallographic data shows that there are no alterations in the positions of the residues within the active site on formation of the α,β-dihydroxyethyl–TPP intermediate, when compared to the native TK structure (holo).

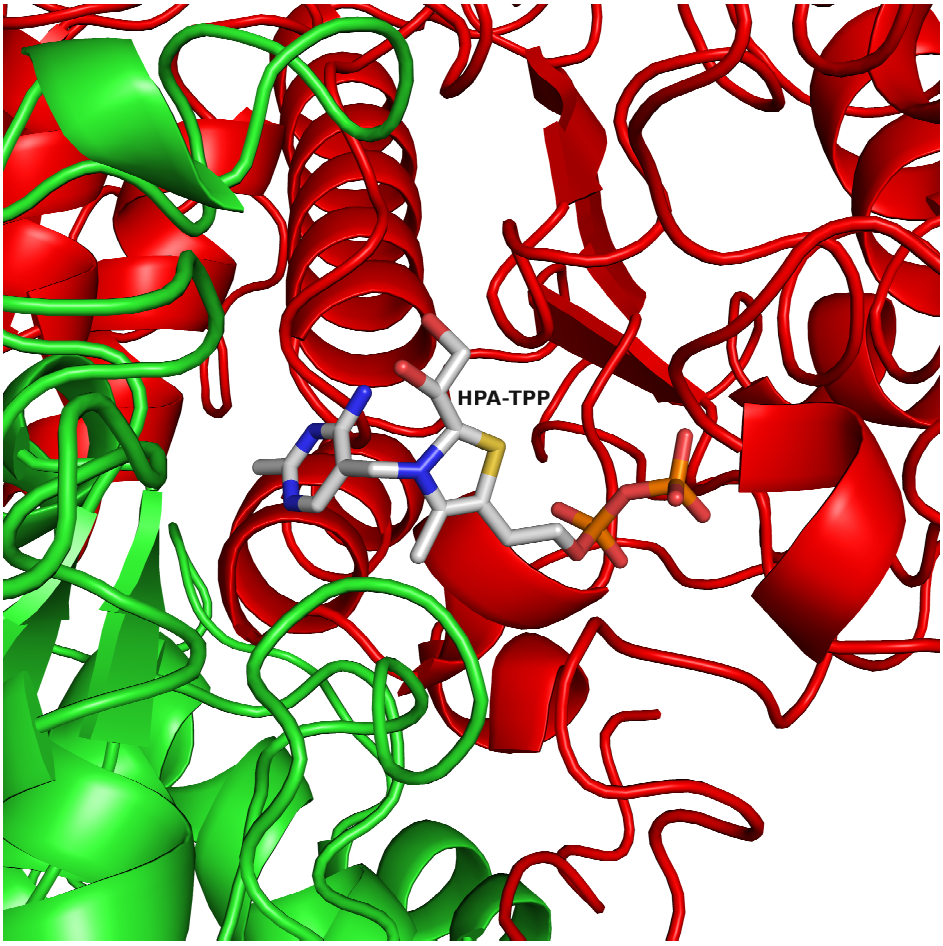


Figure 8.4 The diagram depicting the secondary structure with the α,β-dihydroxyethyl–TPP complex (HPA-TPP)(silver) around the active site of TK formed in the cleft between subunits A (red) and B (green). Figure produced in PYMOL (DeLano, 2006)

Further comparisons reveal slight positional changes of the thiazolium ring. The high-resolution data indicates multiple positions of the N, S and C2 atoms. This indicates possible movement during the formation of the α,β-dihydroxyethyl–TPP intermediate. The positions of the atoms in the phosphate tail and amino-pyrimidine ring remain unchanged throughout the reaction and are all conserved, which in turn allows the V-conformation to be maintained (Fig. 8.7).

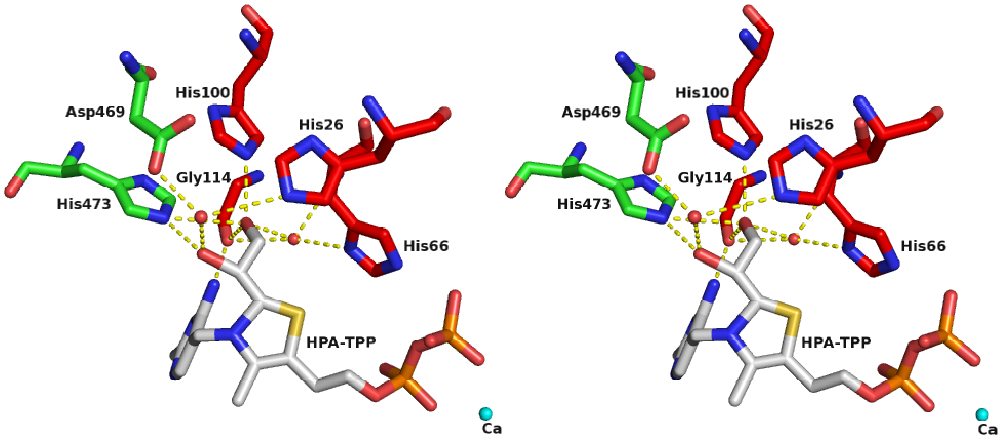


Figure 8.5 A Divergent Stereo representation of the hydrogen-bonding network (yellow) between the conserved residues of subunits A (red) and B (green), depicting the stabilisation of the α,β-dihydroxyethyl–TPP complex (HPA-TPP)(silver). Figure produced in PYMOL (DeLano, 2006).

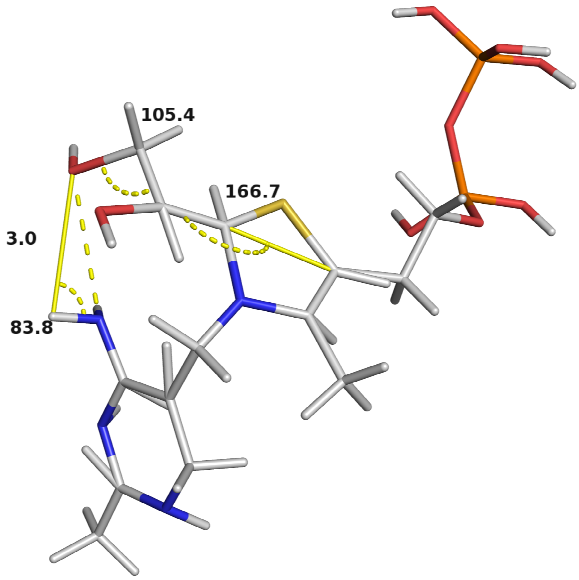


Figure 8.6 The bonding angles formed between the thiazolium ring and hydroxypyruvate on formation of the α,β-dihydroxyethyl–TPP intermediate (HPA-TPP). Also shown is the bond angle and length of the interaction between NH_2 group and Cα oxygen. Figures produced in PYMOL (DeLano, 2006)

8.3 Binding of the fluoropyruvate molecule within the active site

The fluoropyruvate and TPP molecules form a 'Michaelis' like complex within the active site (Fig. 8.7). The fluoropyruvate is firmly held in place within the active site through a series of hydrogen bonds, and electrostatic interactions with the active site residues. The α-fluorine atom hydrogen bonds through His-100 (2.1 Å) and His-473 (3.2 Å). The fluorine is also in close proximity to the 4-NH_2 group of the amino-pyrimidine moiety, and is separated by a distance of 2.9 Å with a bond angle of 87°. The angle is close to the pyramidal hydrogen bond symmetry angle of 89°, therefore raising the possibility of hydrogen bond formation, as with the Oα of the HPA-TPP complex.

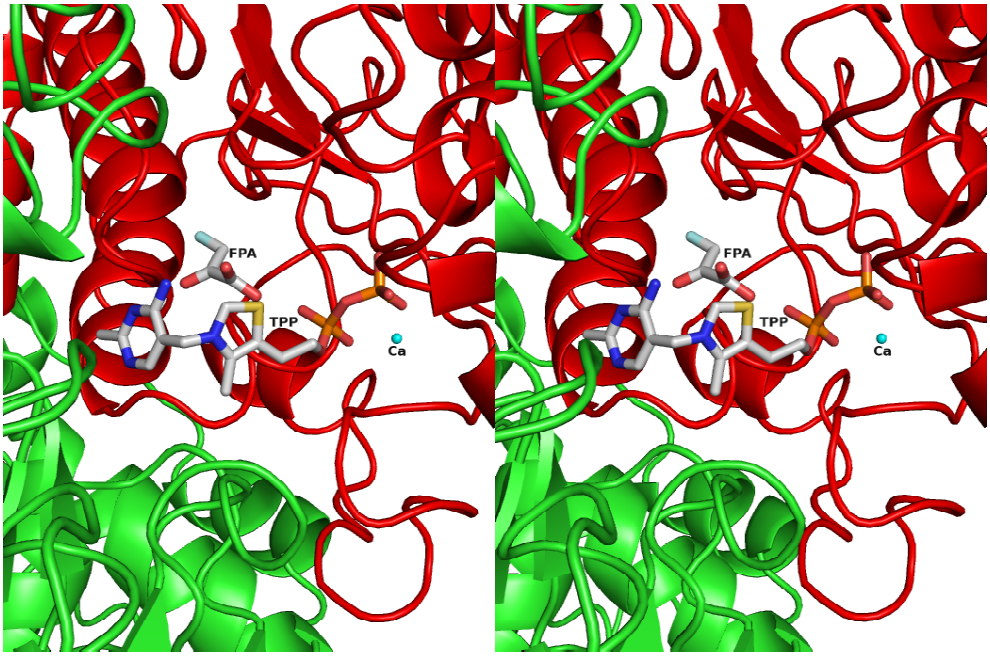


Figure 8.7 A divergent stereo representation of the fluoropyruvate Michaelis complex (FPA-TPP)(silver) located within the active site of TK formed in the cleft between subunits A (red) and B (green). Figure produced in PYMOL (DeLano, 2006).

The Oβ atom forms a hydrogen bond with the His-473 residue (2.67 Å). This anchors the Cβ within a distance of 3.1 Å of the C2 from the thiazolium ring. This would form an angle of nucleophilic attack of 114° by the C2 towards the Cβ, which unusual but close to the accepted Burgi-Dunitz angle of 109° for nucleophilic attack (Burgi *et al.*, 1973). This further indicates the strained C2-Cβ bond formed in the intermediate complex of the TPP and the natural ketol sugar substrates (Asztalos *et al.*, 2007) (Fig. 8.9). The two oxygens Oγ1 and Oγ2 are both hydrogen bonded by the His-26 residue (3.1 and 3.2 Å respectively), with the Oγ1 also hydrogen bonding to the His-261 residue (3.3 Å). This positions the two oxygens (Oγ1, Oγ2) in the same plane as one another, but approximately 90° to the plane of Cα and Cβ (Fig. 8.9).

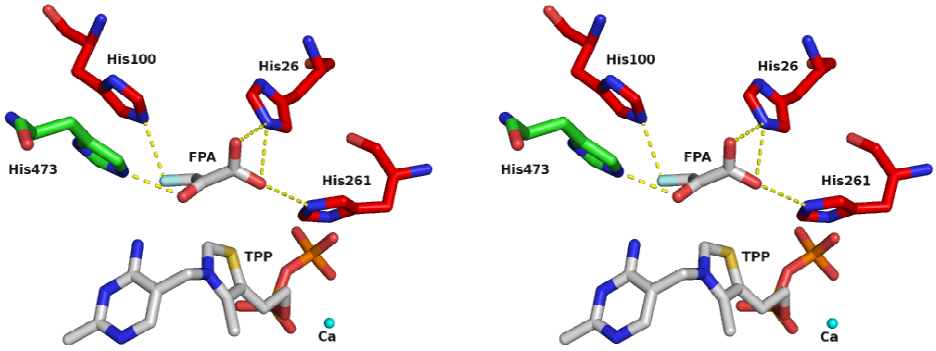


Figure 8.8 The hydrogen bonding network (yellow) between the conserved residues of subunits A (red) and B (green) of the TK enzyme, depicting the stabilisation of the fluoropyruvate Michaelis complex (FPA-TPP)(silver). Divergent Stereo figure produced in PYMOL (DeLano, 2006).

During the formation of the complex, the Cβ must changes hybridisation from sp^2 to sp^3. The high-resolution crystallographic data shows that surprisingly there are no alterations in the positions of the residues within the active site, when compared to the native TK structure (holo) on formation of the Michaelis complex. Observation of the C2 of the thiazolium ring again shows possible movement outside of the thiazolium ring plane towards the Cβ of the fluoropyruvate molecule. This indicates the C2 readiness for nucleophilic attack on the Cβ. This also highlights the kinetically unfavourable formation of the strained covalent C2-Cβ bond in the α,β-dihydroxyethyl–TPP intermediate. The positions of the atoms in the phosphate tail and amino-pyrimidine ring remain unchanged throughout the reaction and are all conserved, which in turn allows the V-conformation to be maintained.

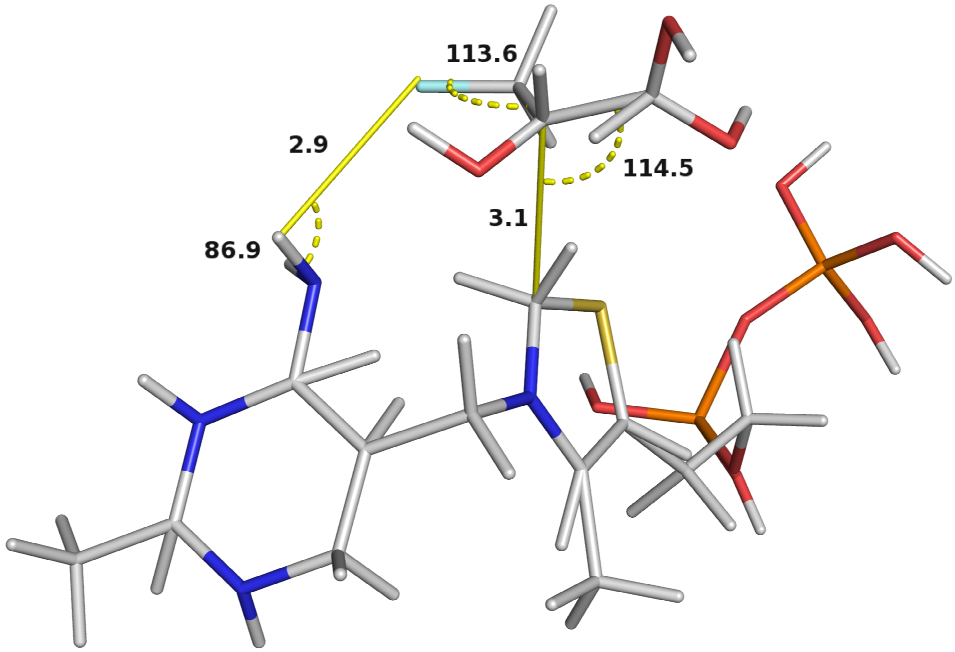


Figure 8.9 Diagram showing the bonding angles formed between the thiazolium ring and fluoropyruvate on formation of the fluoropyruvate Michaelis complex (FPA-TPP). Also shown is the bond angle and length of the interaction between the NH_2 group and Oα. Figures produced in PYMOL (DeLano, 2006)

8.4 Comparisons in the binding of the hydroxypyruvate and fluoropyruvate molecules within the active site

In the transketolase reaction the ketol donor substrate binds to the TPP molecule, forming the α,β-dihydroxyethyl–TPP complex. This in turn releases the remaining moiety of the ketol donor substrate. In this study, hydroxypyruvate (substrate) and fluoropyruvate (substrate analogue) act as the ketol donors. The hydroxypyruvate is able to form the α,β-dihydroxyethyl–TPP complex, whilst the fluoropyruvate forms the Michaelis complex. The fluoropyruvate molecule cannot form the α,β-dihydroxyethyl–TPP complex (hence entrapment) as the Cα hydroxyl group on the hydroxypyruvate has been replaced with a fluorine atom. However, analogous interactions occur due to the similarities in chemical property. This allow comparison between the 'before' (fluoropyruvate Michaelis complex) and the 'after' (α,β-dihydroxyethyl–TPP complex) states of the first stage (two carbon-subunit donation) in the transketolase reaction. This gives an insight into the positioning and movement of the ketol donor during catalysis.

Superimposition of the α,β-dihydroxyethyl–TPP complex structure with the fluoropyruvate Michaelis complex structure demonstrates the conserved positions of active site residues in the binding of the two substrates. The key feature is the positioning of the Oα and Oβ atoms prior to nucleophilic attack (Fig. 8.10, 8.11). The Oα atom of the α,β-dihydroxyethyl–TPP complex overlaps in exactly the same position as the Fα atom in the fluoropyruvate Michaelis complex. Both atoms share the same interactions with the His-100, His-473, Gly-112 residues, and the 4-NH_2 group of the amino-pyrimidine moiety (Fig. 8.10). This implies that these interactions stabilise the Oα atom in this location, thus aiding the positioning of the Cβ. This allows nucleophilic attack by the C2 of thiazolium ring to occur. Similarly, the Oβ atom of α,β-dihydroxyethyl–TPP complex has identical structural coordinates to that of the Oβ atom in the fluoropyruvate Michaelis complex. Both of the oxygens form strong hydrogen bonds with the His-473 residue, leading to further stabilisation of the complex. The key interactions of the Oα/Fα and the Oβ atoms (of the substrate and substrate analogue), with the active site residues allow a possible hypothesis that these atoms remain fixed in position throughout the reaction. This could imply that the Cα and Cβ carbons rotate into the coordinates of the α,β-dihydroxyethyl–TPP complex, from the coordinates of the fluoropyruvate Michaelis complex using the Oα/Fα and Oβ atoms as a rotational axis. The Cβ's on the respective substrates are 1.6 Å apart in the two structures. The

fluoropyruvate Michaelis complex Cβ would be required to rotate through 86° to reach the final coordinates of the α,β-dihydroxyethyl–TPP intermediate (Fig. 8.11).

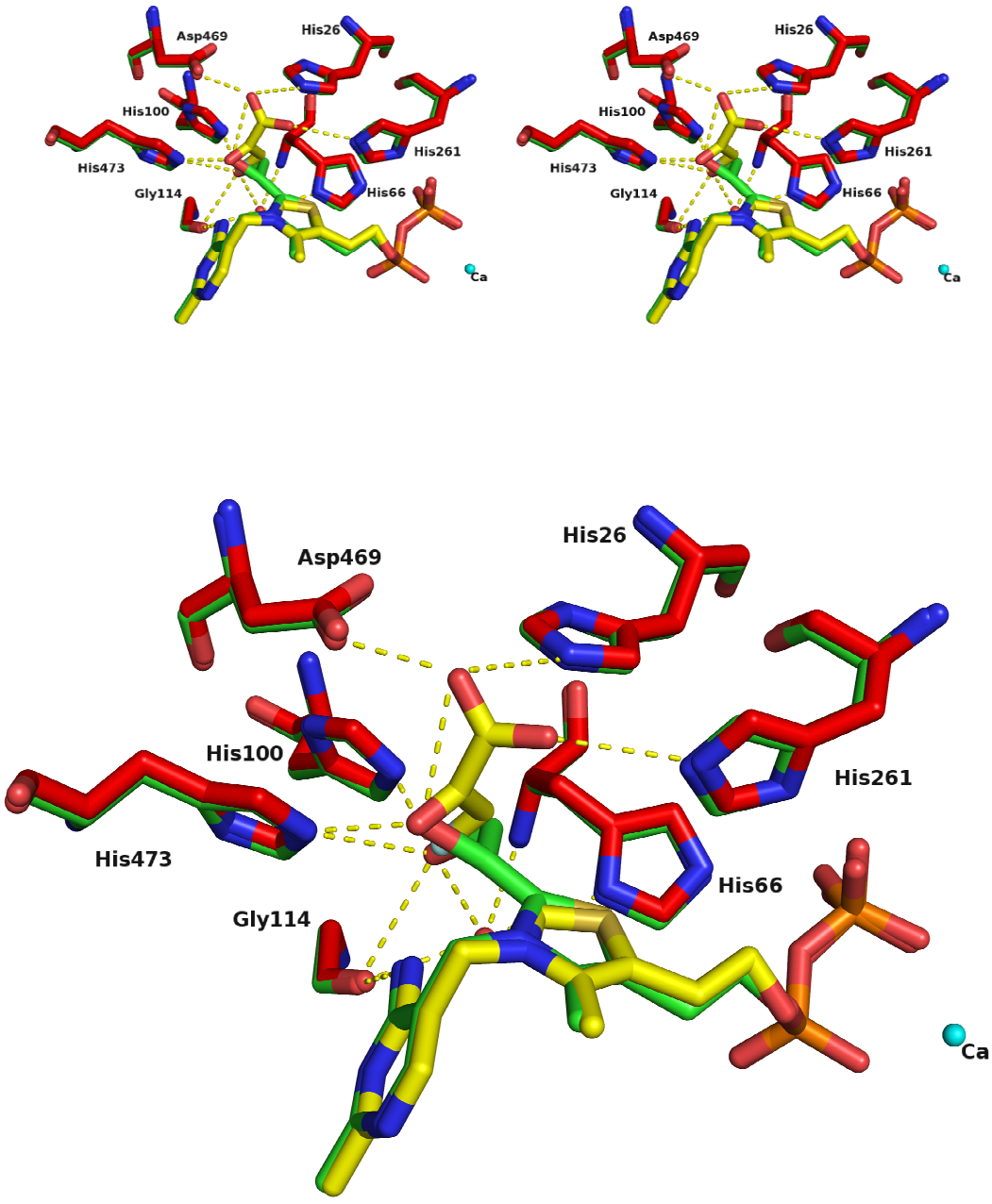


Figure 8.10 Divergent stereo and non-stereo diagrams of the hydrogen bonding network (yellow) between the superimposed structures of fluoropyruvate Michaelis complex (complex yellow, residues red) and α,β-dihydroxyethyl–TPP complex (complex green, residues green) within the active site of the TK enzyme. Figure produced in PYMOL (DeLano, 2006).

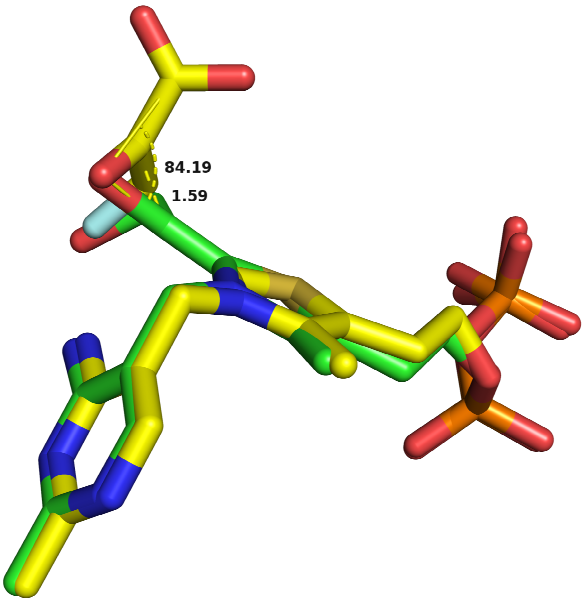


Figure 8.11 Diagram showing the superimposition of fluoropyruvate Michaelis complex (FPA-TPP)(yellow) with the α,β-dihydroxyethyl–TPP complex (HPA-TPP)(green), including the bonding angles formed between the substrate and substrate analogue. Figures produced in PYMOL (DeLano, 2006).

8.5 Comparisons in the binding of the hydroxypyruvate and fluoropyruvate complexes with other TK intermediates

The active site residues within the family of transketolase enzymes are well conserved. All five known holo structures yeast, maize, *L. mexicana*, *T. thermophilus*, and *E. coli* have been shown to exhibit invariant active site residues involved in catalysis, as discussed in section 7.3.1. It can therefore be proposed that during the transketolase reaction the same residues perform the same roles throughout catalysis, independent of the source of transketolase enzyme.

The addition of hydroxypyruvate to transketolase forming the α,β-dihydroxyethyl–TPP complex, is similar to that of the structure of the yeast transketolase in complex with hydroxypyruvate (Fiedler *et al.*, 2002). Superimposition of the two structures reveals an excellent agreement between the two α,β-dihydroxyethyl–TPP complex formed, with an RMS of 0.280 over the 26 atoms. The hybridisation of the Cα and Cβ are well conserved, with both the Oα and Oβ atoms forming the same key interactions with His-100, His-473, Gly-112, 4-NH_2 group and His-473 respectively (Fig. 8.12), thereby stabilising the intermediate. (see Table 7.1 for the yeast residue equivalents).

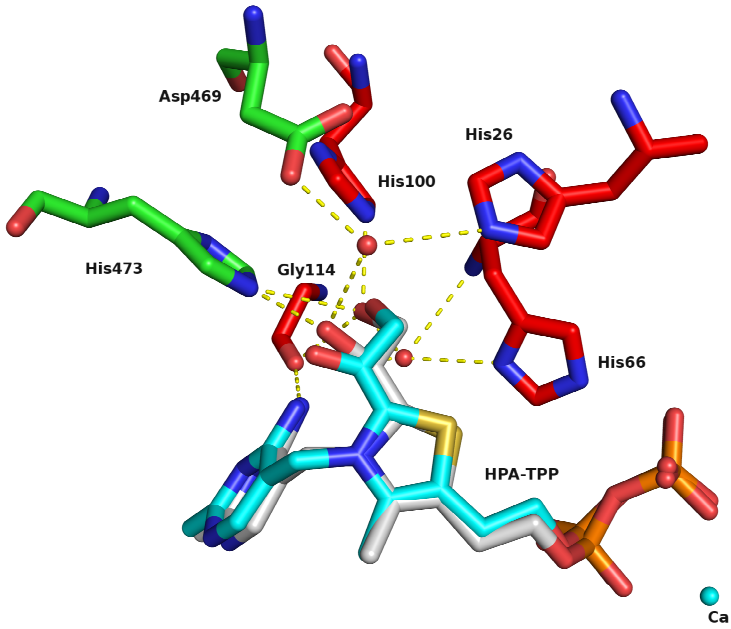


Figure 8.12 Diagram showing the superimposition of *E. coli* α,β-dihydroxyethyl–TPP complex (HPA-TPP)(silver) with that of the yeast α,β-dihydroxyethyl–TPP complex (1GPU)(cyan), showing the hydrogen bonding network (yellow). Figure produced in PYMOL (DeLano, 2006).

It must be noted that in the case of the yeast HPA-TK complex, the Cα and Cβ are in the same plane as the thiazolium ring, but in the *E. coli* HPA-TK complex, the Cα and Cβ are 14° above the plane of the thiazolium ring. This observation is possibly due to the quality of the 1.1 Å high-resolution data of the *E. coli* HPA-TK complex, compared to that of the 2.0 Å resolution of the yeast HPA-TK complex.

In the pentose phosphate pathway xylulose 5-phosphate is the substrate, however in industry the hydroxypyruvate molecule is used as a ketol donor. This is due to the production of carbon dioxide with hydroxypyruvate during the catalysis, rendering the reaction irreversible. The natural substrate xylulose 5-phosphate is trapped within the *E. coli* transketolase active site in the pdb 2R80 (Asztalos *et al.*, 2007). Superimposition and comparisons between the two structures show the C2 of the thiazolium ring in the xylulose 5-phosphate-TPP intermediate shifted 23° out of the plane, when compared to the C2 of the α,β-dihydroxyethyl–TPP complex. A possible rationale for this movement could be due to the additional interactions formed by the larger xylulose 5-phosphate substrate within the active site. The Oα and Oβ atoms of the phosphorylated sugar intermediate overlap with the Oα and Oβ atoms of the α,β-dihydroxyethyl–TPP complex, forming the same residue interactions as seen previously (Fig. 8.13, Fig. 8.14). The phosphate tail of the sugar forms interactions near the entrance of the cleft with the residues Arg-520, His-461, and Ser-385, whilst the remaining sugar xylulose moiety interacts with the residues Glu-469, His-26, and His261. The C2-Cα bond is further strained in the phosphorylated sugar intermediate, thus on elimination of the glycerol 3-phosphate the strain is partially removed by the formation of the α,β-dihydroxyethyl–TPP complex.

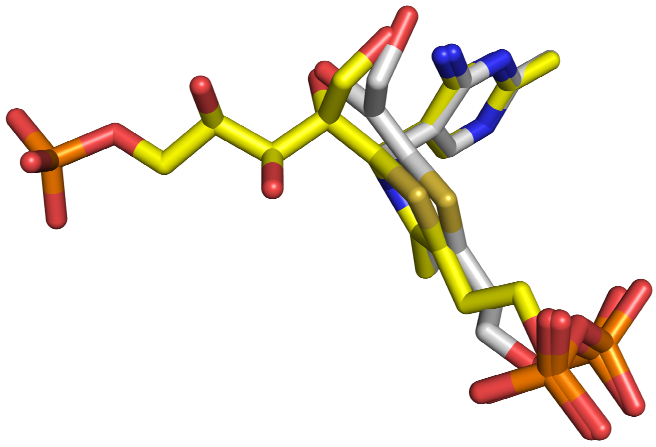


Figure 8.13 Diagram showing the superimposition of α,β-dihydroxyethyl–TPP complex (HPA-TPP) (silver) with Xylulose 5-phosphate-TPP intermediate (yellow) PDB code 2R8O. Figure produced in PYMOL (DeLano, 2006)

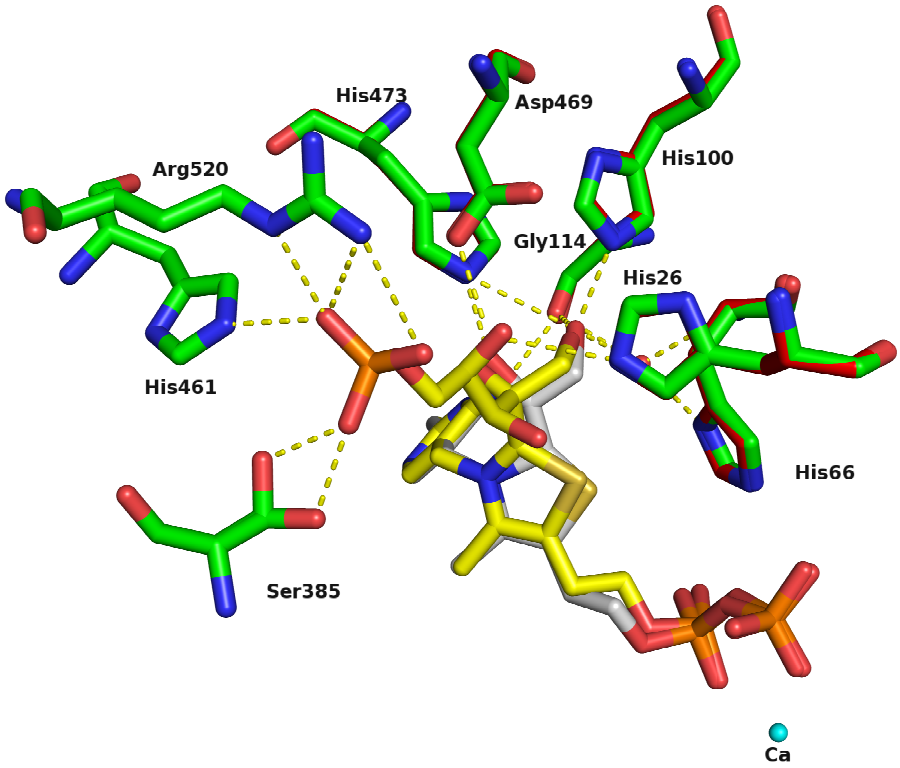


Figure 8.14 Diagram showing the hydrogen bonding network (yellow dashes) between the superimposed structures of α,β-dihydroxyethyl–TPP complex (silver) residues (red) and xylulose 5-phosphate-TPP intermediate (yellow) residues (green) PDB code 2R8O. Figure produced in PYMOL (DeLano, 2006).

The fluoropyruvate Michaelis complex trapped during the donor stage of the catalysis has no parallel with any known intermediate structures of transketolase. However, comparisons maybe drawn with the two structures of the bound donor substrates fructose 6-phosphate and xylulose 5-phosphate in PDB's 2R8P and 2R8O respectively (Asztalos *et al.*, 2007). Superimposition of xylulose 5-phosphate with of α,β-dihydroxyethyl–TPP complex may be seen in figure 8.15. Superimposition of the two structures with the fluoropyruvate Michaelis complex, presents an insight into the movement of the Cα and Cβ atoms of the sugar substrates towards the C2 of the thiazolium ring during the nucleophilic attack. The same active site residues sanction the conserved overlap of the Oα/Fα and Oβ atoms in both bound sugar substrate intermediates as seen in the fluoropyruvate Michaelis complex. The bound sugar substrates form an angle between the C2-Cα-Cγ of 109.5°. However the same angle in the fluoropyruvate complex is 114.5°, which may be seen as the angle pre-nucleophilic attack. This suggests a movement of 5° during the formation of the strained C2-Cβ bond seen in the intermediates (Fig. 8.15, Fig. 8.16).

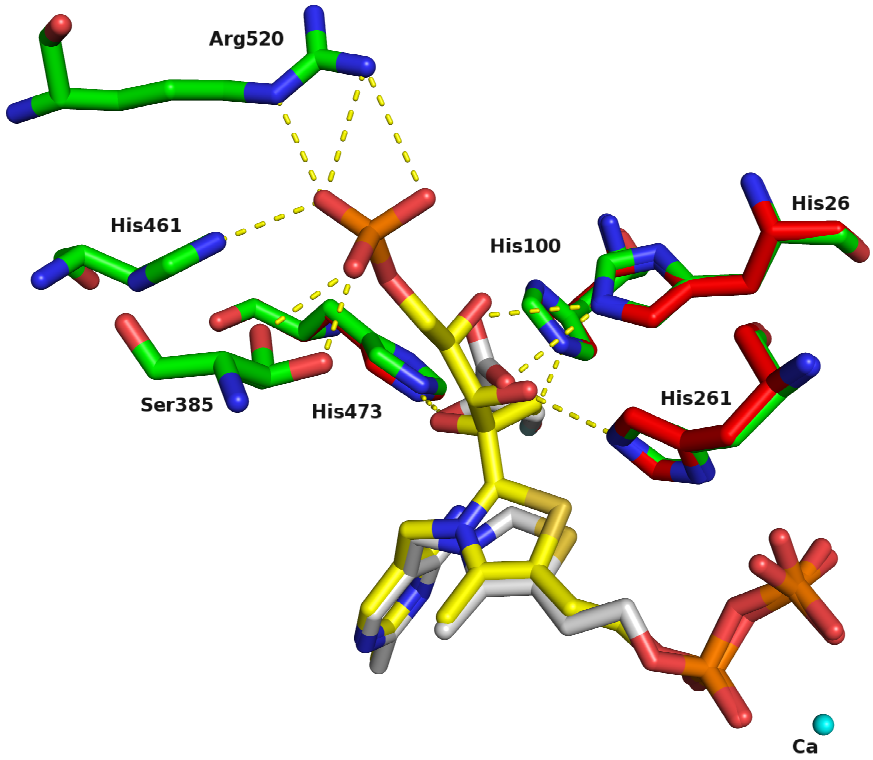


Figure 8.15 Diagram showing the hydrogen bonding network (yellow) between the superimposed structures of fluoropyruvate Michaelis complex (FPA-TPP)(silver) residues (red) and Xylulose 5-phosphate-TPP intermediate (yellow) residues (green) PDB code 2R8O. Figure produced in PYMOL (DeLano, 2006).

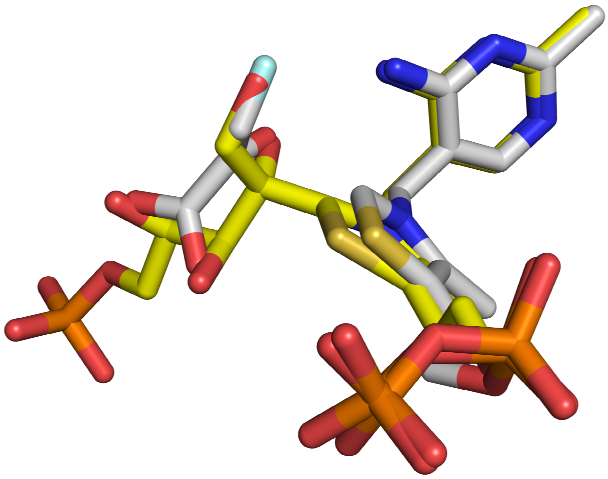


Figure 8.16 Diagram showing the superimposition of fluoropyruvate Michaelis complex (FPA-TPP) (silver) with Xylulose 5-phosphate-TPP intermediate (yellow) PDB code 2R8O. Figures produced in PYMOL (DeLano, 2006).

A noteworthy point regarding the trapped fluoropyruvate Michaelis complex concerns the plane in which the fluoropyruvate is positioned in comparison to that of the xylulose 5-phosphate-TPP intermediate. The fluoropyruvate molecule lies parallel to the plane of the covalently bound xylulose 5-phosphate moiety of the intermediate, with the parallel planes separated by a distance of 0.9 Å (Fig. 8.16).

8.6 Implication of the binding in the hydroxypyruvate and fluoropyruvate complexes on the transketolase mechanism

The evidence gathered from the two structures of the α,β-dihydroxyethyl–TPP complex and the fluoropyruvate Michaelis complex instigates further discussion regarding the donor substrate stage of the transketolase reaction. It is understood that the driving force behind the reaction is the formation of the strained C2-Cβ bond, leading to elimination of CO_2 in the hydroxypyruvate substrate, which in turn reduces the strain upon C2-Cβ bond. The catalytic residues act to stabilise the position of the donor substrate and are therefore conserved throughout catalysis. As a result of the stabilisation the strained C2-Cβ bond is allowed to exist prior to the arrival of the acceptor substrate. Entrapment of the fluoropyruvate within the active site allows a better description of the donor mechanism in the formation of the α,β-dihydroxyethyl–TPP complex.

Initially, the donor substrate enters the cleft of the active site and migrates towards the thiazolium ring of the TPP molecule. The hydroxyl group on Cα forms hydrogen bonds with the His-473 and His-100 residues. The latter residue is vital in recognition of the hydroxyl group, discriminating between the hydroxypyruvate and pyruvate molecules (Wikner *et al.*, 1994). The possible hydrogen bond formed by the NH_2 group of the amino-pyrimidine ring also contributes toward stability during this event. The His-473 residue also forms a hydrogen bond with the Oβ atom contributing further to the stabilisation of the donor substrate. The Cγ–dihydroxyl tail of the hydroxypyruvate is stabilised by two interactions with the His-26 and His-261 residues.

The carbanion ylid is formed as described in section 7.3.2 and acts as a nucleophile. At this point the electronegative Oβ atom draws electron density away from the π-bonded Cβ forming a slight positive charge on the Cβ. The positive charge on the Cβ provokes nucleophilic attack from the carbanion ylid, causing the Cβ's hybridisation to change from sp^2 to sp^3. Thus forming the initial covalent bond of the

donor intermediate. The 4-NH_2 groups importance during the initial formation of the C2-Cβ covalent bond must be noted. As the previously protonated 4-NH_2 group (by deprotonating C2) forms a hydrogen bond with the Oα atom aiding stability. Again augmenting the importance in the role of the 4-NH_2 group of the amino-pyrimidine ring during catalysis. The catalytic importance of the 4-NH_2 group has been highlighted previously (Golbik *et al.*, 1991; Schneider and Lindqvist, 1993).

A key area in this reaction is the angle of attack of the nucleophilic C2 of the thiazolium ring with the Cβ of the substrate during the formation of the C2-Cβ strained covalent bond. The Burgi-Dunitz angle for nucleophilic attack of a carbonyl centre is 109.5°, but the angle formed in the fluoropyruvate complex prior to the bond formation is 114.5°. This suggests that the Cβ is in a form of transient hybridisation between sp^2 (120°) before forming the sp^3 (109.5°) hybrid. The mechanistic 'snapshot' illustrated by the trapped fluoropyruvate in complex with the TK enzyme has not be seen before. This high-resolution structure has allowed hypothesis and further debate regarding the chain of events during the transketolase mechanistic pathway.

However this is a rather optimistic conclusion, as there are many factors that heavily repute this hypothesis. The high concentration of FPA required in crystallisation, and the resulting weak electron density makes concluding any putative attack awkward, especially at the small 5° proposed.

Chapter 9

Structure of transketolase mutant H26Y complexed with a substrate analogue

9 1 Introduction

Mutations can be made in enzymes to further probe into the function of specific residues, in an attempt to elucidate the catalytic mechanism. Mutations can also be made to optimise an enzyme for a specific biotransformation. Such mutations can be used to increase specificity and stability of the enzyme depending upon the requirements of the reaction. There have been many mutations made with the transketolase enzyme, all of which are concerned with the alterations of residues within the active site. The earliest mutation was involving the glutamic acid-418 residue of the yeast transketolase, where the residue has been mutated to alanine and glutamine (Wikner *et al.*, 1994). These mutations were made to examine what effect the Glu-418 residue had on the enzymatic mechanism. The glutamic acid residue was found to hydrogen bond to the N1' nitrogen atom of the pyrimidine ring. This interaction is involved in the generation of the active C2 carbon of the thiazolium ring. It was observed that the Ala-418 mutant had 0.1 % activity, whilst the Gln-418 mutant had 2 % when compared to the wild-type. This research illustrated the requirement for the presence of the glutamic acid residue for catalysis. The next residue subjected to mutagenesis was the His-100 residue in the yeast transketolase (Wikner *et al.*, 1995). The His-100 residue is involved with the stabilisation of the Oα atom of the donor substrate during the formation of the α,β-dihydroxyethyl–TPP complex. A series of mutations were made by Wikner *et al.*, in which they showed the fall in activity of the mutants were due to the different properties of the mutated amino acid residues, and not structural changes of the active site. Meshalkina *et al.*, continued probing into the active site of the yeast transketolase by mutating the Asp-382 to an alanine residue, and the mutation of the Glu-162 to a glutamine (Meshalkina *et al.*, 1997). Here they investigated the interaction that these two residues have on dimer formation in yeast transketolase. They found that mutagenesis of the Glu-162 yielded a reduction in the formation of the dimer. This underlines the residues importance in the formation and

stabilization of the dimer. The Asp-382 mutations provided evidence for the residues role in the binding of the cofactor, most likely via an electrostatic charge compensation of the positively charged thiazolium ring. The effect of another mutation in the active site of the yeast transketolase was in the study of histidine-263 (Wikner *et al.*, 1997). The findings of this study proved that the His-263 residue is primarily involved with the acid/base catalysis, and a lesser effect towards the stabilizing of the substrate. In a separate case, the Asp-477 residue was mutated to an alanine residue (Nilsson *et al.*, 1997). Here, it was found that the Asp-477 residue had a resounding effect upon the enantioselectivity of the yeast transketolase. The Asp-477 is located in the substrate channel of the enzyme and forms a hydrogen bond with the Cβ hydroxyl of the acceptor substrate. Alteration of the Asp-477 to an alanine reduced the production of the D-enantiomer product a thousand fold. Two histidine residues, at position 26 and 261 were mutated to alanine residues in the *E. coli* transketolase. These histidine residues are involved in the stabilization of the intermediate by forming a hydrogen bond network with the hydroxyl group of the substrate. It was found that the removal of the histidine residue reduced the production of the product in comparison to the wild-type enzyme (Asztalos *et al.*, 2007). The most recent mutations performed upon the transketolase enzyme were conducted by Hibbert *et al.* Here many alterations upon the *E. coli* transketolase have been performed, one of which has been investigated during this chapter, the histidine-26 to tyrosine mutation (Hibbert *et al.*, 2007).

9.2 Building of the H26Y variant structure with the bound fluoropyruvate

The electron density map from the data set containing the H26Y TK mutant with the bound fluoropyruvate was resolved. In the data set electron density was observed for the mutated tyrosine residue, and the bound substrate analogue in both clefts of the TK active sites. Simple alteration of the residue was performed in COOT (Emsley and Cowtan, 2004) replacing the histidine residue with a tyrosine residue in the strong positive electron density at position 26 of the structure.

The $2F_o$-F_c difference electron density maps of the TK fluoropyruvate complex showed a strong unassigned electron density within the active site. Again, like the previous entrapment of the fluoropyruvate molecule in chapter 8, no electron density was observed emanating from the C2 of the thiazolium ring. This again would suggest that there was no covalent bond between the TPP and the fluoropyruvate molecule, indicating that the fluoropyruvate molecule was again trapped in a ‘Michaelis’ like

complex. The strong positive electron density for the fluoropyruvate showed the same bonding as in chapter 8. The fluoropyruvate model created in chapter 8 using SKETCHER (Collaborative, 1994) was built into the positive electron density. Refinement featured the fluoropyruvate molecule trapped within the active site. After post refinement the B-factors of the fluoropyruvate were similar to that of the TPP molecule and surrounding residues. This showed a good stoichiometric agreement for occupancy of the fluoropyruvate and tyrosine mutation.

9.3 The H26Y mutant of *E. coli* transketolase

There have been many observed cases where a mutation has been performed within the active site of the transketolase molecule (Wikner *et al.*, 1995; Schneider and Lindqvist, 1998; Hibbert *et al.*, 2007), although only a few have been made upon the His-26 residue (Asztalos *et al.*, 2007). The His-26 residue involvement in the transketolase reaction concerns the binding of the of the donor/acceptor substrate. The Oδ of the substrate forms a hydrogen bond to the His-26 allowing further stability to the complex during the reaction.

The most recent of His-26 alteration concerns the mutation of the histidine to an alanine residue (Asztalos *et al.*, 2007), with the investigation observing a reduction in the overall formation of the TPP-intermediate. This observation occurred due to the removal of the residues ability to hydrogen bond with the hydroxyl of the substrate/donor, which is normally formed between His-26 and the substrate.

In this case study, the His-26 has been mutated to a tyrosine residue. The H26Y mutation of the *E. coli* enzyme transketolase was performed by Prof. John Ward of University College London (UK). Studies with the H26Y mutant conducted by Dr. Paul Dalby of University College London (UK) have discovered that when the H26Y mutant was used with glycolaldehyde and hydroxypyruvate in the biotransformation reaction, a different enantiomeric product was produced from that of the wild-type TK enzyme. The wild-type TK produces the L-enantiomer of erythrulose (Gyamerah and Willetts, 1997), however the product of the H26Y mutant's biotransformation was D-erythulose, with an enantiomeric excess of greater than 90 %. This finding could be used for the production of different sugar enantiomers of interest to the pharmaceutical industry. It is important to understand the structural basis of the enzymatic mechanism for the production of the D-erythulose.

The crystallographic structure of the H26Y TK mutant clearly shows the mutation of the histidine to the tyrosine residue (Fig. 9.1). The B-factors of the tyrosine residue (24.0 $Å^2$) are very similar to those of the adjacent side chain (21.0 $Å^2$), indicating stoichiometricity. The crystallographic data also illustrates a region of extremely weak electron density for the residues Thr-257, His-258, Asp-259, Ser-260 and His-261. When compared to the electron density map of the native transketolase, no electron deficiency was observed. With the exception of His-261, none of these residues are involved in the proposed catalytic mechanism. However, surprisingly there are no major alterations in the positions of the residues within the active site apart from the mutation H26Y (Fig. 9.2). The only noteworthy residue movement is that of the His-261, which effectively rotates the histidine head through 90°. This could be due to some form of structural rearrangement caused by the change in size of the newly mutated residue. Although confirmation of this hypothesis is difficult due to the weak electron density associated with the amino acids discussed previously. Apart from His-261 the key active site residues remain unaffected by the mutation.

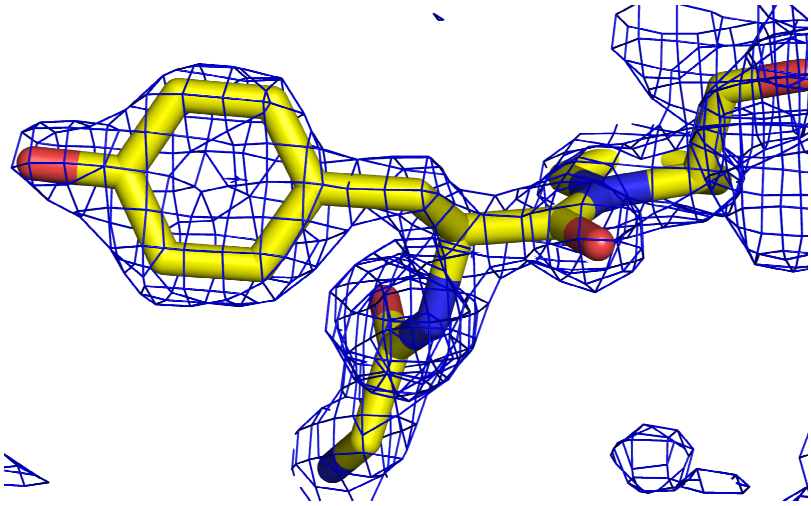


Figure 9.1 The F_o-F_c electron density maps for the H26Y mutation within the active site, also showing adjacent residues Gly-25 and Pro-27, contoured at 2.0σ (blue). Figure produced in PYMOL (DeLano, 2006).

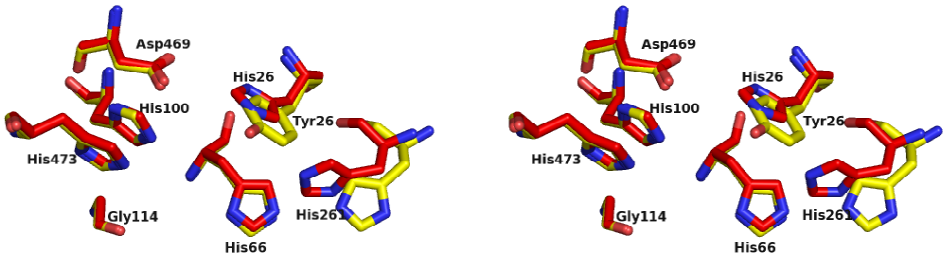
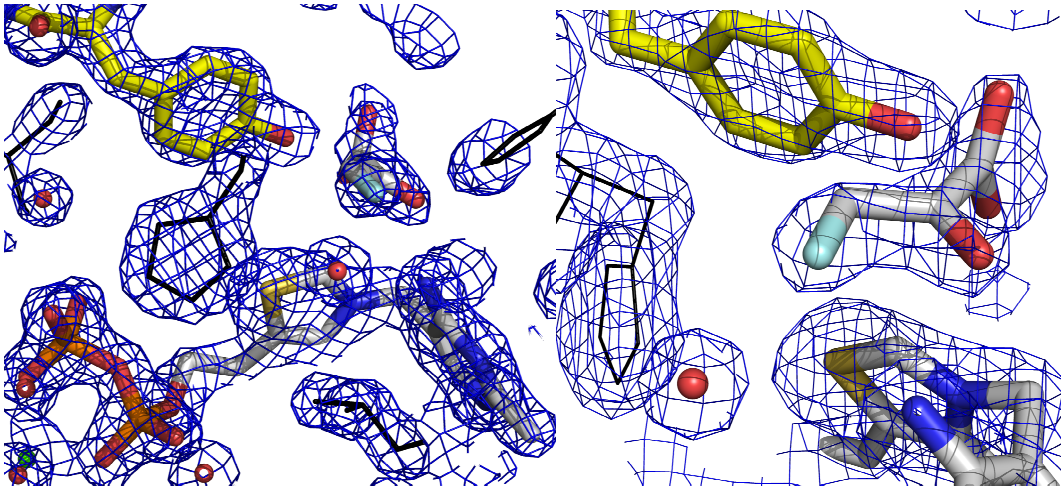


Figure 9.2 Diagram showing the superimposition of the active site residues of the H26Y structure (yellow) and the wild-type transketolase (red). Figure produced in PYMOL (DeLano, 2006).

9.4 Effect of the H26Y mutant transketolase on the fluoropyruvate Michaelis-Menten complex

As in chapter 8, the fluoropyruvate molecule is trapped in the vicinity of the thiazolium ring of the TPP within the active site (Fig. 9.3). The atoms of Cα and Cβ of the fluoropyruvate molecule are well defined but the carboxylate tail of the Cγ seems less defined compared to the previous structure. The coordinate position of the Cα and Cβ atoms are conserved when compared to that of the TK-fluoropyruvate complex. However, the carboxylate tail of the Cγ is shifted approximately 1.2 Å (Fig. 9.5). This result is possibly due to the increased size of the mutated residue, although validation of this proposition is difficult due to the weak electron density surrounding the carboxylate tail of the Cγ. As with the TK-fluoropyruvate complex, the H26Y-fluoropyruvate complex is formed through a series of hydrogen bonds and electrostatic interactions with the active site residues. The α-fluorine atom hydrogen bonds through His-100 (2.8 Å), whilst the Oβ atom forms a hydrogen bond with His-473 (2.4 Å), positioning the Cβ in within a distance of 3.0 Å to the C2 of the thiazolium ring. The carboxylate tail of the Cγ forms hydrogen bonds with the Tyr-26 mutated residue (2.2 and 2.5 Å) and Asp-469 (2.7 Å) (Fig. 9.4).



(a) **(b)**

Figure 9.3 $2F_o$-F_c electron density maps of the fluoropyruvate complex (silver) within the H26Y active site, contoured at 1.0σ (blue). (a) Depicts the fluoropyruvate complex (silver) and Calcium ion (green), with variant Tyr-26 (yellow), His-66, His-473, Leu-116, and Ala-29 (black). (b) Depicts the thiazolium ring of the TPP molecule and

fluoropyruvate molecule (silver), with variant Tyr-26 (yellow) and His-66 (black) residue. Figures produced in PYMOL (DeLano, 2006).

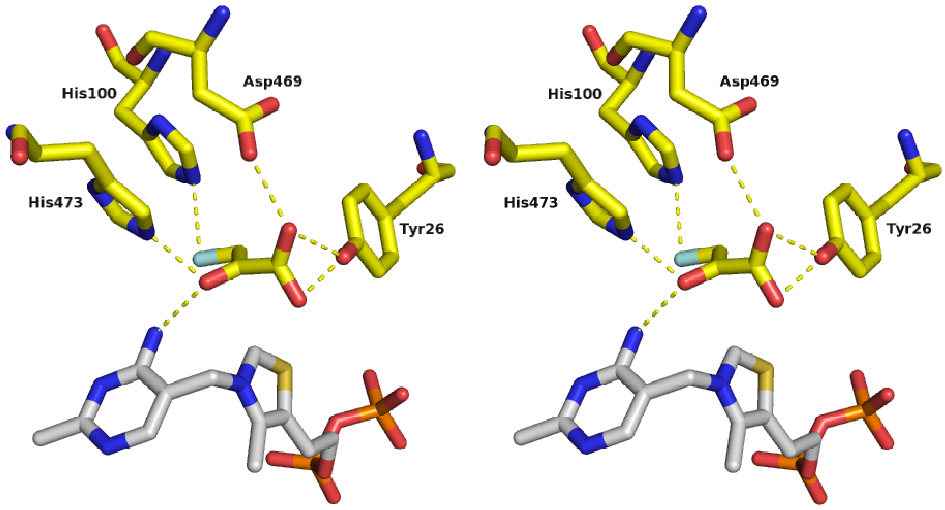


Figure 9.4 Divergent diagram showing the hydrogen-bonding network (yellow dashes) between the active site residues of the H26Y mutant and the fluoropyruvate molecule. Figure produced in PYMOL (DeLano, 2006)

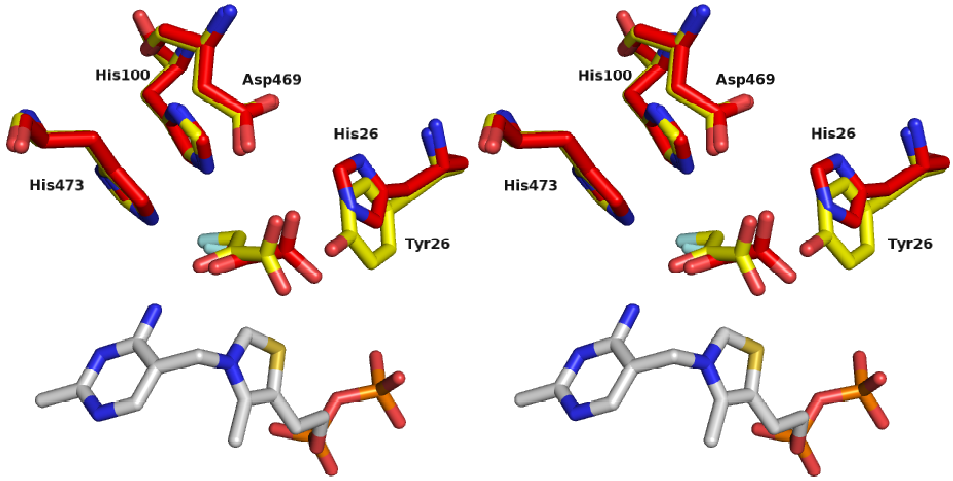


Figure 9.5 Divergent diagram showing the superimposition of the two fluoropyruvate Michaelis complexes, wild-type TK (Red) and the H26Y mutant (Yellow). Figure produced in PYMOL (DeLano, 2006).

9.5 Production of alternative enantiomer from the H26Y mutant transketolase

As discussed previously, the research conducted by Dr. Paul Dalby of University College London (UK) shows the H26Y mutant to produce an alternative enantiomeric product to that of the wild-type *E. coli* transketolase enzyme. The research performed pointed to the production of D-enantiomers with an enantiomeric excess of greater than 90 %, when the H26Y mutant conducted the biotransformation reaction using the standard reagents glycolaldehyde and hydroxypyruvate (unpublished work). The results gained from the resolved structure of the H26Y mutant complexed with fluoropyruvate, give some form of hypothesis as to why the production of the alternative enantiomer has been invoked by the H26Y mutant.

Studying the position and relative size of the tyrosine mutation, it is clear that one possible reckoning for the production of the D-enantiomer may be caused by the physical steric hindrance of the mutated residue, when compared to the native histidine residue. Further probing of the pathway and the chemistry into how a D-enantiomer is formed may bring about a conclusion to the mechanism (Fig. 9.6).

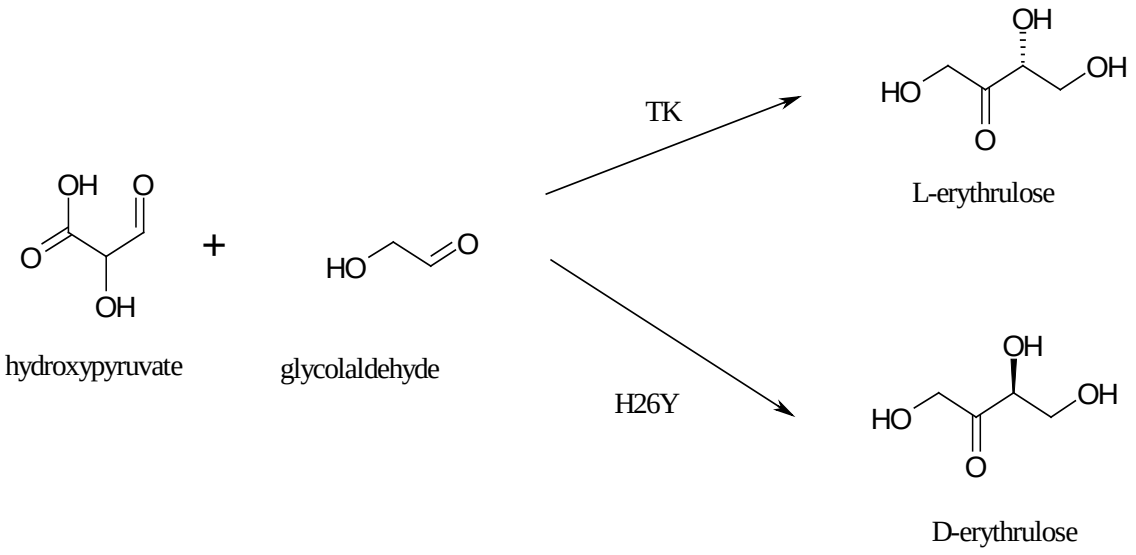


Figure 9.6 Diagram showing the transketolase biotransformation reaction using glycolaldehyde and hydroxypyruvate to form L or D erythrulose using the wild-type TK or the H26Y mutant TK, respectively.

Reflecting on the transketolase catalytic mechanism, it is plausible to suggest that the enantiomeric nature of the product is decided during the second stage of the reaction, when the acceptor substrate enters the active site. Modelling of the α,β-dihydroxyethyl–TPP intermediate into the active site of the H26Y-fluoropyruvate complex (Fig.9.7), indicates that the aldol acceptor would be forced to enter below the tyrosine residue in order to reach the dihydroxyethyl moiety. This would compel the acceptor to position itself unfavourably below the α,β-dihydroxyethyl carbanion, thereby creating the D-enantiomer, from the acceptor substrate.

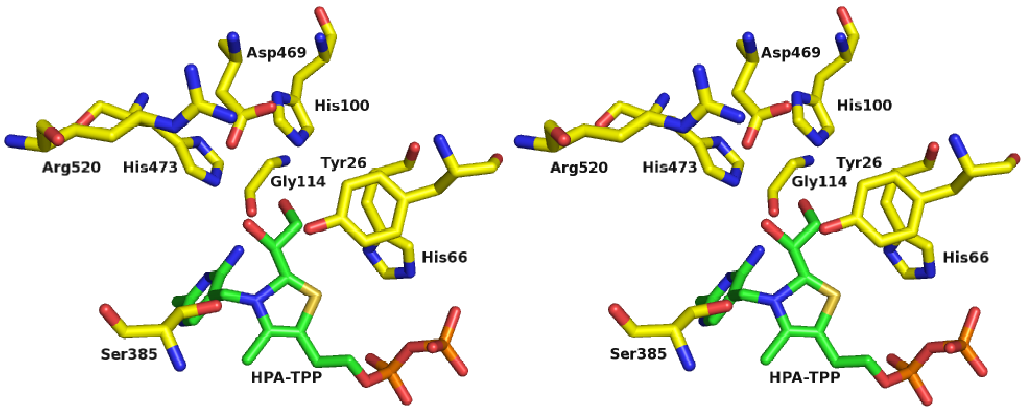


Figure 9.7 Divergent diagram to show the α,β-dihydroxyethyl–TPP intermediate (green) within the H26Y mutant TK active site (yellow). Figure produced in PYMOL (DeLano, 2006).

Other research conducted by Dr. Paul Dalby of University College London (London, UK) shows that the mutation of aspartic acid to a tyrosine residue at position 469 also produces the D-enantiomer in biotransformation reactions. It can again be hypothesised that this mutation forces the acceptor substrate to an unfavoured position below the α,β-dihydroxyethyl moiety, thus forming the alternative enantiomer.

Chapter 10

Purification, Characterisation and X-ray diffraction studies on *Pseudomonas putida* 2,5-Diketocamphane 1,2-Monooxygenase

10.1 Introduction

10.1.1 The baeyer-villiger oxidation

In 1899, Adolf von Baeyer and Victor Villiger discovered a reaction, which involved in the conversion of ketones into esters or cyclic ketones into lactones. The reaction was latter named the Baeyer-Villiger reaction (Baeyer and Villiger, 1899). During the reaction nucleophilic attack of the peroxy acid occurs to form a tetrahedral 'Criegee intermediate' (Fig. 10.1) (Criegee, 1948). The Criegee intermediate is unstable and undergoes intramolecular rearrangement. This results in the migration of an alkyl group to an oxygen atom, yielding the lactone product. A number of factors influence the rate of rearrangement and the migration preferences. These factors include the type of peroxy acid, the steric nature, the conformation, and electronic factors (Renz and Meunier, 1999). Due to the nature of the Baeyer-Villiger reaction, research has been conducted, with an aim for the synthesis of lactones and esters.

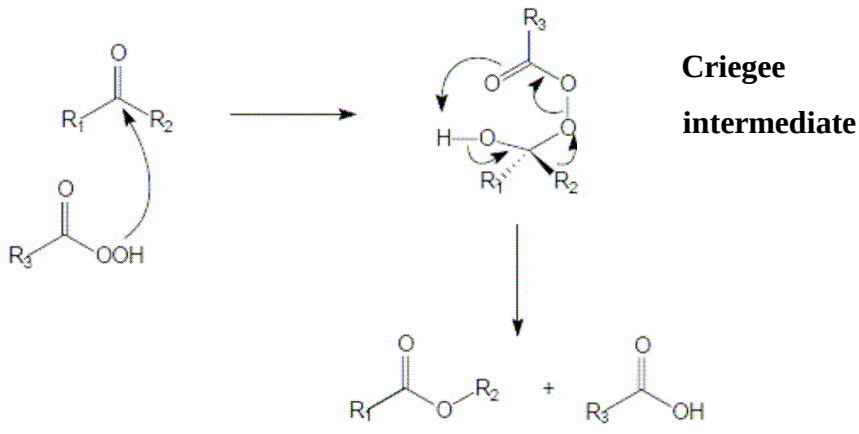


Figure 10.1 Mechanism of the Baeyer-Villiger oxidation by peroxy acids.

The first Baeyer-Villiger biotransformation reaction was reported in 1948 where fungi was used in the degradation of steroids (Turfitt, 1948). It has been reported that Baeyer-Villiger oxidation reactions have been observed in many different organisms acting on different pathways. Over the past decade a number of applications may be found in literature, such as the synthesis of iridoids and brassinosteroids in plants, and toxin synthesis in shellfish (Damtoft *et al.*, 1995; Wright *et al.*, 1996; Winter *et al.*, 1999; Kamerbeek *et al.*, 2003). A large number of enzymes have also been observed to catalyse the Baeyer-Villiger reaction, these enzymes have been characterized and named the Baeyer–Villiger monooxygenases (BVMO's).

10.1.2 The baeyer-villiger monooxygenases (BVMO)

Baeyer-Villiger monooxygenases (BVMOs) are NAD(P)H-dependent flavoproteins that catalyse a variety of oxidative reactions. They employ NAD(P)H and molecular oxygen to catalyse the insertion of an oxygen atom into a carbon–carbon bond of their substrates, and the other atom is reduced to water. Two types of Baeyer–Villiger monooxygenases have been classified depending on the nature of their flavin cofactor: Type I and Type II BVMOs.

The type I BVMOs are FAD-dependent proteins that use NADPH as source of electrons and consist of a single polypeptide chain. A well-known example of Type I BVMO is cyclohexanone monooxygenase (CHMO). It has been shown that CHMO is responsible for catalysing the conversion of cyclohexanone to epsilon-caprolactone. In sequence similarity analysis, it has been demonstrated that Type I BVMOs share sequence homology with CHMO and can be identified using a specific protein sequence motif, FXGXXXHXXXW(P/D) (Fraaije *et al.*, 2002).

In contrast, type II BVMOs are FMN-dependent proteins receiving electrons from an NADH source. They consist of α2β trimers, a dimeric oxygenating subunit and a NADH dehydrogenase. There are a limited number of type II BVMOs enzymes classified to date. These include 2,5-diketocamphane 1,2-monooxygenase (2,5-DKMO) and 3,6-diketocamphane 1,6-monooxygenase (3,6-DKMO) (Taylor and Trudgill, 1986; Jones *et al.*, 1993).

By combination of the crystallographic analysis and the kinetic data measured on several Baeyer–Villiger monooxygenases, the mechanism of the catalytic reaction has been proposed (Sheng *et al.*, 2001; Malito *et al.*, 2004). Two key intermediates are formed in the catalytic process of BVMOs: a flavin-peroxide and the "Criegee" intermediate. The flavin-peroxide is caused by the reaction of the reduced flavin with molecular oxygen. The Criegee intermediate is formed on attack of the flavin peroxide upon the substrate. Rearrangement of the Criegee intermediate brings in the oxygenated product and a hydroxy-flavin molecule. This hydroxy-flavin needs to be hydrolyzed to regenerate the oxidized flavin (Fig. 10.2).

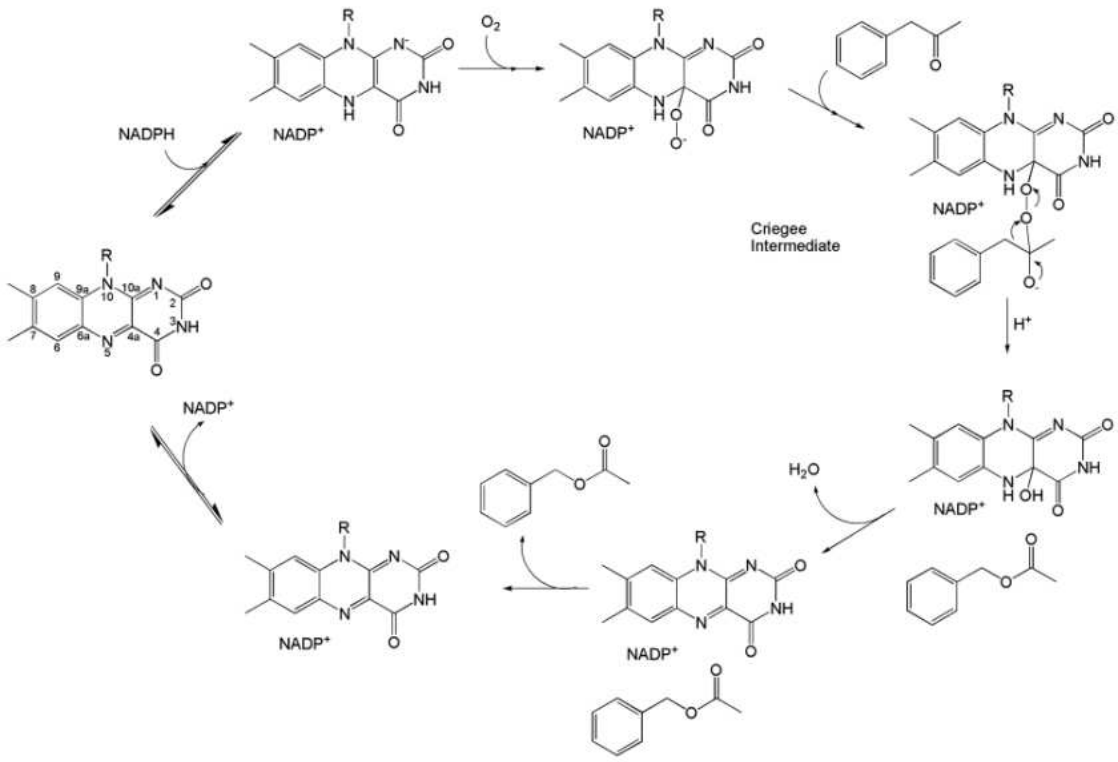


Figure 10.2 Schematic representation of the overall catalytic reaction of the Baeyer–Villiger monooxygenases. The atomic numbering of the flavin ring is shown on the left (corresponding to the initial step of the reaction) (Malito *et al.*, 2004; Sheng *et al.*, 2001).

In a structural study of the type I Baeyer–Villiger monooxygenase, the phenthylacetone monooxygenase from the thermophilic bacteria *Thermobifida fusca*, has been shown to consist of two domains, the FAD-binding domain and the NADP-binding domain. Each domain exhibits the typical dinucleotide-binding fold. The active site is placed in a cleft between two domains. An arginine residue plays a part in stabilizing the negatively charged flavin-peroxide and Criegee intermediates. The arginine residue is situated above the flavin ring and is predicted to exist in two positions. An 'in' and an 'out' position. The 'in' position is observed in the structure when stabilising the intermediate, whilst the 'out' position allows positioning of the NADPH for reduction of the flavin (Malito *et al.*, 2004). Alteration between the two positions allows the reaction to occur (Fig.10.3).

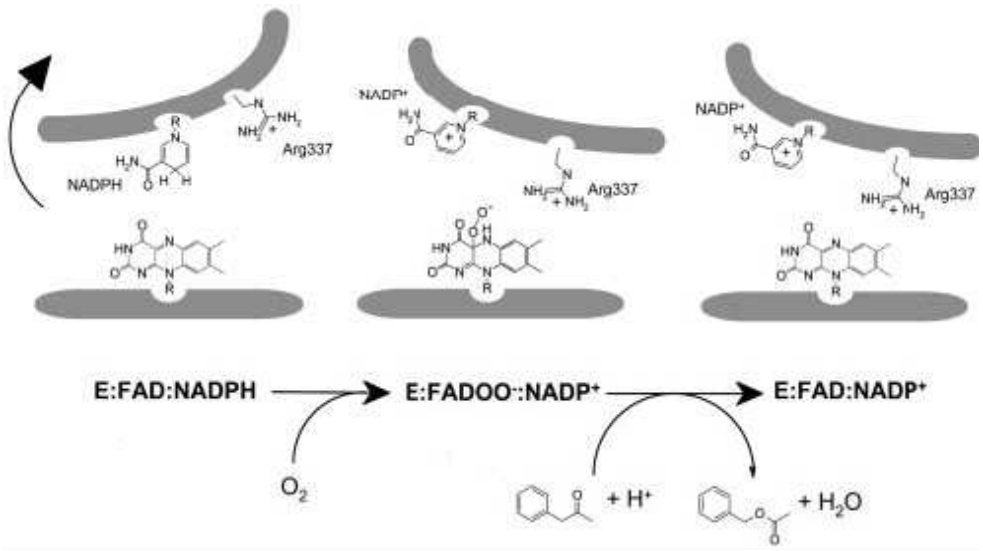


Figure 10.3 Schematic representation of the main structural features and conformational changes occurring in catalytic steps of the overall reaction of type I Baeyer–Villiger monooxygenases enzyme, taken from (Malito *et al.*, 2004).

10.1.3 Baeyer–Villiger monooxygenases involved in camphor degradation pathway

An example of BVMOs acting on the degradation pathway in microorganisms may be seen in *Pseudomonas putida* ATCC 17453 (NCIMB 10007). This strain is able to grow upon camphor as its sole carbon source. Camphor is a natural bicyclic ketone which exists in two enantiomeric forms, (+)-camphor and (-)- camphor. It has been shown that *P. putida* ATCC 17453 carries a plasmid named 'CAM' plasmid which encodes enzymes involved in the pathway for catabolism of (+) or (-)-camphor (Rheinwald *et al.*, 1973). This would allow the *P. putida* ATCC 17453 to grow on both (+) and (-) camphor. In order to degrade the (+/-) camphor the *P. putida* uses the type II BVMOs. The 2,5-DKMO is used for ring expansion of the (+) camphor and the 3,6-DKMO for the (-) camphor.

10.1.4 The 2,5-diketocamphane 1,2-monooxygenase

The 2,5-diketocamphane 1,2-monooxygenase is made up by an oxygenase component and a flavin reductase component. The oxygenase component (78 kDa) is composed of two identical subunits and binds one FMN molecule. The other component is an NADH-dependent reductase (36 kDa), which is attached to the FMN very weakly (Kd 0.45 μM) and responsible for reducing the FMN (Taylor and Trudgill, 1986). It is generally agreed that the components that form the active complex are only weakly associated and dissociate in a reversible manner during purification or on dilution (Taylor and Trudgill, 1986).

10.1.5 The 3,6-diketocamphane 1,6-monooxygenase

The structure of the oxygenating subunit of the 3,6-DKMO from *P. putida* ATCC 17453 has been solved recently (McGhie, 1998; Isupov and Lebedev, 2008). Using crystals grown from sodium acetate and polyethylene glycol (PEG) 8,000 using the

vapour-phase diffusion method. The crystals belong to $P2_12_12_1$ space group, with cell dimensions a = 55.8, b = 94.5 and c = 163.7 Å. The crystal diffracted to a 2.8Å resolution.

10.1.6 Project aims

The aim of this project was to purify the recombinant 2,5-DKMO from *P. putida.* The purified recombinant 2,5-DKMO protein will be used for crystallisation. Any crystals obtained will be subjected to X-ray crystallography studies in order to determine the structure of the 2,5-DKMO enzyme. Results from these studies will enhance knowledge of the structure of active sites and thus rationalise the reaction mechanism of Type II BVMOs.

10.2 Purification and characterisation of the recombinant 2,5-diketocamphane monooxygenase

The *P. putida* 2,5-Diketocamphane Monooxygenase enzyme was received from Prof. John Ward of University College London (UK), as a cell paste, which was stored at 4°C until needed.

10.2.1 Materials and methods

10.2.1.1 Buffers

Buffer H	21 mM KH_2PO_4/K_2HPO_4 buffer pH 7.1, 0.1 mM EDTA, 6 mM β-Mercaptoethanol, and $1x10^{-5}$ M PMSF.
Buffer I	21 mM KH_2PO_4/K_2HPO_4 buffer pH 7.1, 0.1 mM EDTA, 6 mM β-Mercaptoethanol, and $1x10^{-5}$ M PMSF. Containing 2 M KCl.
Buffer J	21 mM KH_2PO_4/K_2HPO_4 buffer pH 7.1, 0.1 mM EDTA, 6mM β-Mercaptoethanol, and $1x10^{-5}$ M PMSF. Containing 0.5 M NaCl.

10.2.1.2 Preparation of lysed cell extract

The *P. putida* 2,5-DKMO cell paste was suspended in Buffer H. The cells were broken by sonication using a Sanyo Soniprep 150 MSE at full power for six-30 seconds intervals on ice, the resultant extract was centrifuged at 12500 rpm for 20 minutes to remove the cellular debris, as described in section 2.1.10.2.

10.2.1.3 Protamine sulfate precipitation

The cleared yellow supernatant (due to flavin present) was subjected to a protamine sulfate precipitation stage, in order to remove the nucleic acids. The protamine sulfate (10 % [w/v] stock solution) was added to the cell suspension to give a final concentration of 0.05 % [w/v]. The resulting solution was stirred for 30 minutes at 4 °C and then centrifuged at 12000 g for 20 minutes to remove precipitated nucleic acids. The pellet was discarded and the supernatant retained and stored at 4 °C.

10.2.1.4 Hydrophobic chromatography

The supernatant from section 10.2.1.3 was subjected to hydrophobic chromatography and was performed using a Pharmacia Phenyl Sepharose™ CL-4B (column volume 40 cm^3) with a flow rate of 3.5ml/min. KCl was added to the supernatant (section 10.2.1.3) yielding a final 2M KCl concentration. The resultant solution was loaded on to the phenyl sepharose column, with a decreasing concentration gradient of Buffer I over 600ml. 10 ml fractions were collected and run on SDS-PAGE. The cleanest fractions were combined and dialysed against buffer H.

10.2.1.5 Ion-exchange chromatography

The 2,5-DKMO fractions from section 10.2.1.4 were concentrated (section 2.1.6) to 50 ml and subjected to ion-exchange chromatography. The ion-exchange chromatography was performed using a Pharmacia FFQ anion exchange column (column volume 80 cm^3) with a flow rate of 4 ml/min, as described in section 2.1.10.4 (using respective Buffers H and J, with no ammonium sulfate precipitation)

10.2.1.6 Gel filtration chromatography

The fractions from section 10.2.1.5 were concentrated (section 2.1.6) to 2 ml and subjected to gel filtration chromatography. The gel filtration chromatography was performed using a Pharmacia Superdex 200 gel filtration column (column volume 120 cm^3) with a flow rate of 1ml/min, as described in section 2.1.10.5 (using respective Buffer H).

10.2.1.7 Dynamic light scattering analysis

The Dynamic Light scattering analysis was conducted using 2,5-DKMO after purification to ensure homogeneity. Refer to section 2.2.2 in the materials and methods chapter for protocol.

10.2.2 Purification and characterisation results for recombinant 2,5-diketocamphane monooxygenase

10.2.2.1 Hydrophobic chromatography

Phenyl sepharose hydrophobic chromatography was the first purification column used in the homologous purification of the 2,5-DKMO enzyme. The elution of the column produced a broad peak over 100 ml (Fig. 10.4). SDS-PAGE gel electrophoresis showed the 2,5-DKMO monomer with a molecular weight of 40 kDa (Fig. 10.5).

10.2.2.3 Ion-exchange chromatography

FFQ Sepharose anion exchange chromatography was the second purification column used in the homologous purification of the 2,5-DKMO enzyme. The elution of the column produced two significant peaks, the first peak was eluted at approximately 0.185 M NaCl and the other at 0.225 M NaCl (Fig. 10.6). The latter of the two contained the 2,5-DKMO enzyme. SDS-PAGE gel electrophoresis showed the 2,5-DKMO monomer with a molecular weight of 40 kDa (Fig. 10.7).

10.2.2.4 Gel-filtration chromatography

Gel filtration chromatography was the final purification column used in the homologous purification of the 2,5-DKMO enzyme. The column produced one significant peak, eluting at approximately 78 ml with a UV_{280} absorbance of 1.80 Au. (Fig. 10.8) The peak of the 2,5-DKMO enzyme was assessed for purity on an SDS-PAGE gel (10.9). Calculation of the K_{av} and the log molecular weight from the elution volume of the 2,5-DKMO enzyme (see appendix I) estimate that the protein with an elution of 78 ml had a molecular mass of approximately 80 kDa, which would indicate that the 2,5-DKMO protein was in a dimeric state.

10.2.2.5 Dynamic light scattering analysis

The analysed 2,5-DKMO sample showed a monomodal distribution with no polydispersity (Fig. 10.10). This implies that the 2,5-DKMO sample is homogeneous.

The other data sourced from the analysis shows the 2,5-DKMO to have an estimated size of 40 kDa, which would imply the enzyme is in its monomer form. However, as discussed in section 4.3.1 the size estimation capabilities of this apparatus are debatable.

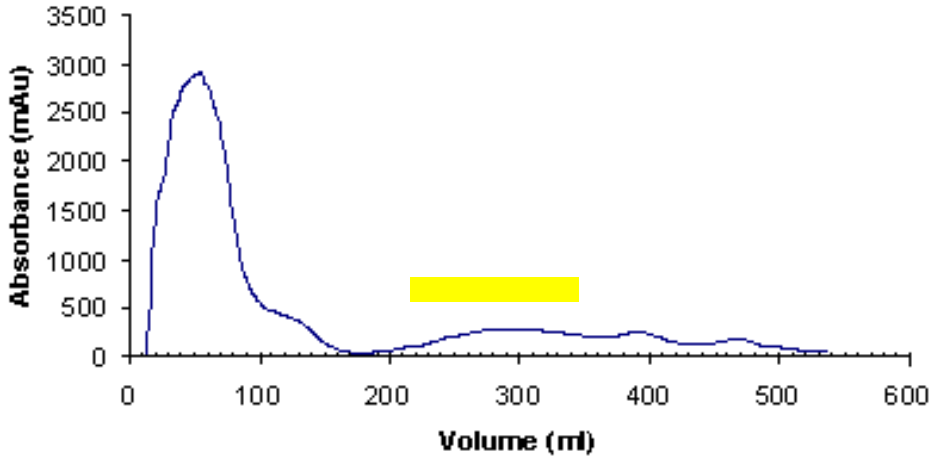


Figure 10.4 Chromatography trace of Phenyl sepharose column for 2,5-DKMO. The absorbance at 280 nm is shown in blue, and active volume shown in yellow.

Figure 10.5 SDS PAGE gel showing the Phenyl sepharose purification of 2,5-DKMO. The molecular weight markers (Bio-Rad) on the left, with protein post Phenyl sepharose purification on the right.

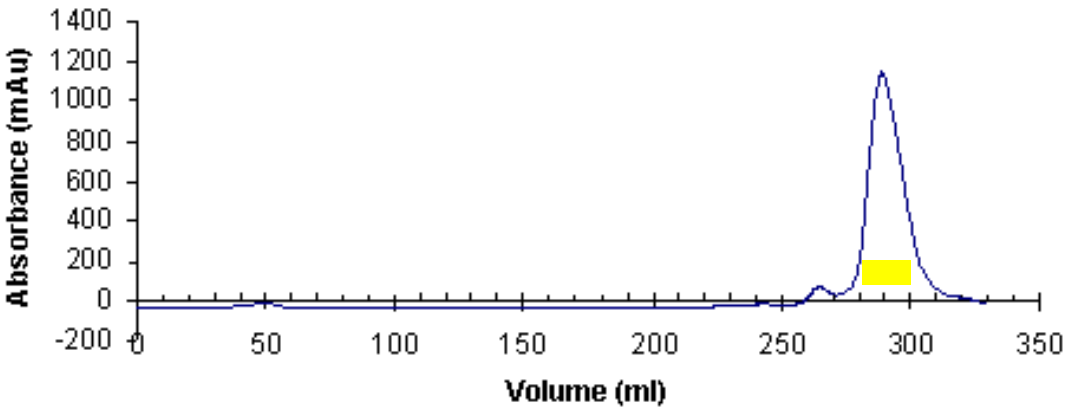


Figure 10.6 Chromatography trace of FFQ column for 2,5-DKMO. The absorbance at 280 nm is shown in blue, and active volume shown in yellow.

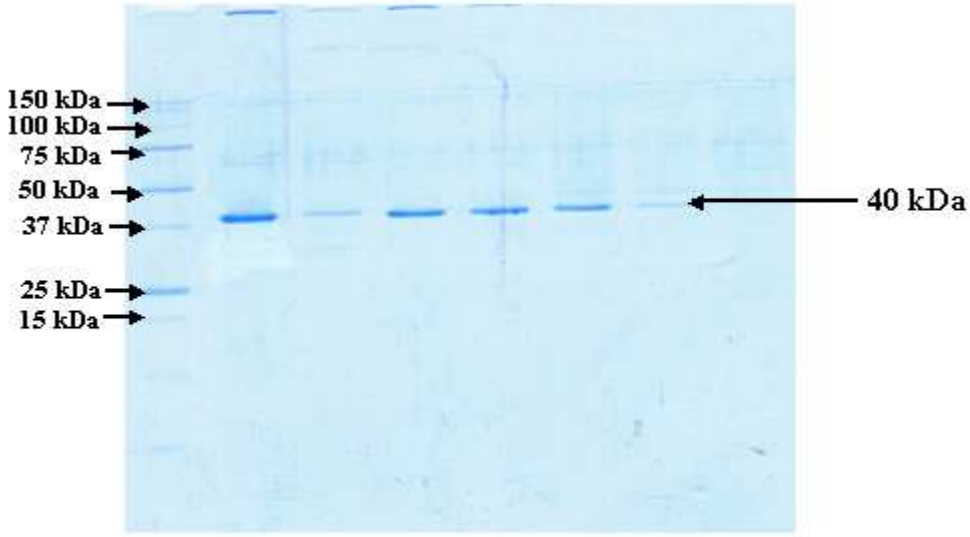


Figure 10.7 SDS PAGE gel showing the FFQ purification of 2,5-DKMO. The molecular weight markers (Bio-Rad) on the left, with protein post FFQ purification on the right.

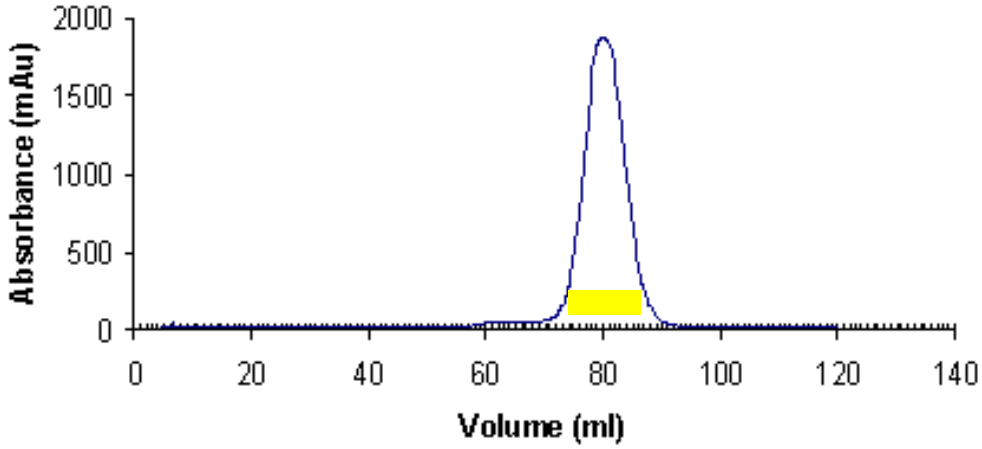


Figure 10.8 Chromatography trace of GF column for 2,5-DKMO. The absorbance at 280 nm is shown in blue, and active volume shown in yellow.

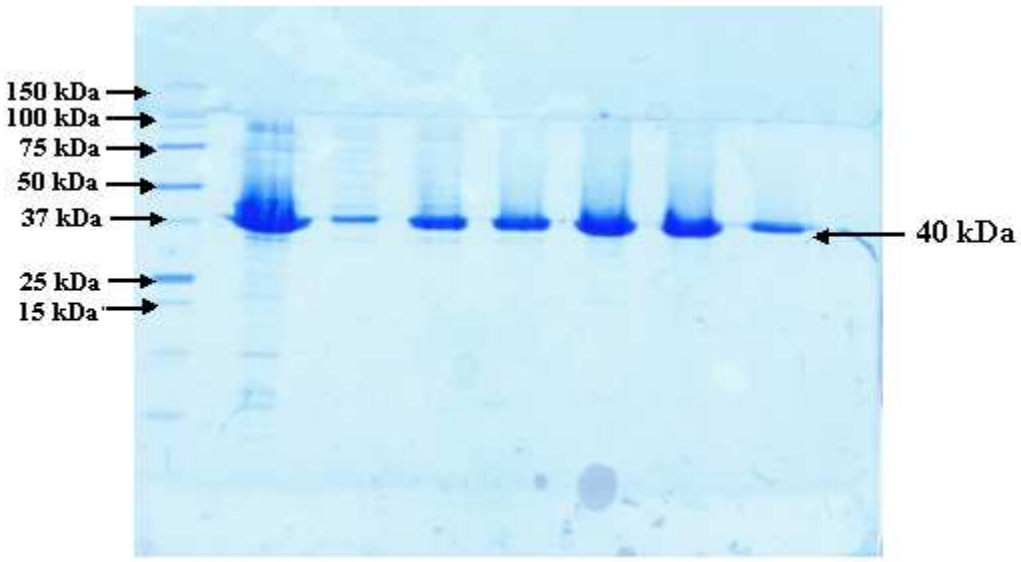
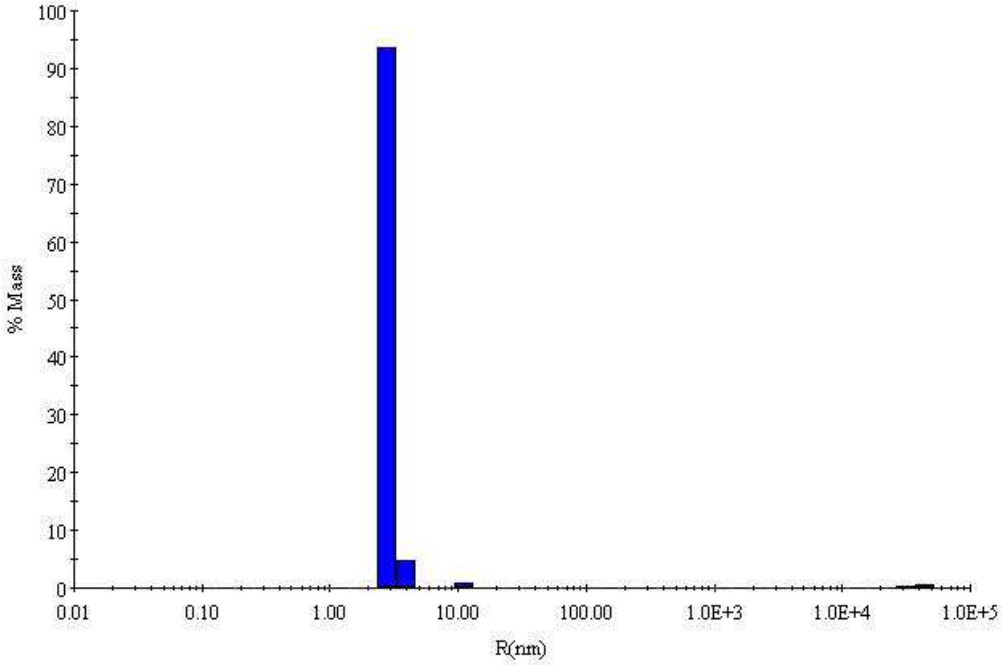


Figure 10.9 SDS PAGE gel showing the GF purification of 2,5-DKMO. The molecular weight markers (Bio-Rad) on the left, with protein post GF purification on the right.



Item	R	Pd	%Pd	MW-R	%Int	%Mass
	(nm)	(nm)		(kDa)		
☑ Peak 1	2.9	0.2	8.6	39.586	28.4	99.2
☑ Peak 2	11.2	0.0	0.0	959.441	12.1	0.7
☑ Peak 3	57.7	8.3	14.4	44467.000	43.7	0.0

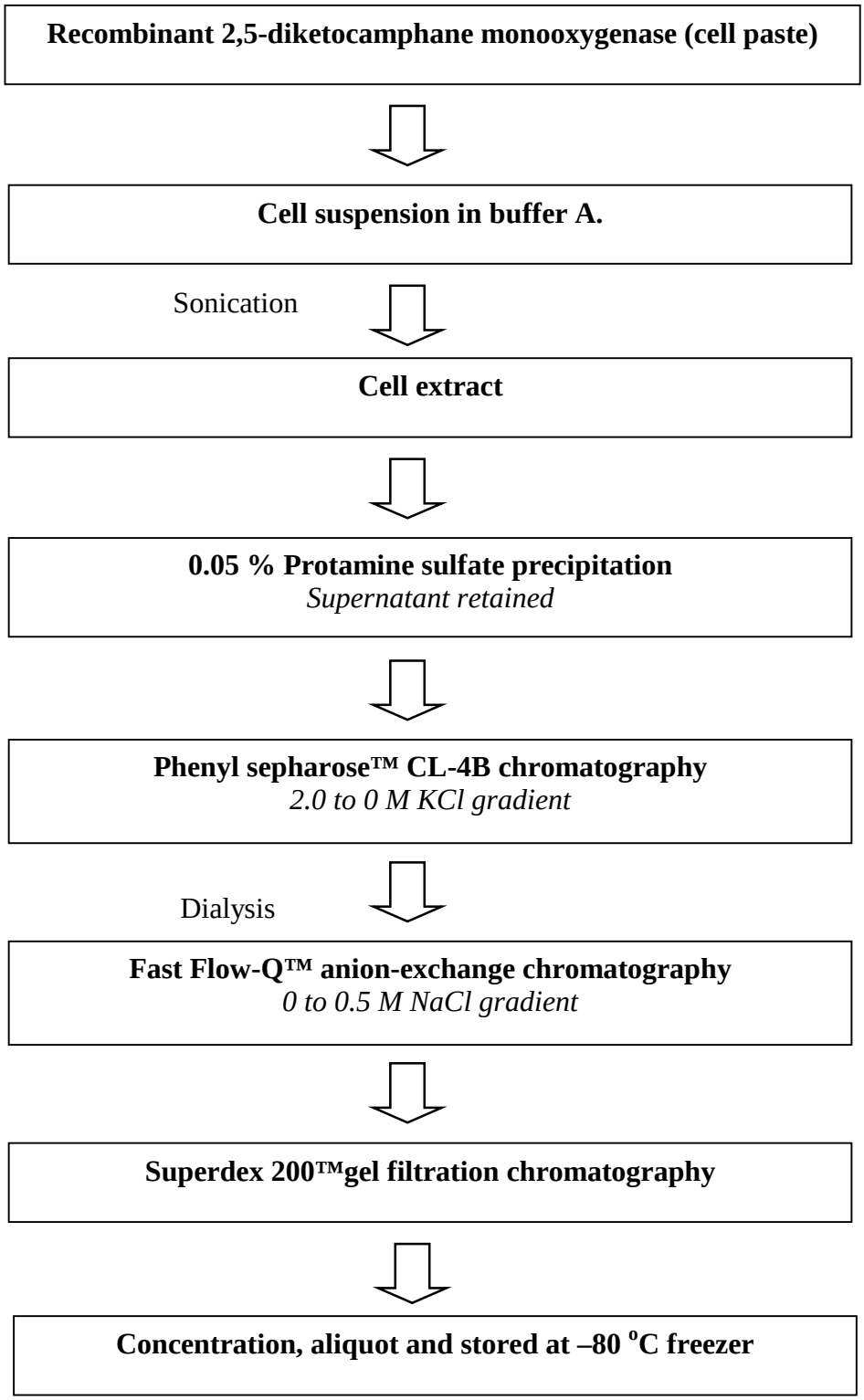
Figure 10.10 Graph produced from dynamic light scattering for the 2,5-DKMO sample.

10.2.3 Discussion

A purification protocol for the 2,5-DKMO has been established (section 10.2.4) with 13.5 mg of pure 2,5-DKMO protein produced per litre of LB culture.

In the early stages of the purification of the 2,5-DKMO protein an ammonia sulfate precipitation step was used. Firstly, in order to separate the cell proteins post cell lysis, and again to prepare the enzyme for gel filtration post FFQ. The subsequent supernatant produced had no yellow colour, indicating the absences of the FMN cofactor. On study of the literature, it was found that by removing the ammonium sulfate step in purification would allow the protein to retain the yellow FMN cofactor (Tanner and Hopper, 2000). A purification protocol without the ammonia sulfate precipitation was developed, which resulted in a yellow supernatant throughout the purification process.

10.2.4 Purification summary of 2,5-diketocamphane monooxygenase



10.3 Crystallisation of recombinant 2,5-diketocamphane monooxygenase

The purified 2,5-DKMO protein from section 10.2.2.4 was used in protein crystallisation studies.

10.3.1 Materials and methods

10.3.1.1 Sample preparation

The purified 2,5-DKMO protein was prepared as described in section 2.3.1 using a protein concentration of 10 mg/ml in buffer H. The 2,5-DKMO protein was incubated in a 125 mM FMN solution at 4 °C for 1 hour prior to crystallisation experiments.

10.3.1.2 Crystallisation trials

The prepared protein solution was screened against NeXtal 'The pHclear™' (Qiagen, UK) Structure Screen 1 & 2 (Molecular Dimensions Ltd., Cambridgshire, UK.), Crystallization Basic Kit for Proteins 82009, and Crystallization Extension Kit for Proteins 70437 (Sigma-Aldrich, UK) (See details in appendix II) using the Oryx6 system crystallisation robot (Douglas Instruments, Cambridge, UK) using the method described in section 2.3.2.2.

10.3.1.3 Crystallisation results for recombinant 2,5-Diketocamphane Monooxygenase

Initial crystallisation trials produced multiple results within a range of alternative conditions. Crystals from Crystallization Basic Kit for Proteins 82009, Crystallization Extension Kit for Proteins 70437 and Structure Screen 1 & 2 (MDL 1 & 2) grew within a 10 day period, and varied in shape and form. The results obtained from these trials are summarised in table 10.1.

Two crystals suitable for X-ray diffraction studies were grown in:

A) 0.2 M Calcium acetate hydrate, 0.1 M Na Cacodylate pH 6.5, and 18 % w/v PEG 8000, (Fig.10.11).

B) 0.2 M Ammonium sulfate, 0.1M MES pH 6.5, and 30 % w/v PEG monomethylether 5000, (Fig. 10.12).

Both crystals were yellow indicating the presence of the FMN cofactor, which had been preserved throughout purification.

Crystals from both conditions diffracted using the 'in house' X-ray source (section 6.2.3). A complete data set was collected from the crystal grown in condition A.

Kit	Conditions	Comment
MDL1	0.2M Sodium acetate trihydrate, 0.1 M sodium cacodylate pH 6.5, 30 % w/v PEG 8000	Crystal clusters
	0.2 M Zinc acetate dihydrate, 0.1M Na cacodylate pH 6.5, 18 % w/v PEG 8000	Small crystal clusters
	0.2 M Calcium acetate hydrate, 0.1 M Na cacodylate pH 6.5, 18 % w/v PEG 8000	Crystal clusters
MDL2	0.2 M Ammonium sulfate, 0.1M MES pH 6.5, 30 % w/v PEG monomethylether 5000	Big chunky crystal
	0.2 M Ammonium sulfate, 0.1 M Sodium acetate pH 4.6, 30 % w/v PEG monomethylether 2000	Fine needles
82009	sodium-citrate 0.2 M, TRIS-HCl (pH 8.5) 0.1 M, PEG 400 30 %	
	HEPES sodium-salt (pH 7.5) 0.1 M, Li-sulfate 1.5 M	Microcrystals
	HEPES sodium-salt (pH 7.5) 0.1M, PEG 400 2 %, ammonium sulfate 2.0 M	Needles
	sodium-citrate (pH 5.6) 0.1 M, 2-Propanol 20 %, PEG 4000 20 %	Fine needles
	Zn-acetate 0.2 M, sodium-cacodylate (pH 6.5) 0.1 M, PEG 8000 18 %	Long crystal
	Ca-acetate 0.2 M, sodium-cacodylate (pH 6.5) 0.1 M, PEG 8000 18 %	Rectangular crystal
	Ni(II)-chloride 0.01 M, Tris-HCl (pH 8.5) 0.1 M, PEG monomethylether 2000 20 %	Cubic crystal

Table 10.1 Positive results for crystallisation trials for 2,5-DKMO using the Sigma Aldrich crystallization Basic Kit for Proteins 82009 and MDL structure screen 1 & 2.

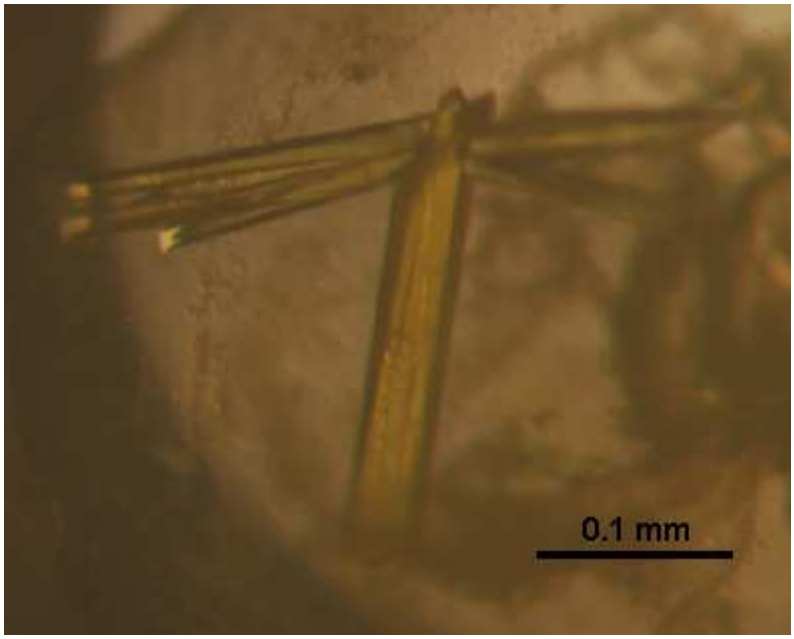


Figure 10.11 Photograph of the 2,5-DKMO crystal grown from 0.2 M calcium acetate hydrate, 0.1 M sodium cacodylate pH 6.5, and 18 % w/v PEG 8000.

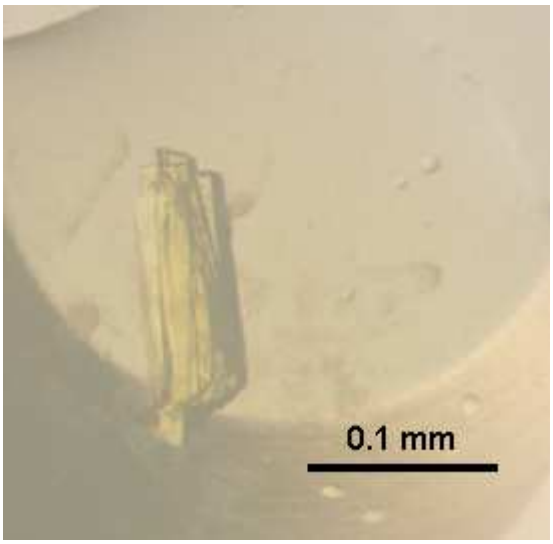


Figure 10.12 Photograph of the 2,5-DKMO crystal grown from 0.2 M ammonium sulfate, 0.1 M MES pH 6.5, and 30 % w/v PEG monomethylether 5000.

10.4 X-ray diffraction studies of recombinant 2,5-diketocamphane monooxygenase

The 2,5-DKMO crystal grown in 0.2 M calcium acetate hydrate, 0.1 M Na cacodylate pH 6.5, and 18 % w/v PEG 8000 was frozen under silicon oil (MDL, UK) as no suitable cryoliquor could be found after multiple trials. The crystal diffracted to 1.73 Å resolution 'in house'. Indexing and scaling (see section 6.2.5) revealed the crystal belonged to the primitive monoclinic space group $P2_1$ with unit cell parameters a = 51.9 Å, b = 90.1 Å, c= 75.4 Å, β=100.5°. 67219 unique reflections were observed and data set had 94.4 % completeness. The statistics for this data is shown in table 10.2.

	2,5-DKMO
Space group	$P2_1$
Unit Cell (a,b,c Å)	51.9; 90.1, 75.4 β=100.5
Wavelength	1.542
Resolution range (Å)	15 – 1.73 (1.76 - 1.73)
Number of unique reflections	67219
Completeness	94.4 (88.0)
Redundancy	4.5 (3.6)
B-Factor of data from Wilson Plot ($Å^2$)	35.5
I > 3σ(I) (%)	79.3 (35.7)
(I)/σ (I)	25.2 (3.2)
R_{sym} (%)	5.5 (45.1)

Table 10.2 Data processing statistics for the 2,5-DKMO data set ($R_{sym} = \Sigma_h\Sigma_J|<I_h>-I_J(h)|/ \Sigma_h\Sigma_J I(h)$. Where (h) is the intensity of the reflection h. Σ_h is the sum over all reflections and Σ_J is the sum over J measurements of the reflection. Values given in the parentheses are given for the highest shell resolution.

Molecular replacement studies were carried out with the program MOLREP (Vagin and Teplyakov, 1997) using chain A of 3,6 DKMO as a model, which has 45 % amino acid sequence identity with the 2,5-DKMO (Fig.10.13). The cross-rotation function was calculated for the dimeric model with a radius of 28 Å within a resolution range of 15-3 Å. The rotation solutions were the top peaks between 9.53 and 9.0 σ with a noise peak at 4.0 σ. The correlation coefficient for the translation solution was 37.2 %, while the correlation coefficient for the wrong position was not higher than 24.0 %. The subunit model was subjected to rigid-body refinement implemented in MOLREP. The resulting model was refined in REFMAC 5.2 (Murshudov *et al.*, 1997), resulting in a model with R-factor 35.3 and R_{FREE} 40.5. The model was then rebuild using the sequence of 2,5-DKMO with MOLREP. The 2,5-DKMO model was then refined using REFMAC 5.2, with manual model building performed in COOT (Emsley and Cowtan, 2004). The resultant 2,5-DKMO model yielded an R-factor 19.3 and R_{FREE} 24.5. A further data set was collected at Daresbury SRS on beamline 10.1, which diffracted to 1.4 Å. The data set is currently being processed and refined by Dr. Rob Gibson in Prof. Littlechild's group.

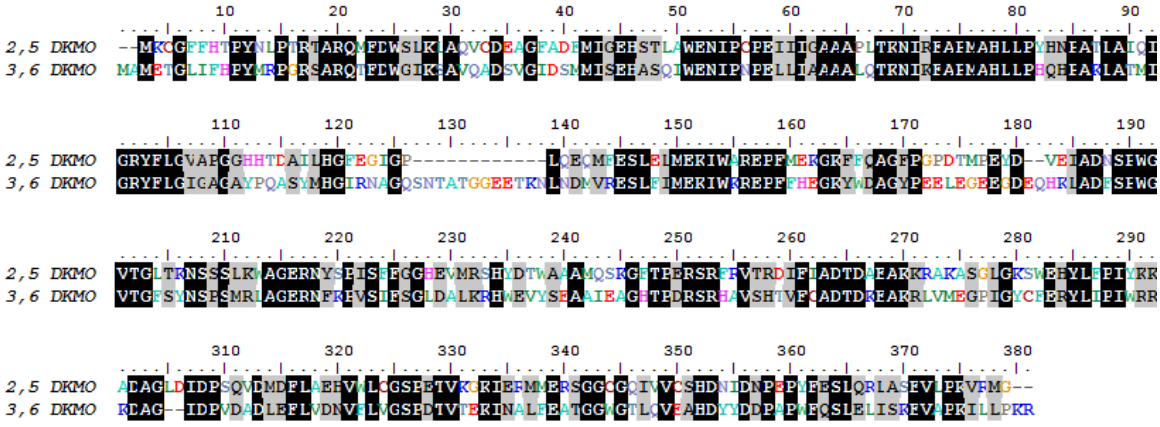


Figure 10.13 Sequence alignment of the 2,5-DKMO and 3,6 DKMO from *P. putida*.

Chapter 11

Summary and further work

11.1 Summary

The work presented in this thesis has described the purification, characterisation, crystallisation and structural determination of the *E. coli* TK and active site mutant H26Y. The characterisation and structural determination studies of the *E. coli* TK and H26Y mutant were conducted with substrate hydroxypyruvate and potential inhibitor fluoropyruvate to probe the catalytic mechanism of the enzyme. The purification and crystallisation of the *E. coli* TK active site mutant D469Y was also carried out. In addition, the purification, crystallisation, characterisation and structural studies of the commercially useful oxygenating enzyme, *P. putida* 2,5-DKMO was performed.

The recombinant *E. coli* TK was easily purified to homogeneity in four stages. Under reproducible conditions crystals grew easily within two weeks and diffracted to a very high resolution. Characterisation and kinetic studies of the *E. coli* TK with its industrial substrate hydroxypyruvate gave a K_m of 40 mM, the same value as seen previously. However, investigations using the potential inhibitor fluoropyruvate showed slight competitive inhibition towards the donor ketol hydroxypyruvate, but significantly higher competitive inhibition towards the acceptor aldol glycolaldehyde. These results implied that even though fluoropyruvate competed with hydroxypyruvate for the thiazolium ylide, the enamine was still produced. The fluoropyruvate continued its competitive inhibition once the enamine was formed, but was more effective against the glycolaldehyde thereby inhibiting the TK reaction.

The recombinant *E. coli* TK crystals were complexed with the substrate hydroxypyruvate and potential inhibitor fluoropyruvate producing two further structures. The native TK and the TK-HPA structures were resolved to a 1.18 and 1.05 Å resolution respectively, whilst the novel TK-FPA structure was solved to 1.60 Å resolution, produced a snapshot image of the ketol donor prior to formation of the enamine. The three structures were solved by molecular replacement using the original *E. coli* TK structure 1QGD as a model. The three structures were of high quality,

refining to R-factors of 12.9, 12.4, 17.2 % and R-free of 15.7, 14.9, 19.1 % for the TK, TK-HPA and TK-FPA models respectively. Examination of the three structures has helped to learn more about the mechanism for the donor stage of the transketolase mechanism. Analysis of the TK-HPA structure showed that the hydroxypyruvate substrate was covalently bound to the thiamine pyrophosphate with the creation of the α,β-dihydroxyethyl–thiamine pyrophosphate complex within the enzyme active site. Whilst analysis of the novel TK-FPA structure, showed an unusual angle of 114° for nucleophilic attack of the thiamine pyrophosphate prior to the formation of the α,β-dihydroxyethyl–thiamine pyrophosphate complex. The angle of 114° varies from the accepted Burgi-Dunitz angle of 109.5° for nucleophilic attack. However as commented on in chapter 8, theses conclusions are rather optimistic due to the occupancy and poor electron density of the FPA and questionable diffraction data results. In all of the TK structures, a series of water molecules were observed between the two active sites of the enzyme. They are proposed to act as a 'proton-wire' and are linked via a network of hydrogen bonds. The proton-wire is thought to facilitate proton transfer between the two thiamine pyrophosphate molecules, thereby providing a method of communication between the two active sites of the enzyme.

Structural analysis of the *E. coli* TK mutant H26Y complexed with fluoropyruvate resulted in a 1.66 Å resolution structure. The H26Y mutant structure was again solved by molecular replacement as with the other three TK structure. The complexed H26Y mutant structure was well refined to an R-factor of 14.6 % and R-free of 17.5 %. Comparisons of the four structures, aside from the mutated tyrosine residue, showed only subtle movement within the active site. The structural differences observed between the mutant and native TK structures have allowed an hypothesis to be developed with regard to the formation of the alternative enantiomeric product with the mutant enzyme. The *E. coli* TK mutant D469Y was easily purified, and small enzyme crystals could be obtained, however these were not of suitable quality for X-ray diffraction studies.

The *P. putida* 2,5-DKMO was purified to homogeneity, whilst maintaining the cofactor FMN by removal of the ammonium sulfate precipitation step. Crystals were grown in high molecular weight PEG precipitant. The crystals diffracted to 1.73 Å, and the final 2,5-DKMO structure was refined to an R-factor of 19.3 % and R-free of 24.5 %.

11.2 Further work

There is much research continuing over the globe into the TK enzymes and thiamine chemistry as a whole, showing that this field of research is still very alive. From the studies conducted during this thesis, there are many routes in which this research may be continued.

In the short term work on the *E. coli* TK mutant D469Y could be continued in order to understand the formation of the alternative enantiomeric product. Also further structural studies and comparison between the mutant H26Y structure and it's fluoropyruvate complexed structure to determine whether the disorder between the residues Tyr-257 and His-261 observed in H26Y-FPA structure, is also observed in the un-complexed form. This would determine the significance of the disorder with regard to the change in enantiomeric specificity.

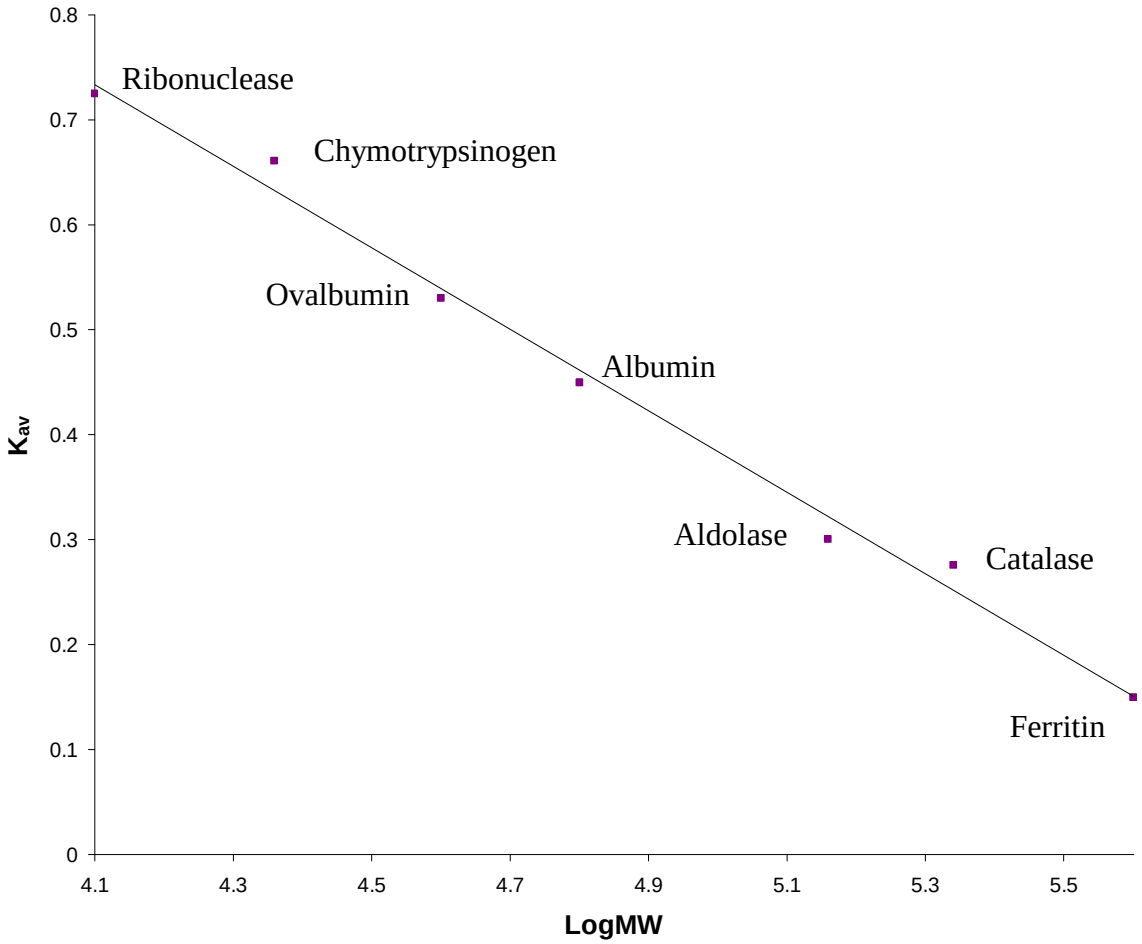
In the longer term, investigations into an extended range of substrates may be carried out, based on the findings from this thesis, by taking into consideration the position of the donor substrate during catalysis, and the residues involved. With the knowledge gained from the structural studies of the H26Y-FPA complex, other mutations maybe conducted in order to allow the TK to accept larger or different types of substrates producing the D-enantiomeric product, which would be useful for the pharmaceutical industry.

The high-resolution structural studies of the *E. coli* TK has enhanced our understanding of the catalytic mechanism of this enzyme. However, during refinement of the high-resolution structure interesting observations were made. A small area of electron density was observed in between the two oxygen atoms of the glutamate residue at position 411. Coupled with the understanding of the deprotonation of the thiamine pyrophosphate, it may be envisaged that this area of density belongs to the proton from the amino-pyrimidine ring. It has been suggested that it is possible to observed protons at such high structural resolution. Further evidence would be required for this conclusion to be made. One possible method could be to take advantage of the size of the *E. coli* TK crystals, and to subject them to neutron diffraction studies, enabling the 'possible' proton to be viewed. Another observation of the high-resolution structural studies was the non-symmetrical nature of the electron densities for the substrates. Although, the positions and occupancy of the substrates appeared to be the

same, one active site always seemed to be more defined than the other. Although work has been conducted on this previously, it would be an interesting project to link substrate occupancy of the active sites via the proposed ‘proton wire’ in communication between the two active sites.

The 2,5 DKMO research is still at an early stage, as the structure determined in this thesis was only of the dimeric oxygenase subunits. Further work to determine the structure in complex with the dehydrogenase component is being carried out by Dr. Rob Gibson in Prof. Littlechild’s group.

Appendix I



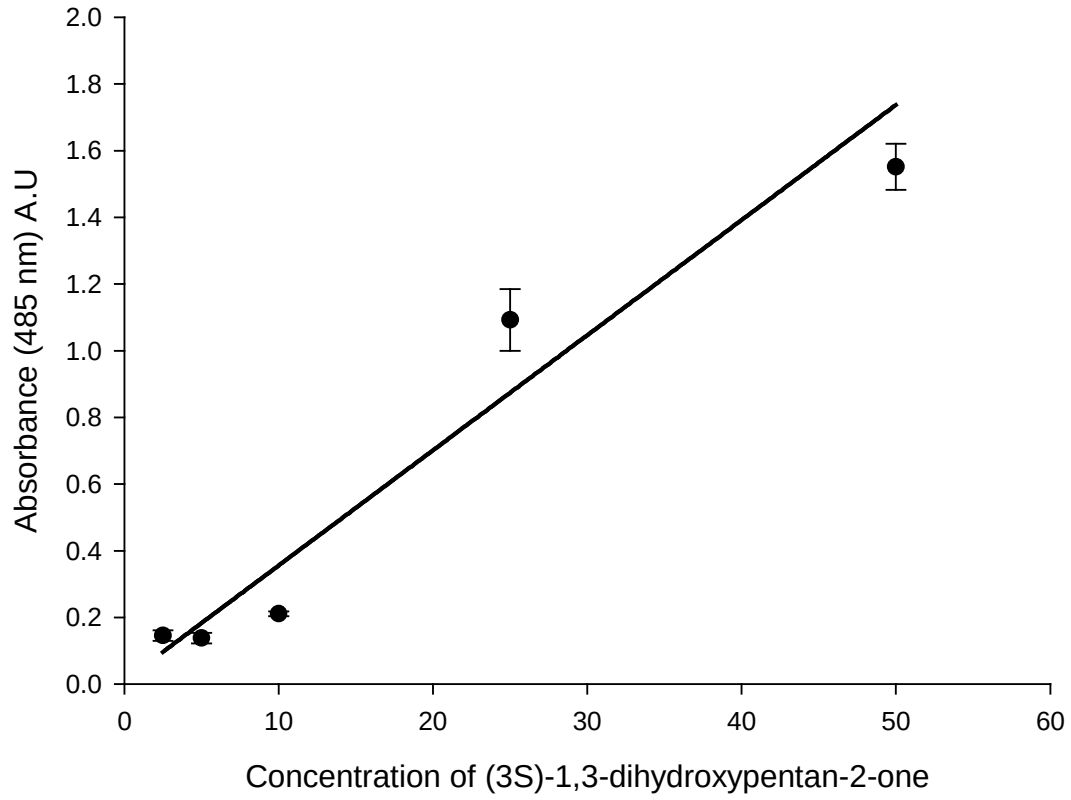
Kav = Ve – Vo / Vt – Vo
Where Ve is elution volume of peak
Vo is void volume of column with blue dextran = 44ml
Vt is the total bed colume of the column = 122ml

Equation of line:

Kav = 2.34409 – 0.38912 x log MW

2.34409 – Kav / 0.38912 = log MW

Appendix Ia - Superdex 200 gel filtration column calibration



Appendix Ib – Concentration - absorbance standard curve for colormetric assay section 2.2.1

Appendix II

Crystallisation buffers

Buffer C 50 mM PIPES buffer pH 6.4 containing 20 mM Thiamine pyrophosphate and 90 mM $CaCl_2$.

Buffer D 50 mM PIPES buffer pH 6.4 containing 10 mM Thiamine pyrophosphate and 45 mM $CaCl_2$.

Buffer E 50 mM PIPES buffer pH 6.4 containing 2 mM Thiamine pyrophosphate and 9 mM $CaCl_2$.

Buffer F 50 mM PIPES buffer pH 6.4 containing 10 mM Thiamine pyrophosphate, 45 mM $CaCl_2$ and 0.3 mM β-Fluoropyruvic Acid.

Buffer G 50 mM PIPES buffer pH 6.4 containing 2mM Thiamine pyrophosphate, 9 mM $CaCl_2$ and 0.3 mM β-Fluoropyruvic Acid.

Vapour diffusion crystallisation conditions

%$(NH_4)_2SO_4$	**Buffer + pH in Solvent Well**	**Sitting Drop Composition**	**Initial protein concentration + Buffer**
40	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
40	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
40	50 mM PIPES, pH 6.4	5 µl +1 µl	20 mg/ml + Buffer C
42	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
42	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
42	50 mM PIPES, pH	5 µl +1 µl	20 mg/ml + Buffer C

	6.4		
44	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
44	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
44	50 mM PIPES, pH 6.4	5 µl +1 µl	20 mg/ml + Buffer C
46	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
46	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
46	50 mM PIPES, pH 6.4	5 µl +1 µl	20 mg/ml + Buffer C
48	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
48	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
48	50 mM PIPES, pH 6.4	5 µl +1 µl	20 mg/ml + Buffer C
50	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
50	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
50	50 mM PIPES, pH 6.4	5 µl +1 µl	20 mg/ml + Buffer C

%$(NH_4)_2SO_4$	**Buffer + pH in Solvent Well**	**Sitting Drop Composition**	**Initial protein concentration + Buffer**
40	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
40	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D
40	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer E
44	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
44	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D
44	50 mM PIPES, pH 6.4	9 µl +1µl	20 mg/ml + Buffer E
48	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
48	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D
48	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer E
52	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
52	50 mM PIPES, pH	9 µl +1 µl	20 mg/ml + Buffer D

	6.4		
52	50 mM PIPES, pH 6.4	9 µl +1µl	20 mg/ml + Buffer E
56	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
56	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D
56	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer E
60	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer C
60	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D
60	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer E

%$(NH_4)_2SO_4$	Buffer + pH in Solvent Well	Sitting Drop Composition	Initial protein concentration + Buffer
46	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D
46	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer E
46	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer F
46	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer G
50	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D
50	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer E
50	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer F
50	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer G
54	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D
54	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer E
54	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer F
54	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer G
58	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D
58	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer E
58	50 mM PIPES, pH 6.4	9 µl +1µl	20 mg/ml + Buff er F
58	50 mM PIPES, pH 6.4	9 µl +1µl	20 mg/ml + Buffer G
62	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D
62	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer E
62	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer F
62	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer G
66	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer D
66	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer E
66	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer F
66	50 mM PIPES, pH 6.4	9 µl +1 µl	20 mg/ml + Buffer G

Vapour diffusion TK-FPA co-crystallisation conditions

Microbatch screening conditions

No.	Buffer	Precipitant	pH
1	0.1 M Citric acid	1.0 M NaCl	4.0
2	0.1 M Citric acid	1.0 M NaCl	5.0
3	0.1 M MES	1.0 M NaCl	6.0
4	0.1 M HEPES	1.0 M NaCl	7.0
5	0.1 M TRIS	1.0 M NaCl	8.0
6	0.1 M BICINE	1.0 M NaCl	9.0
7	0.1 M Citric acid	2.0 M NaCl	4.0
8	0.1 M Citric acid	2.0 M NaCl	5.0
9	0.1 M MES	2.0 M NaCl	6.0
10	0.1 M HEPES	2.0 M NaCl	7.0
11	0.1 M TRIS	2.0 M NaCl	8.0
12	0.1 M BICINE	2.0 M NaCl	9.0
13	0.1 M Citric acid	3.0 M NaCl	4.0
14	0.1 M Citric acid	3.0 M NaCl	5.0
15	0.1 M MES	3.0 M NaCl	6.0
16	0.1 M HEPES	3.0 M NaCl	7.0
17	0.1 M TRIS	3.0 M NaCl	8.0
18	0.1 M BICINE	3.0 M NaCl	9.0
19	0.1 M Citric acid	4.0 M NaCl	4.0
20	0.1 M Citric acid	4.0 M NaCl	5.0
21	0.1 M MES	4.0 M NaCl	6.0
22	0.1 M HEPES	4.0 M NaCl	7.0
23	0.1 M TRIS	4.0 M NaCl	8.0
24	0.1 M BICINE	4.0 M NaCl	9.0
25	0.1 M Citric acid	5 % (w/v) PEG 6000	4.0
26	0.1 M Citric acid	5 % (w/v) PEG 6000	5.0
27	0.1 M MES	5 % (w/v) PEG 6000	6.0
28	0.1 M HEPES	5 % (w/v) PEG 6000	7.0
29	0.1 M TRIS	5 % (w/v) PEG 6000	8.0
30	0.1 M BICINE	5 % (w/v) PEG 6000	9.0
31	0.1 M Citric acid	10 % (w/v) PEG 6000	4.0
32	0.1 M Citric acid	10 % (w/v) PEG 6000	5.0
33	0.1 M MES	10 % (w/v) PEG 6000	6.0
34	0.1 M HEPES	10 % (w/v) PEG 6000	7.0
35	0.1 M TRIS	10 % (w/v) PEG 6000	8.0
36	0.1 M BICINE	10 % (w/v) PEG 6000	9.0
37	0.1 M Citric acid	20 % (w/v) PEG 6000	4.0
38	0.1 M Citric acid	20 % (w/v) PEG 6000	5.0
39	0.1 M MES	20 % (w/v) PEG 6000	6.0
40	0.1 M HEPES	20 % (w/v) PEG 6000	7.0
41	0.1 M TRIS	20 % (w/v) PEG 6000	8.0
42	0.1 M BICINE	20 % (w/v) PEG 6000	9.0
43	0.1 M Citric acid	30 % (w/v) PEG 6000	4.0
44	0.1 M Citric acid	30 % (w/v) PEG 6000	5.0
45	0.1 M MES	30 % (w/v) PEG 6000	6.0
46	0.1 M HEPES	30 % (w/v) PEG 6000	7.0
47	0.1 M TRIS	30 % (w/v) PEG 6000	8.0
48	0.1 M BICINE	30 % (w/v) PEG 6000	9.0
49	0.1 M Citric acid	0.8 M NH4SO4	4.0
50	0.1 M Citric acid	0.8 M NH4SO4	5.0
51	0.1 M MES	0.8 M NH4SO4	6.0
52	0.1 M HEPES	0.8 M NH4SO4	7.0
53	0.1 M TRIS	0.8 M NH4SO4	8.0
54	0.1 M BICINE	0.8 M NH4SO4	9.0
55	0.1 M Citric acid	1.6 M NH4SO4	4.0

56	0.1 M Citric acid	1.6 M NH4SO4	5.0
57	0.1 M MES	1.6 M NH4SO4	6.0
58	0.1 M HEPES	1.6 M NH4SO4	7.0
59	0.1 M TRIS	1.6 M NH4SO4	8.0
60	0.1 M BICINE	1.6 M NH4SO4	9.0
61	0.1 M Citric acid	2.4 M NH4SO4	4.0
62	0.1 M Citric acid	2.4 M NH4SO4	5.0
63	0.1 M MES	2.4 M NH4SO4	6.0
64	0.1 M HEPES	2.4 M NH4SO4	7.0
65	0.1 M TRIS	2.4 M NH4SO4	8.0
66	0.1 M BICINE	2.4 M NH4SO4	9.0
67	0.1 M Citric acid	3.2 M NH4SO4	4.0
68	0.1 M Citric acid	3.2 M NH4SO4	5.0
69	0.1 M MES	3.2 M NH4SO4	6.0
70	0.1 M HEPES	3.2 M NH4SO4	7.0
71	0.1 M TRIS	3.2 M NH4SO4	8.0
72	0.1 M BICINE	3.2 M NH4SO4	9.0
73	0.1 M Citric acid	10 % (v/v) MPD	4.0
74	0.1 M Na acetate	10 % (v/v) MPD	5.0
75	0.1 M MES	10 % (v/v) MPD	6.0
76	0.1 M HEPES	10 % (v/v) MPD	7.0
77	0.1 M TRIS	10 % (v/v) MPD	8.0
78	0.1 M BICINE	10 % (v/v) MPD	9.0
79	0.1 M Citric acid	20 % (v/v) MPD	4.0
80	0.1 M Na acetate	20 % (v/v) MPD	5.0
81	0.1 M MES	20 % (v/v) MPD	6.0
82	0.1 M HEPES	20 % (v/v) MPD	7.0
83	0.1 M TRIS	20 % (v/v) MPD	8.0
84	0.1 M BICINE	20 % (v/v) MPD	9.0
85	0.1 M Citric acid	40 % (v/v) MPD	4.0
86	0.1 M Na acetate	40 % (v/v) MPD	5.0
87	0.1 M MES	40 % (v/v) MPD	6.0
88	0.1 M HEPES	40 % (v/v) MPD	7.0
89	0.1 M TRIS	40 % (v/v) MPD	8.0
90	0.1 M BICINE	40 % (v/v) MPD	9.0
91	0.1 M Citric acid	65 % (v/v) MPD	4.0
92	0.1 M Na acetate	65 % (v/v) MPD	5.0
93	0.1 M MES	65 % (v/v) MPD	6.0
94	0.1 M HEPES	65 % (v/v) MPD	7.0
95	0.1 M TRIS	65 % (v/v) MPD	8.0
96	0.1 M BICINE	65 % (v/v) MPD	9

Nextal Biotech - The pH clear crystallisation screen conditions.

No.	Reagent composition
1.	0.02M Calcium chloride dehydrate, 0.1M Na Acetate trihydrate pH 4.6, 30% v/v 2-methyl-2,4 pentanediol
2.	0.2 M Ammonium acetate, 0.1M Na Acetate trihydrate pH 4.6, 30 % w/v PEG 4000
3.	0.2 M Ammonium sulfate , 0.1M Na acetate trihydrate pH 4.6, 25 % w/v PEG 4000
4.	0.1 M Na acetate trihydrate pH 4.6 , 2.0M Sodium formate
5.	0.1 M Na acetate trihydrate pH 4.6, 2.0M Ammonium sulfate
6.	0.1 M Na acetate trihydrate pH 4.6, 8 % w/v PEG 4000
7.	0.2 M Ammonium acetate, 0.1 M tri-sodium citrate dihydrate pH 5.6, 30 % w/v PEG 4000
8.	0.2 M Ammonium acetate , 0.1 M tri-sodium citrate dihydrate pH 5.6, 30 % v/v 2-methyl-2,4-pentanediol
9.	0.1 M tri-Sodium citrate dihydrate pH 5.6, 20 % w/v 2-propanol, 20 % w/v PEG 4000
10.	0.1 M Na Citrate pH 5.6, 1.0M Ammonium dihydrogen phosphate
11.	0.2 M Calcium chloride dihydrate, 0.1 M Na acetate trihydrate pH 4.6, 20 % v/v 2-propanol
12.	0.1 M Na Cacodylate pH 6.5, 1.4 M Na acetate trihydrate
13.	0.2 M tri-sodium citrate dihydrate, 0.1 M Na Cacodylate pH 6.5, 30 % v/v 2-propanol
14.	0.2 M Ammoniium sulfate, 0.1 M Na Cacodylate pH 6.5, 30 % w/v PEG 8000
15.	0.2 M Magnesium acetate tetrahydrate, 0.1 M Na Cacodylate pH 6.5, 20 % PEG 8000
16.	0.2 M Magnesium acetate tetrahydrate , 0.1 M Na Cacodylate pH 6.5, 30 % v/v 2-methyl-2,4-pentanediol
17.	0.1 M Imidazole pH 6.5, 1.0M Sodium acetate trihydrate
18.	0.2 M Sodium acetate trihydrate, 0.1 M Na Cacodylate pH 6.5, 30 % w/v PEG 8000
19.	0.2 M Zinc acetate dihydrate, 0.1 M Na Cacodylate pH 6.5, 18 % w/v PEG 8000
20.	0.2 M Calcium acetate hydrate, 0.1 M Na Cacodylate pH 6.5, 18 % w/v PEG 8000
21.	0.2 M tri-sodium citrate dihydrate, 0.1 M Na Hepes pH 7.5, 30 % v/v 2-methyl-2,4-pentanediol
22.	0.2 M Magnesium chloride hexahydrate, 0.1 M Na Hepes pH 7.5, 30 % v/v 2-propanol
23.	0.2 M Calcium chloride dehydrate, 0.1 M Na Hepes pH 7.5, 28 % v/v PEG 400
24.	0.2 M Magnesium chloride hexahydrate, 0.1M Na Hepes pH 7.5, 30 % v/v PEG 400
25.	0.2 M tri-sodium citrate dihydrate , 0.1M Na Hepes pH 7.5, 20 % v/v 2-propanol
26.	0.1 M Na Hepes pH 7.5, 0.8 M K, Na tartrate tetrahydrate
27.	0.1 M Na Hepes pH 7.5, 1.5 M Lithium sulfate monohydrate
28.	0.1 M Na Hepes pH 7.5, 0.8 M Na dihydrogen phosphate, 0.8M K dihydrogen phosphate monohyd.
29.	0.1 M Na Hepes pH 7.5 , 1.4 M tri-Sodium citrate dihydrate
30.	0.1 M Na Hepes pH 7.5 , 2 % v/v PEG 400, 2.0M Amm sulfate
31.	0.1 M Na Hepes pH 7.5, 10 % v/v 2-propanol, 20 % w/v PEG 4000
32.	0.1 M Tris HCl pH 8.5, 2.0M Ammoniium sulfate
33.	0.2 M Magnesium chloride hexahydrate, 0.1M Tris HCl pH 8.5, 30 % w/v PEG 4000
34.	0.2 M tri-sodum citrate dihydrate, 0.1 M Tris HCl pH 8.5, 30 % v/v PEG 400
35.	0.2 M Lithium sulfate monohydrate, 0.1 M Tris HCl pH 8.5, 30 % w/v PEG 4000
36.	0.2 M Ammonium acetate, 0.1M Tris HCl pH 8.5, 30 % v/v 2-propanol
37.	0.2 M Sodium acetate trihydrate, 0.1M Tris HCl pH 8.5, 30 % w/v PEG 4000
38.	0.1 M Tris HCl pH 8.5, 8 % w/v PEG 8000
39.	0.1 M Tris HCl pH 8.5, 2.0M Ammonium dihydrogen phosphate
40.	0.4 M K, Na Tartrate tetrahydrate
41.	0.4 M Ammonium dihydrogen phosphate
42.	0.2 M Ammonium sulfate, 30 % w/v PEG 8000
43.	0.2 M Ammonium sulfate , 30 % w/v PEG 4000
44.	2.0 M Ammonium sulfate
45.	4.0 M Sodium formate
46.	0.05 M Potassium dihydrogen phosphate, 20 % w/v PEG 8000
47.	30 % w/v PEG 1500
48.	0.2 M Magnesium formate
49.	1.0 M Lithium sulfate monohydrate, 2 % w/v PEG 8000
50.	0.5 M Lithium sulfate monohydrate, 15 % w/v PEG 8000

Molecular Dimensions Ltd Structure Screen 1 crystallisation screen conditions

No.	Reagent composition
1.	0.1 M Sodium chloride, 0.1M Bicine pH 9.0, 30 % w/v PEG monomethylether 550
2.	0.1 M Bicine pH 9.0, 2.0M Magnesium chloride hexahydrate
3.	2 % w/v Dioxane, 0.1M Bicine pH 9.0, 10 % w/v PEG 20,000
4.	0.2 M Magnesium chloride hexahydrate, 0.1 M Tris pH 8.5, 3.4M 1,6 Hexanediol
5.	0.1 M Tris pH 8.5, 25 % v/v tert-Butanol
6.	0.01 M Nickel chloride hexahydrate, 0.1 M Tris pH 8.5, 1.0M Lithium sulfate
7.	1.5M Ammonium sulfate, 0.1 M Tris pH 8.5, 12 % v/v Glycerol
8.	0.2M Ammonium phosphate monobasic, 0.1 M Tris pH 8.5, 50 % v/v MPD
9.	0.1M Tris pH 8.5, 20 % v/v Ethanol
10.	0.01M Nickel chloride hexahydrate, 0.1M Tris pH 8.5, 20 % w/v PEG monomethylether 2000
11.	0.5M Ammonium sulfate, 0.1M Hepes pH 7.5, 30 % v/v MPD
12.	0.1 M Hepes pH 7.5, 10 % w/v PEG 6000, 5 % v/v MPD
13.	0.1 M Hepes pH 7.5, 20 % v/v Jeffamine M-600
14.	0.1 M Sodium chloride, 0.1 M Hepes pH 7.5, 1.6 M Ammonium sulfate
15.	0.1 M Hepes pH 7.5, 2.0 M Ammonium formate
16.	0.05 M Cadmium sulfate octahydrate, 0.1 M Hepes pH 7.5, 1.0M Sodium acetate
17.	0.1 M Hepes pH 7.5, 70 % v/v MPD
18.	0.1 M Hepes pH 7.5, 4.3M Sodium chloride
19.	0.1 M Hepes pH 7.5, 10% w/v PEG 8000, 8 % v/v Ethylene glycol
20.	0.1 M Mes pH 6.5, 1.6 M Magnesium sulfate heptahydrate
21.	0.1 M Na phosphate monobasic, 0.1M Mes pH 6.5, 2.0 M Sodium Chloride, 0.1 M K phosphate monobasic
22.	0.1 M Mes pH 6.5, 12 % w/v PEG 20,000
23.	1.6 M Ammonium sulfate, 0.1 M Mes pH 6.5, 10 % v/v Dioxane
24.	0.05 M Cesium chloride, 0.1 M Mes pH 6.5. 30 % v/v Jeffamine M-600
25.	0.01 M Cobalt chloride hexahydrate, 0.1 M Mes pH 6.5, 1.8M Ammonium sulfate
26.	0.2 M Ammonium sulfate, 0.1 M Mes pH 6.5, 30% w/v PEG monomethylether 5000
27.	0.01 M Zinc sulfate heptahydrate, 0.1 M Mes pH 6.5, 25% v/v PEG monomethylether 550
28.	0.1 M Hepes pH 7.5, 20 % w/v PEG 10,000
29.	0.2 M K/Na Tartrate, 0.1M Sodium citrate pH 5.6, 2.0M Ammonium sulfate
30.	0.5 M Ammonium sulfate, 0.1M Sodium citrate pH 5.6, 1.0M Lithium sulfate
31.	0.5 M Sodium chloride, 0.1M Sodium citrate pH 5.6, 4 % v/v polyethyleneimine
32.	0.1 M Sodium citrate pH 5.6, 35 % v/v tert-butanol
33.	0.01 M Ferric chloride hexahydrate, 0.1 M Sodium citrate pH 5.6, 10 % v/v Jeffamine M-600
34.	0.01 M Manganese chloride tetrahydrate, 0.1 M Sodium citrate pH 5.6, 2.5 M 1,6 Hexanediol
35.	0.1 M Sodium acetate pH 4.6, 2.0 M Sodium chloride
36.	0.2 M Sodium Chloride, 0.1 M Sodium acetate pH 4.6, 30 % v/v MPD
37.	0.01 M Cobalt Chloride hexahydrate, 0.1 M Sodium acetate pH 4.6, 1.0M 1,6 Hexanediol
38.	0.1 M Cadmium chloride, 0.1 M Sodium acetate pH 4.6, 30 % v/v PEG 400
39.	0.2 M Ammonium sulfate, 0.1M Sodium acetate pH 4.6, 30 % w/v PEG monomethylether 2000
40.	2.0 M Sodium Chloride, 10 % w/v PEG 6000
41.	0.01 M Cetyl trimethyl ammoniumbromide, 0.5 M Sodium chloride, 0.1 M Magnesium chloride hexahydrate
42.	25 % v/v Ethylene glycol
43.	35 % v/v Dioxane
44.	2.0 M Ammonium Sulfate, 5 % v/v Isopropanol
45.	1.0 M Imidazole pH 7.0
46.	10 % w/v PEG 1000, 10 % w/v PEG 8000
47.	1.5 M Sodium Chloride, 10 % v/v Ethanol
48.	1.6 M Sodium citrate pH 6.5
49.	15 % w/v Polyvinylpyrolidone
50.	2.0 M Urea

Molecular Dimensions Ltd Structure Screen 2 crystallisation screen conditions

No.	Reagent composition
1.	Ca-chloride 0.02 M, Na-acetate (pH 4.6) 0.1 M, 2-Methyl-2,4-pentanediol 30 %
2.	K-,Na-tartrate 0.4 M
3.	NH4-dihydrogenphosphate 0.4 M
4.	TRIS-HCl (pH 8.5) 0.1M, NH4-sulfate 2.0 M
5.	Na-citrate 0.2M, HEPES Na-salt (pH 7.5) 0.1 M, 2-Methyl-2,4-pentanediol 30%
6.	Mg-chloride 0.2 M, TRIS-HCl (pH 8.5) 0.1 M, PEG 4000 30 %
7.	Na-cacodylate (pH 6.5) 0.1 M, Na-acetate 1.4 M
8.	Na-citrate 0.2 M, Na-cacodylate (pH 6.5) 0.1 M, 2-Propanol 30 %
9.	NH4-acetate 0.2 M, Na-citrate (pH 5.6) 0.1 M, PEG 4000 30 %
10.	NH4-acetate 0.2 M, Na-acetate (pH 4.6) 0.1 M, PEG 4000 30 %
11.	Na-citrate (pH 5.6) 0.1 M, NH4-dihyrogenphosphate 1.0 M
12.	Mg-chloride 0.2 M, HEPES Na-salt (pH 7.5) 0.1M, 2-Propanol 30 %
13.	Na-citrate 0.2 M, TRIS-HCl (pH 8.5) 0.1 M, PEG 400 30 %
14.	Ca-chloride 0.2 M, HEPES Na-salt (pH 7.5) 0.1 M, PEG 400 28 %
15.	NH4-sulfate 0.2 M, Na-cacodylate (pH 6.5) 0.1M, PEG 8000 30 %
16.	HEPES Na-salt (pH 7.5) 0.1 M, Li-sulfate 1.5 M
17.	Li-sulfate 0.2 M, TRIS-HCl (pH 8.5) 0.1 M, PEG 4000 30 %
18.	Mg-acetate 0.2 M, Na-cacodylate pH (6.5) 0.1M, PEG 8000 20 %
19.	NH4-acetate 0.2 M, TRIS-HCl (pH 8.5) 0.1 M, 2-Propanol 30 %
20.	NH4-sulfate 0.2 M, Na-acetate (pH 4.6) 0.1 M, PEG 4000 25 %
21.	Mg-acetate 0.2 M, Na-cacodylate (pH 6.5) 0.1 M, 2-Methyl-2,4-pentanediol 30%
22.	Na-acetate 0.2 M, TRIS-HCl (pH 8.5) 0.1 M, PEG 4000 30 %
23.	Mg-chloride 0.2 M, HEPES Na-salt (pH 7.5) 0.1 M, PEG 400 30 %
24.	Ca-chloride 0.2 M, Na-acetate (pH 4.6) 0.1 M, 2-Propanol 20%
25.	Imidazole (pH 6.5) 0.1 M, Na-acetate 1 M
26.	NH4-acetate 0.2 M, Na-citrate (pH 5.6) 0.1 M, 2-Methyl-2,4-pentanediol 30 %
27.	Na-citrate 0.2 M, HEPES Na-salt (pH 7.5) 0.1 M, 2-Propanol 20 %
28.	Na-acetate 0.2 M, Na-cacodylate pH (6.5) 0.1 M, PEG 8000 30 %
29.	HEPES Na-salt (pH 7.5) 0.1 M, K-,Na-tartrate 0.8 M
30.	NH4-sulfate 0.2 M, PEG 8000 30 %
31.	NH4-sulfate 0.2 M, PEG 4000 30 %
32.	NH4-sulfate 2 M
33.	Na-formiate 4 M
34.	Na-acetate (pH 4.6) 0.1 M,Na-formiate 2 M
35.	HEPES Na-salt (pH 7.5) 0.1 M, K-dihydrogenphosphate 0.8 M, Na-dihydrogenphosphate 0.8 M
36.	TRIS-HCl (pH 8.5) 0.1 M, PEG 8000 8 %
37.	Na-acetate (pH 4.6) 0.1 M, PEG 4000 8 %
38.	HEPES Na-salt (pH 7.5) 0.1 M, Na-citrate 1.4 M
39.	HEPES Na-salt (pH 7.5) 0.1 M, PEG 400 2 %, NH4-sulfate 2.0 M
40.	Na-citrate (pH 5.6) 0.1 M, 2-Propanol 20 %, PEG 4000 20 %
41.	HEPES Na-salt (pH 7.5) 0.1 M, 2-Propanol 10 %, PEG 4000 20 %
42.	K-dihydrogenphosphate 0.05 M, PEG 8000 20 %
43.	PEG 1500 30 %
44.	Mg-formiate 0.2 M
45.	Zn-acetate 0.2 M, Na-cacodylate (pH 6.5) 0.1 M, PEG 8000 18 %
46.	Ca-acetate 0.2 M, Na-cacodylate (pH 6.5) 0.1 M, PEG 8000 18 %
47.	Na-acetate (pH 4.6) 0.1 M, NH4-sulfate 2.0 M
48.	TRIS-HCl (pH 8.5) 0.1 M, NH4-dihydrogenphosphate 2.0 M
49.	Li-sulfate 1.0 M, PEG 8000 2 %
50.	Li-sulfate 0.5 M, PEG 8000 15 %

Sigma 82009 Crystallization Basic Kit crystallisation screen conditions

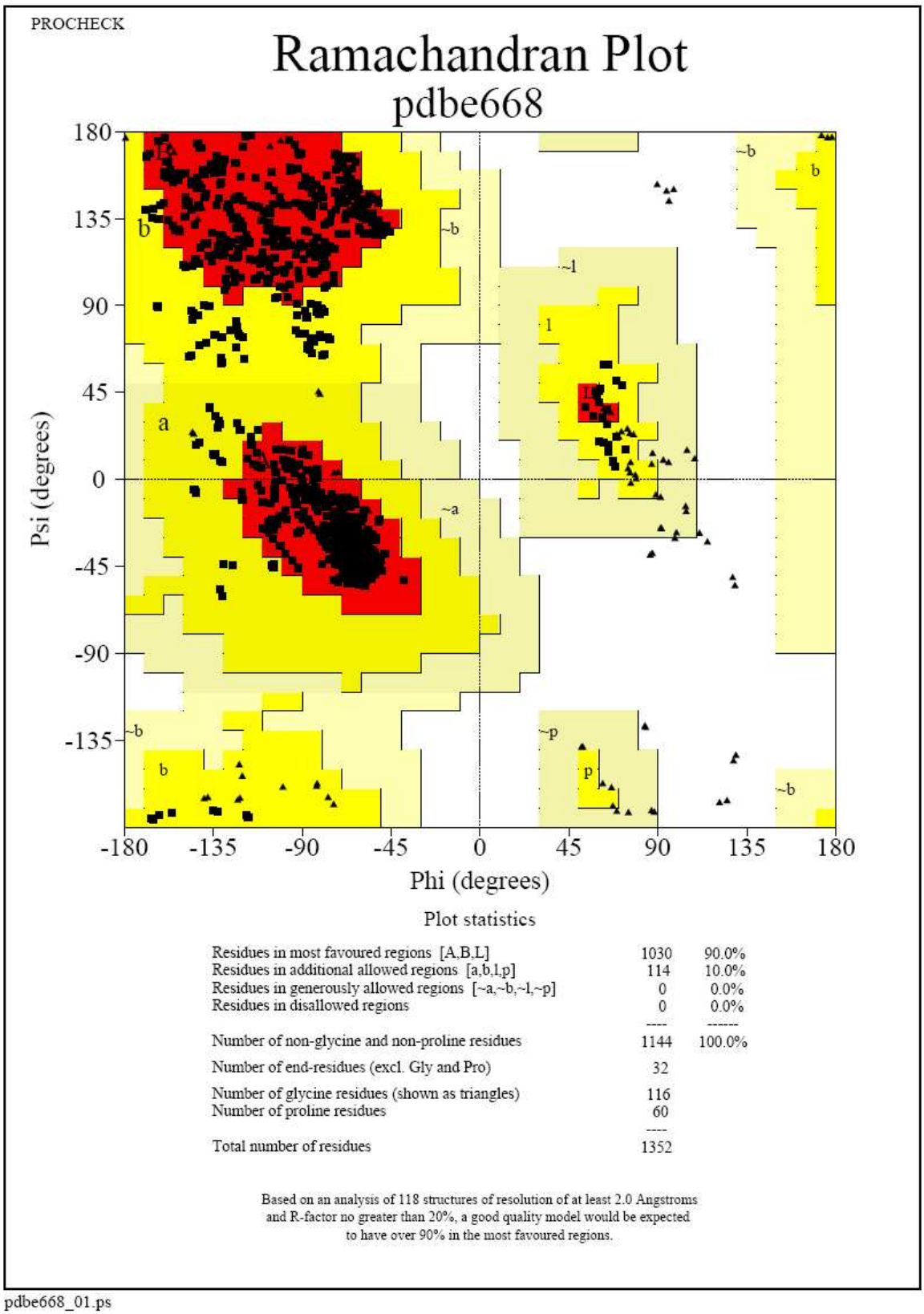
No.	Reagent composition
1.	Na-chloride 2 M, PEG 6000 10 %
2.	Na-chloride 0.5 M, Mg-chloride 0.01 M, CTAB 0.01 M
3.	Ethylene glycol 25 %
4.	Dioxane 35 %
5.	NH4 -sulfate 2 M, 2-Propanol 5 %
6.	Imidazole (pH 7.0) 1 M
7.	PEG 1000 10 %, PEG 8000 10 %
8.	Na-chloride 1.5 M, Ethanol 10 %
9.	Na-chloride 2 M, Na-acetate (pH 4.6) 0.1 M
10.	Na-chloride 0.2 M, Na-acetate (pH 4.6) 0.1 M, 2-Methyl-2,4-pentanediol 30 %
11.	Co-chloride 0.01 M, Na-acetate (pH 4.6) 0.1 M, 1,6-Hexanediol 1 M
12.	Cd-chloride 0.1 M, Na-acetate (pH 4.6) 0.1 M, PEG 400 30 %
13.	NH4 -sulfate 0.2 M, Na-acetate (pH 4.6) 0.1 M, PEG MME 2000 30 %
14.	K-,Na-tartrate 0.2 M, NH4 -sulfate 2 M, Na-citrate (pH 5.6) 0.1 M
15.	Li-sulfate 1 M, NH4 -sulfate 0.5 M, Na-citrate (pH 5.6) 0.1 M
16.	Na-chloride 0.5 M, Na-citrate (pH 5.6) 0.1 M, Polyethylenimin 2 %
17.	Na-citrate (pH 5.6) 0.1 M, tert.-Butanol 35 %
18.	Fe(III)-chloride 0.01 M, Na-citrate (pH 5.6) 0.1 M, Jeffamine M-600 10 %
19.	Na-citrate (pH 5.6) 0.1M, 1,6-Hexanediol 2.5 M
20.	Mg-sulfate 1.6 M, MES (pH 6.5) 0.1 M
21.	Na-chloride 2 M, Ka-,Na-dihydrogenphosphate each 0.1 M, MES (pH 6.5) 0.1 M
22.	MES (pH 6.5) 0.1 M, PEG 20'000 12 %
23.	NH4 -sulfate 1.6 M, MES (pH 6.5) 0.1 M, Dioxane 10 %
24.	Cs-chloride 0.05 M, MES (pH 6.5) 0.1 M, Jeffamine M-600 30 %
25.	Co-chloride 0.01 M, NH4 -sulfate 1.8M, MES (pH 6.5) 0.1 M
26.	NH4 -sulfate 0.2 M, MES (pH 6.5) 0.1 M, PEG MME 5000 30 %
27.	Zn-sulfate 0.01 M, MES (pH 6.5) 0.1 M, PEG MME 550 25 %
28.	Na-citrate (pH 6.5) 1.6 M
29.	NH4 -sulfate 0.5 M, HEPES Na-salt (pH 7.5) 0.1 M, 2-Methyl-2,4-pentanediol 30 %
30.	HEPES Na-salt (pH 7.5) 0.1 M, 2-Methyl-2,4-pentanediol 5 %, PEG 6000 10 %
31.	HEPES Na-salt (pH 7.5) 0.1 M, Jeffamine M-600 20 %
32.	NH4 -sulfate 1.6 M, Na-chloride 0.1 M, HEPES Na-salt (pH 7.5) 0.1 M
33.	NH4 -formate 2.0 M, HEPES Na-salt (pH 7.5) 0.1 M
34.	Na-acetate 1.0 M, Cd-sulfate 0.05 M, HEPES Na-salt (pH 7.5) 0.1 M
35.	HEPES Na-salt (pH 7.5) 0.1 M, 2-Methyl-2,4-pentanediol 70 %
36.	Na-chloride 4.3 M, HEPES Na-salt (pH 7.5) 0.1 M
37.	HEPES Na-salt (pH 7.5) 0.1 M, Ethylene glycol 8 %, PEG 8000 10 %
38.	HEPES Na-salt (pH 7.5) 0.1M, PEG 10'000 20 %
39.	Mg-chloride 0.2 M, TRIS-HCl (pH 8.5) 0.1 M, 1,6-Hexanediol 3.4 M
40.	TRIS-HCl (pH 8.5) 0.1M, tert.-Butanol 25 %
41.	Li-sulfate 1 M, Ni(II)-chloride 0.01 M, TRIS-HCl (pH 8.5) 0.1 M
42.	NH4 -sulfate 1.5 M, TRIS-HCl (pH 8.5) 0.1 M, Glycerol 12 %
43.	NH4 -phosphate 0.2 M, TRIS-HCl (pH 8.5) 0.1 M, 2-Methyl-2,4-pentanediol 50 %
44.	TRIS-HCl (pH 8.5) 0.1M, Ethanol 20 %
45.	Ni(II)-chloride 0.01 M, TRIS-HCl (pH 8.5) 0.1 M, PEG MME 2000 20 %
46.	Na-chloride 0.1 M, Bicine (pH 9.0) 0.1 M, PEG MME 550 20 %
47.	Mg-chloride 2 M, Bicine (pH 7.5) 0.1 M
48.	Bicine (pH 9.0) 0.1 M, Dioxane 2 %, PEG 20'000 10 %
49.	Mg-chloride 0.1 M, TRIS-HCl (pH 8.5) 0.1M, PEG 20000 15 %
50.	PEG 20000 20 %

Sigma 70437 Crystallization Extension Kit crystallisation screen conditions

Appendix III

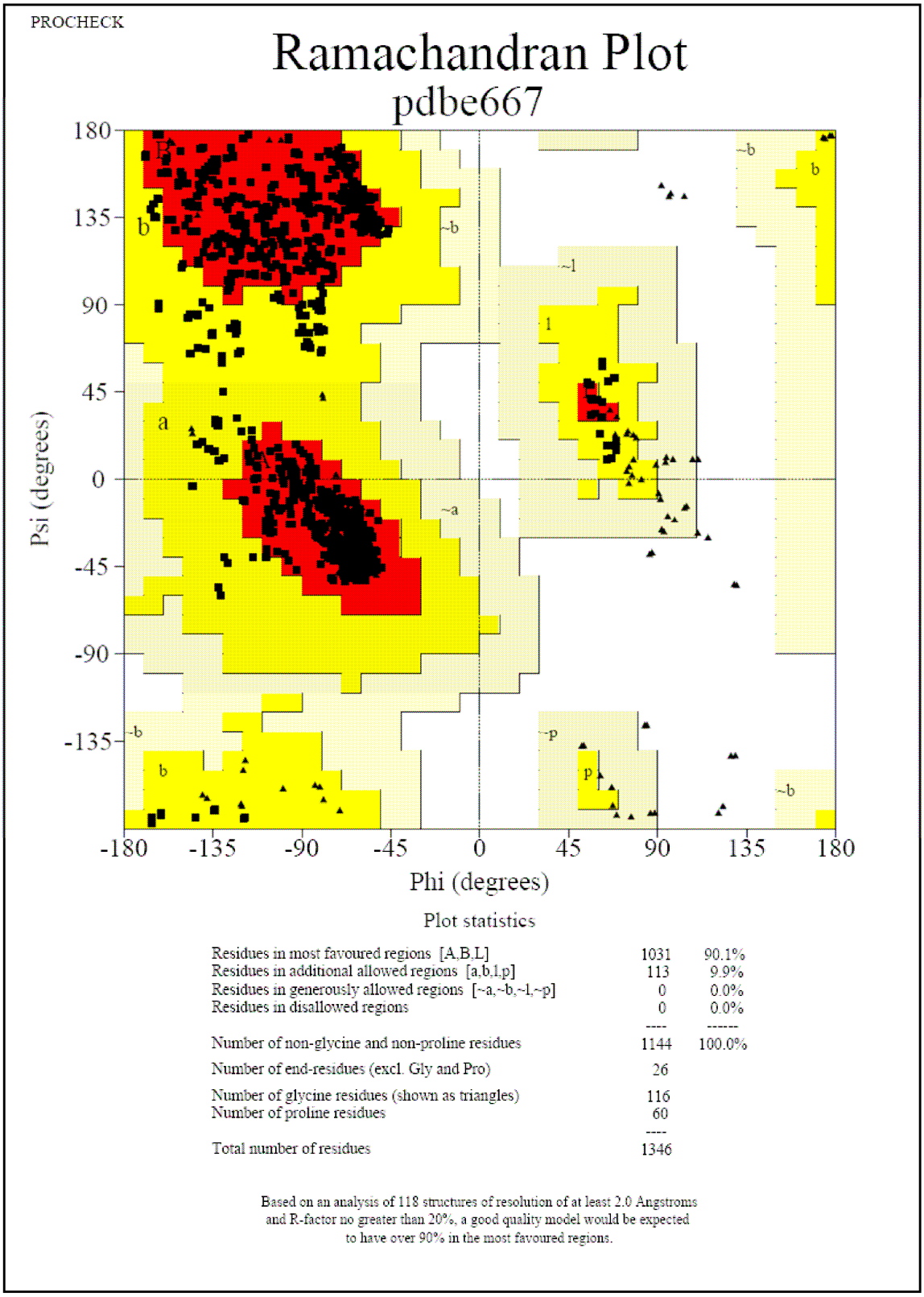
Ramachandran Diagram	Model
pdbe665	Final model of H26Y-HPA
pdbe666	Final model of TK-FPA
pdbe667	Final model of TK-HPA
pdbe668	Final model of TK

All models were produced using PROCHECK (Laskowski *et al.*, 1993).



pdbe668_01.ps

Appendix IIIa. Ramachandran plot of final model of TK



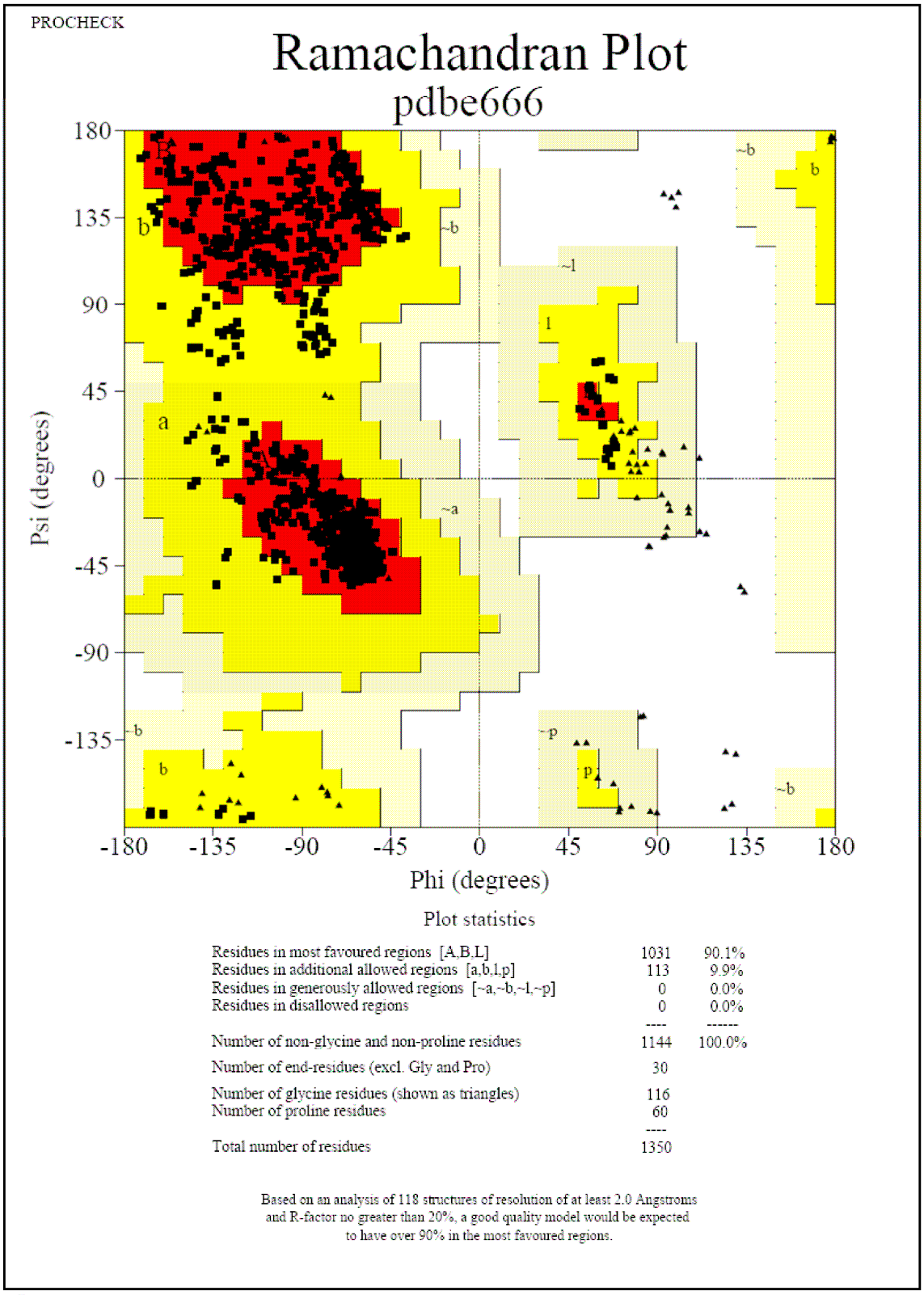
Plot statistics

Residues in most favoured regions [A,B,L]	1031	90.1%
Residues in additional allowed regions [a,b,l,p]	113	9.9%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	1144	100.0%
Number of end-residues (excl. Gly and Pro)	26	
Number of glycine residues (shown as triangles)	116	
Number of proline residues	60	
Total number of residues	1346	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

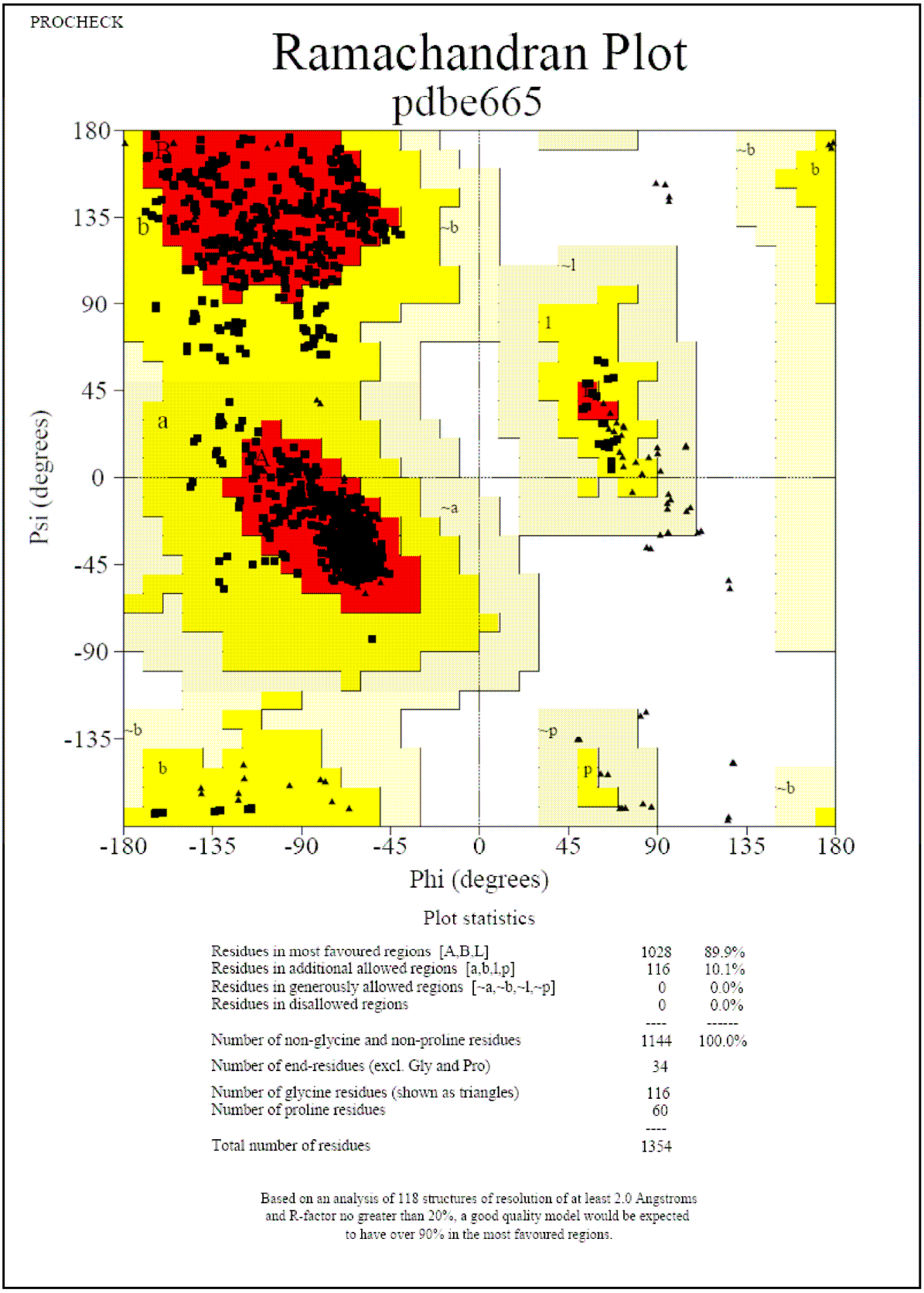
pdbe667_01.ps

Appendix IIIb. Ramachandran plot of final model of TK-HPA



pdbe666_01.ps

Appendix IIIc. Ramachandran plot of final model of TK-FPA

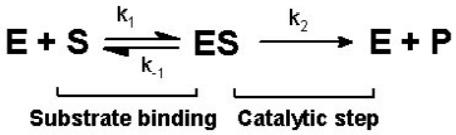


pdbe665_01.ps

Appendix IIId. Ramachandran plot of final model of H26Y-FPA

Appendix IV

Deriving the Michaelis-Menten Equation



From **Fig.1.1,** Where substrate S reversibly binds to the enzyme E in the Substrate binding step, with a proportion of the resulting complex ES breaking down and yielding the product P and the free enzyme E (Catalytic step). Where k_1, k_{-1}, and k_2 are constants.

Two assumptions must be made for this model,

1) As long as initial velocity is considered, the concentration of product can be neglected (compared to that of the substrate, thus [P] << [S]).

2) The concentration of substrate is in large excess over that of the enzyme ([E] << [S]).

These assumptions hold true for low enzyme concentrations, and allow $[S]_0 = [S]_{free} + [ES] + [P]$ to approximate to $[S]_0 = [S]$. With the enzyme approximating to $[E]_{total} = [E]_{free} + [ES]$.

The initial rate of the reaction maybe given by the first order equation,

$$\mathbf{v_0 = k_2[ES]} \qquad \textbf{(Eqn 1)}$$

where v_0 is the initial rate of the reaction, k_2 is catalytic constant and [ES] is concentration of enzyme/substrate complex.

For the initial phase of the reaction is said to be in steady state, and according to the model from Fig.1.X the rate at which [ES] is being formed and consumed is the same.

Rate of formation of [ES] = k_1[E][S]

Rate of consumption of [ES] = k_{-1}[ES] + k_2 [ES]

So in the steady state,

$$\mathbf{k_{-1}[ES] + k_2[ES] = k_1[E][S]} \qquad \textbf{(Eqn 2)}$$

Substituting Equation 2 into Equation 1.

$(k_{-1} + k_2)$ [ES] = k_1 [E][S], and $(k_{-1} + k_2)/k_1$ = [E][S]/[ES] **(Eqn 3)**

The kinetic constants are grouped as K_m.

$K_m = (k_{-1} + k_2)/k_1$ **(Eqn 4)**

[E] is expressed in terms of [ES] and $[E]_{total}$, to limit the number of unknowns.

$[E] = [E]_{total} - [ES]$ **(Eqn 5)**

Equations 5 and 4 are substituted into 3.

$K_m = ([E]_{total} - [ES])\,[S]/[ES]$ **(Eqn 6)**

Multiply both sides by [ES], collect [ES] on the left and factorise them. Then divide both sides by (K_m + [S]), making [ES] the subject.

$[ES] = [E]_{total}\,[S]/(K_m + [S])$ **(Eqn 7)**

Substitute equation 7 into 1

$v_0 = k_2[E]_{total}\,[S]/(K_m + [S])$ **(Eqn 8)**

V_{max} occurs when all of the enzyme is bound with substrate, make [ES] = $[E]_{total}$ giving equation 9

$V_{max} = k_2\,[E]_{total}$ **(Eqn 9)**

Thus substituting equation 9 into 8 giving the Michaelis-Menten Equation.

$v_0 = V_{max}\,[S]/(K_m + [S])$

References

Anfinsen, C. B. (1973). "Principles that Govern the Folding of Protein Chains." *Science* **181**: 223-230.

Asztalos, P., Parthier, C., Golbik, R., Kleinschmidt, M., Hubner, G., Weiss, M. S., Freidemann, R., Wille, G. and Tittmann, K. (2007). "Strain and Near Attack Conformers in Enzymic Thiamin Catalyis: X-ray Crystallographic Snapshots of Bacterial Transketolase in Covalent Complex with Donor Ketoses Xylulose 5-phosphate and Fructose 6-phosphate, and a Noncovalent Complex with Acceptor Aldose Ribose 5-phosphate." *Biochemistry* **46**: 12037-12052.

Baeyer, A. V. and Villiger, V. (1899). "Einwirkung des Caro'schen Reagens auf Ketone." *Berichte der deutschen chemischen Gesellschaft* **32**: 3625-3633.

Barkla, C. G. (1904). "Polarised Rontgen Radiation." *Proceedings of the Royal Society of London* **74**: 474-475.

Barkla, C. G. (1909). "Ionisation by Röntgen Rays." *Nature* **80**: 187.

Bassham, J., Benson, A. and Calvin, M. (1950). "The path of carbon in photosynthesis." *Journal of Biological Chemistry* **185**: 781-787.

Bergfors, T., Ed. (1999). Protein Crystallization Techniques, Strategies, and Tips. A Laboratory Manual.

Bernal, J. D. and Crowfoot, D. (1934). "X-Ray Photographs of Crystalline Pepsin." *Nature* **133**: 794-795.

Blewett, J. (1998). "Synchrotron Radiation - Early History." *Journal of Synchrotron Radiation* **5**: 135-139.

Bragg, W. L. (1912). "The Specular Reflection of X-rays." *Nature* **90**: 410.

Brent, R. L. (2004). "Environmental Causes of Human Congenital Malformations: The Pediatrician's Role in Dealing With These Complex Clinical Problems Caused by a Multiplicity of Environmental and Genetic Factors." *Pediatrics* **113**: 957-968.

Breslow, R. (1958). "On the Mechanism of Thiamine Action. IV.1 Evidence from Studies on Model Systems." *Journal of the American Chemical Society* **80**: 3719-3726.

Breslow, R. (1962). "The Mechanism of Thiamine Action: Predictions from Model Experiments." *Annals of the New York Academy of Sciences* **98**: 445-452.

Breslow, R. and McNelis, E. (1959). "Studies on Model Systems for Thiamine Action. Synthesis of Reactive Intermediates, and Evidence on the Function of the Pyrimidine Ring." *Journal of the American Chemical Society* **81**: 3080-3082.

Brocklebank, S., Woodley, J. M. and Lilly, M. D. (1999). "Immobilised transketolase for carbon-carbon bond synthesis: biocatalyst stability." *Journal of Molecular Catalysis B: Enzymatic* **7**: 223-231.

Brünger, A. T. (1997). Free R value: Cross-validation in crystallography. Methods in Enzymology, Academic Press. **Volume 277**, 366-396.

Burgi, H. B., Dunitz, J. D. and Shefter, E. (1973). "Geometrical Reaction Coordinates. II. Nucleophilic addition to a carbonyl group." *Journal of the American Chemical Society* **95**: 5065-5067.

CCP4 (1994). "The CCP4 suite: programs for protein crystallography." *Acta Crystallographica D* **50**: 760-763.

Charmantray, F., Hélaine, V., Legeret, B. and Hecquet, L. (2008). "Preparative scale enzymatic synthesis of d-sedoheptulose-7-phosphate from [beta]-hydroxypyruvate and d-ribose-5-phosphate." *Journal of Molecular Catalysis B: Enzymatic* **In Press, Corrected Proof**.

Charmantray, F., Hélaine, V., Legeret, B. and Hecquet, L. (2009). "Preparative scale enzymatic synthesis of d-sedoheptulose-7-phosphate from [beta]-hydroxypyruvate and d-ribose-5-phosphate." *Journal of Molecular Catalysis B: Enzymatic* **57**: 6-9.

Cleland, W. W. (1973). "Derivation of Rate Equations for Multisite Ping-Pong Mechanisms with Ping-Pong Reactions at One or More Sites." *Journal of Biological Chemistry* **248**: 8353-8355.

Collaborative (1994). "The CCP4 suite: programs for protein crystallography." *Acta Crystallographica Section D* **50**: 760-763.

Comin-Anduix, B., Boros, L. G., Marin, S., Boren, J., Callol-Massot, C., Centelles, J. J., Torres, J. L., Agell, N., Bassilian, S. and Cascante, M. (2002). "Fermented Wheat Germ Extract Inhibits Glycolysis/Pentose Cycle Enzymes and Induces Apoptosis through Poly(ADP-ribose) Polymerase Activation in Jurkat T-cell Leukemia Tumor Cells." *Journal of Biological Chemistry* **277**: 46408-46414.

Compton, A. (1923). "A Quantum Theory of the Scattering of X-Rays by Light Elements." *The Physical Review* **21**: 483-502.

Copeland, R. A. (2000). *Kinetics of Single-Substrate Enzyme Reactions*. Weinheim, Germany, Wiley-VCH-Verlag GmbH & Co. KGaA.

Criegee, R. (1948). "Die Umlagerung der Dekalin-peroxydester als Folge von kationischem Sauerstoff." *Liebigs Annalen der Chemie* **560**: 127-135.

Crowther, R. A. (1972). The Molecular Replacement Method. M. G. Rossmann. New York, Gordon & Breach, 173-178.

Crowther, R. A. and Blow, D. M. (1967). "A method of positioning a known molecule in an unknown crystal structure." *Acta Crystallographica* **23**: 544-548.

D'Arcy, A., Mac Sweeney, A., Stihle, M. and Haber, A. (2003). "The advantages of using a modified microbatch method for rapid screening of protein crystallization conditions." *Acta Crystallographica D* **59**: 396-399.

D'Arcy, A., Sweeney, A. M. and Haber, A. (2004). "Practical aspects of using the microbatch method in screening conditions for protein crystallization." *Methods* **34**: 323-328.

Damtoft, S., Franzyk, H. and Jensen, S. R. (1995). "Biosynthesis of iridoids in Syringa and Fraxinus: Secoiridoid precursors." *Phytochemistry* **40**: 773-784.

Datta, A. G. and Racker, E. (1961). "Mechanism of Action of Transketolase. I. Properties of the Crystalline Yeast Enzyme." *Journal of Biological Chemistry* **236**: 617-623.

de la Haba, G., Leder, I. G. and Racker, E. (1955). "Crystalline Transketolase from Bakers Yeast: Isolation and Properties." *Journal of Biological Chemistry* **214**: 409-426.

de Vries, J. G. (2001). "The Heck reaction in the production of fine chemicals." *Canadian Journal of Chemistry* **79**: 1086-1092.

DeLano, W. L. (2006). "PYMOL." *Delano Scientific LLC*.

Diels, O. and Alder, K. (1928). "Synthesen in der hydroaromatischen Reihe." *Liebigs Annalen der Chemie* **460**: 98-122.

Douglas-Instruments-Ltd. (2008). "http://www.douglas.co.uk/f_ftp1/scr_oryx.pdf."

Drenth, J. (1999). *Principles of protein X-ray crystallography*. Heidelberg, Springer-Verlag.

Du, M. X., Sim, J., Fang, L., Yin, Z., Koh, S., Stratton, J., Pons, J., Wang, J. J.-X. and Carte, B. (2004). "Identification of Novel Small-Molecule Inhibitors for Human Transketolase by High-Throughput Screening with Fluorescent Intensity (FLINT) Assay." *Journal of Biomolecular Screening* **9**: 427-433.

Elder, F. R., Gurewitsch, A. M., Langmuir, R. V. and Pollock, H. C. (1947). "Radiation from Electrons in a Synchrotron." *Physical Review* **71**: 829.

Emsley, P. and Cowtan, K. (2004). "Coot: model-building tools for molecular graphics." *Acta Crystallographica D* **60**: 2126-2132.

Esakova, O. A., Meshalkina, L. E., Golbik, R., Hubner, G. and Kochetov, G. A. (2004). "Donor substrate regulation of transketolase." *European Journal of Biochemistry* **271**: 4189-4194.

Esakova, O. A., Meshalkina, L. E. and Kochetov, G. A. (2005). "Effects of transketolase cofactors on its conformation and stability." *Life Sciences* **78**: 8-13.

Ewald, P. P. (1979). "A review of my papers on crystal optics 1912 to 1968." *Acta Crystallographica Section A* **35**: 1-9.

Fessner, W. D. (1998). "Enzyme mediated C-C bond formation." *Current Opinion in Chemical Biology* **2**: 85-97.

Fiedler, E., Thorell, S., Sandalova, T., Golbik, R., Konig, S. and Schneider, G. (2002). "From the Cover: Snapshot of a key intermediate in enzymatic thiamin catalysis: Crystal structure of the alpha -carbanion of (alpha ,beta -dihydroxyethyl)-thiamin diphosphate in the active site of transketolase from Saccharomyces cerevisiae." *Proceedings of the National Academy of Sciences* **99**: 591-595.

Fox, K. M. and Karplus, P. A. (1993). "Crystallization of Old Yellow Enzyme Illustrates an Effective Strategy for Increasing Protein Crystal Size." *Journal of Molecular Biology* **234**: 502-507.

Fraaije, M. W., Kamerbeek, N. M., van Berkel, W. J. H. and Janssen, D. B. (2002). "Identification of a Baeyer-Villiger monooxygenase sequence motif." *FEBS Letters* **518**: 43-47.

Frank, R., Leeper, F. and Luisi, B. (2007). "Structure, mechanism and catalytic duality of thiamine-dependent enzymes." *Cellular and Molecular Life Sciences (CMLS)* **64**: 892-905.

French, C. and Ward, J. M. (1995). "Improved production and stability of E. coli recombinants expressing transketolase for large scale biotransformation." *Biotechnology Letters* **17**: 247-252.

Friedrich, W., Knipping, P. and von Laue, M. (1912). "Interferenz-Erscheinungen bei Röntgenstrahlen." *Sitzungsberichte der Mathematisch-Physikalischen Classe der Königlich-Bayerischen Akademie der Wissenschaften zu München*: 303–322.

Ganem, B. (1978). "From glucose to aromatics: recent developments in natural products of the shikimic acid pathway." *Tetrahedron* **34**: 3353-3383.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. and Bairoch, A. (2008). "ProtParam Protein Identification and Analysis Tools on the ExPASy Server."

Gerhardt, S., Echt, S., Busch, M., Freigang, J., Auerbach, G., Bader, G., Martin, W., Bacher, A., Huber, R. and Fischer, M. (2003). "Structure and Properties of an Engineered Transketolase from Maize." *Plant Physiology* **132**: 1941-1949.

Golbik, R., Neef, H., Hubner, G., Konig, S., Seliger, B., Meshalkina, L., Kochetov, G. A. and Schellenberger, A. (1991). "Function of the aminopyrimidine part in thiamine pyrophosphate enzymes." *Bioorganic Chemistry* **19**: 10-17.

Grignard, V. (1900). "Sur quelques nouvelles combinaisons organométaliques du magnésium et leur applicatione à des synthèses d'alcools et d'hydrocabures." *Comptes Rendus de l'Académie des Sciences* **130**: 1322–1325.

Guy, J. E., Isupov, M. N. and Littlechild, J. A. (2003). "The Structure of an Alcohol Dehydrogenase from the Hyperthermophilic Archaeon Aeropyrum pernix." *Journal of Molecular Biology* **331**: 1041-1051.

Gyamerah, M. and Willetts, A. J. (1997). "Kinetics of overexpressed transketolase from Escherichia coli JM 107/pQR 700." *Enzyme and Microbial Technology* **20**: 127-134.

Heck, R. F. and Nolley, J. J. P. (1972). "Palladium-catalyzed vinylic hydrogen substitution reactions with aryl, benzyl, and styryl halides." *Journal of Organic Chemistry* **37**: 2320–2322.

Hecquet, L., Bolte, J. and Demuynck, C. (1996). "Enzymatic synthesis of "natural-labeled" 6-deoxy-L-sorbose precursor of an important food flavor." *Tetrahedron* **52**: 8223-8232.

Heinrich, P. C., Steffen, H., Janser, P. and Wiss, O. (1972). "Studies on the Reconstitution of Apotransketolase with Thiamine Pyrophosphate and Analogs of the Coenzyme." *European Journal of Biochemistry* **30**: 533-541.

Hibbert, E. G., Senussi, T., Costelloe, S. J., Lei, W., Smith, M. E. B., Ward, J. M., Hailes, H. C. and Dalby, P. A. (2007). "Directed evolution of transketolase activity on non-phosphorylated substrates." *Journal of Biotechnology* **131**: 425-432.

Hobbs, G., Mitra, R., Chauhan, R., Woodley, J. and Lilly, M. (1996). "Enzyme-catalysed carbon-carbon bond formation: Large-scale production of Escherichia coli transketolase." *Journal of Biotechnology* **45**: 173-179.

Horecker, B. L., Gibbs, M., H., K. and Smyrniotis, P. Z. (1954). "The mechanism of pentose phosphate conversion to hexose monophosphate." *The Journal of Biological Chemistry* **207**: 393-403.

Horecker, B. L. and Smyrniotis, P. Z. (1952). "The enzymatic formation of sedoheptulose phosphate from pentose phosphate." *Journal of the American Chemical Society* **74**: 2123.

Hubner, G., Tittmann, K., Killenberg-Jabs, M., Schaffner, J., Spinka, M., Neef, H., Kern, D., Kern, G., Schneider, G., Wikner, C. and Ghisla, S. (1998). "Activation of thiamin diphosphate in enzymes." *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1385**: 221-228.

Isupov, M. N. and Lebedev, A. A. (2008). "NCS-constrained exhaustive search using oligomeric models." *Acta Crystallographica D* **64**: 90-98.

Iwanenko, D. and Pomeranchuk, I. (1944). "On the Maximal Energy Attainable in a Betatron." *Physical Review* **65**: 343.

Jones, J. B., Sih, C. J. and Perlman, D. (1976). *Applications of biochemical systems in organic chemistry. Part 2*. United States, John Wiley and Sons,New York.

Jones, K. H., Smith, R. T. and Trudgill, P. W. (1993). "Diketocamphane enantiomer-specific 'Baeyer-Villiger' monooxygenases from camphor-grown Pseudomonas putida ATCC 17453." *Journal of General Microbiology* **139**: 797-805.

Kamerbeek, N. A., Janssen, D. B., van Berkel, W. J. H. and Fraaije, M. W. (2003). "Baeyer-Villiger Monooxygenases, an Emerging Family of Flavin-Dependent Biocatalysts." *Advanced Synthesis & Catalysis* **345**: 667-678.

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H. and Phillips, D. C. (1958). "A three-dimensional model of the myoglobin molecule obtained by X-ray analysis." *Nature* **181**: 662-666.

Koch, K., van den Berg, R. J. F., Nieuwland, P. J., Wijtmans, R., Schoemaker, H. E., van Hest, J. C. M. and Rutjes, F. P. J. T. (2008). "Enzymatic enantioselective C-C-bond formation in microreactors." *Biotechnology and Bioengineering* **99**: 1028-1033.

Kochetov, G. A. (1982). "Determination of transketolase activity via ferricyanide reduction." *Methods Enzymology* **89**: 43-44.

Kochetov, G. A. (2001). "Functional Flexibility of the Transketolase Molecule." *Biochemistry (Moscow)* **66**: 1077-1085.

Kochetov, G. A., Meshalkina, L. E. and Usmanov, R. A. (1976). "The Number of Active Sites in a Molecule of Transketolase " *Biochemical and Biophysical Research Communications* **69**: 839-843.

Kovina, M. V. and Kochetov, G. A. (1998). "Cooperativity and flexibility of active sites in homodimeric transketolase." *FEBS Letters* **440**: 81-84.

Kovina, M. V., Selivanov, V. A., Kochevova, N. V. and Kochetov, G. A. (1997). "Kinetic mechanism of active site non-equivalence in transketolase." *FEBS Letters* **418**: 11-14.

Kremer, A. B., Egan, R. M. and Sable, H. Z. (1980). "The active site of transketolase. Two arginine residues are essential for activity." *Journal of Biological Chemistry* **255**: 2405-2410.

Krishna, S. H. (2002). "Developments and trends in enzyme catalysis in nonconventional media." *Biotechnology Advances* **20**: 239-267.

Kuimov, A. N., Kovina, M. V. and Kochetov, G. A. (1988). "Inhibition of transketolase by N-acetylimidazole." *Biochemistry International* **17**: 517-521.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993). "PROCHECK: a program to check the stereochemical quality of protein structures." *Journal of Applied Crystallography* **26**: 283-291.

Leeper, F. J., Hawksley, D., Mann, S., Perez Melero, C. and Wood, M. D. H. (2005). "Studies on thiamine diphosphate-dependant enzymes." *Biochemical Society Transactions* **33**: 772-775.

Lineweaver, H. and Burk, D. (1934). "The Determination of Enzyme Dissociation Constants." *Journal of the American Chemical Society* **56**: 658-666.

Littlechild, J., Turner, N., Hobbs, G., Lilly, M., Rawas, A. and Watson, H. (1995). "Crystallization and preliminary X-ray crystallographic data with Escherichia coli transketolase." *Acta Crystallographica D* **51**: 1074-1076.

Lye, G. J. and Woodley, J. M. (1999). "Application of in situ product-removal techniques to biocatalytic processes." *Trends in Biotechnology* **17**: 395-402.

Malito, E., Alfieri, A., Fraaije, M. W. and Mattevi, A. (2004). "Crystal structure of a Baeyer-Villiger monooxygenase." *Proceedings of the National Academy of Sciences of the United States of America* **101**: 13157-13162.

Manfred, K. (2006). "White Biotechnology: Ready to partner and invest in." *Biotechnology Journal* **1**: 787-794.

Matthews, B. W. (1968). "Solvent content of protein crystals." *Journal of Molecular Biology* **33**: 491-497.

McGhie, E. J. (1998). "PhD Thesis." *University of Exeter*.

McMahon, D. J. and Brown, R. J. (1982). "Evaluation of Formagraph for Comparing Rennet Solutions." *Journal of Dairy Science* **65**: 1639-1642.

McPherson, A. (1991). "A brief history of protein crystal growth." *Journal of Crystal Growth* **110**: 1-10.

McPherson, A. (2004). "Introduction to protein crystallization." *Methods* **34**: 254-265.

Meshalkina, L., Nilsson, U., Wikner, C., Kostikowa, T. and Schneider, G. (1997). "Examination of the thiamin diphosphate binding site in yeast transketolase by site-directed mutagenesis." *European Journal of Biochemistry* **244**: 646–652.

Michaelis, L. and Menten, M. L. (1913). "Kinetik der Invertinwirkung." *Biochemische Zeitschrift* **49**: 333–369.

Minor, T. C., Milch, J. R. and Reynolds, G. T. (1974). "Protein crystallography using the rotation method and an image-intensifier-aided detector." *Journal of Applied Crystallography* **7**: 323-330.

Mitra, R. K. and Woodley, J. M. (1996). "A useful assay for transketolase in asymmetric syntheses." *Biotechnology Techniques* **10**: 167-172.

Mitra, R. K., Woodley, J. M. and Lilly, M. D. (1998). "Escherichia coli transketolase-catalyzed carbon-carbon bond formation: biotransformation characterization for reactor evaluation and selection." *Enzyme and Microbial Technology* **22**: 64-70.

Morgan, B., Dodds, D. R., Zaks, A., Andrews, D. R. and Klesse, R. (1997). "Enzymatic Desymmetrization of Prochiral 2-Substituted-1,3-propanediols: A Practical Chemoenzymatic Synthesis of a Key Precursor of SCH51048, a Broad-

Spectrum Orally Active Antifungal Agent." *Journal of Organic Chemistry* **62**: 7736-7743.

Morris, K. G., Smith, M. E. B., Turner, N. J., Lilly, M. D., Mitra, R. K. and Woodley, J. M. (1996). "Transketolase from Escherichia coli: A practical procedure for using the biocatalyst for asymmetric carbon-carbon bond synthesis." *Tetrahedron: Asymmetry* **7**: 2185-2188.

Morrison, J. F. and Stone, S. R. (1988). "Mechanism of the reaction catalyzed by dihydrofolate reductase from Escherichia coli: pH and deuterium isotope effects with NADPH as the variable substrate." *Biochemistry* **27**: 5499-5506.

Muench, H. (1934). "Treatment of hides." *Patent number: 1967679*: Issue date: 24 Jul 1934.

Muller, Y. A., Lindqvist, Y., Furey, W., Schulz, G. E., Jordan, F. and Schneider, G. (1993). "A thiamin diphosphate binding fold revealed by comparison of the crystal structures of transketolase, pyruvate oxidase and pyruvate decarboxylase " *Structure* **1**: 95-103.

Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997). "Refinement of Macromolecular Structures by the Maximum-Likelihood Method." *Acta Crystallographica D* **53**: 240-255.

Navaza, G. (1994). "AMORE: an automated package for molecular replacement." *Acta Crystallographica A* **50**: 157-163.

Nikkola, M., Lindqvist, Y. and Schneider, G. (1994). "Refined Structure of Transketolase from Saccharomyces cerevisiae at 2.0 A Resolution." *Journal of Molecular Biology* **238**: 387-404.

Nilsson, U., Lindqvist, Y., Kluger, R. and Schneider, G. (1993). "Crystal structure of transketolase in complex with thiamine thiazolone diphosphate, an analogue of the reaction intermediate, at 2.3 A resolution." *FEBS Letters* **326**: 145-148.

Nilsson, U., Meshalkina, L., Lindqvist, Y. and Schneider, G. (1997). "Examination of Substrate Binding in Thiamine Diphosphate-dependant Transketolase by Protein Crystallography and Site-directed Mutagenisis." *Journal of Biological Chemistry* **272**: 1864-1869.

Nixon, P. F., Diefenbach, R. J. and Duggleby, R. G. (1992). "Inhibition of transketolase and pyruvate decarboxylase by omeprazole." *Biochemical Pharmacology* **44**: 177-179.

Otwinowski, Z. and Minor, W. (2007). "Processing of X-ray diffraction data collected in oscillation mode " *Methods in Enzymology* **276**: 307-326.

Parratt, L. G. (1959). *Review of Scientific Instruments* **30**: 297-299.

Patterson, A. L. (1934). "A Fourier series for the method of determination of the components of interatomic distances in crystals." *Physical Review* **46**: 372-376.

Patterson, A. L. (1934). "A Fourier Series Method for the Determination of the Components of Interatomic Distances in Crystals." *Physical Review* **46**: 372.

Pauling, L. and Corey, R. B. (1951a). "Atomic coordinates and structur factors of two helical configurations of poly peptide chains." *Proceedings of the National Academy of Sciences* **37**: 235-240.

Pauling, L. and Corey, R. B. (1951b). "The pleated sheet, a new layer configuration of polypeptide chains." *Proceedings of the National Academy of Sciences* **37**: 251-256.

Perutz, M. (1964). "The Hemoglobin Molecule." *Scientific American* **211**: 64 - 76.

Petsko, G. A. (1975). "Protein crystallography at sub-zero temperatures: cryo-protective mother liquors for protein crystals." *Journal of Molecular Biology* **96**: 381-392.

Phillips, J. C., Wlodawer, A., Yevitz, M. M. and Hodgson, K. O. (1976). "Applications of synchotron radiation to protein crystallography: primary results." *Proceedings of the National Academy of Sciences* **73**: 128-132.

Pollock, H. C. (1983). "The Discovery of Synchrotron Radiation." *American Journal of Physics* **51**: 278-280.

Pusey, M. L., Liu, Z.-J., Tempel, W., Praissman, J., Lin, D., Wang, B.-C., Gavira, J. A. and Ng, J. D. (2005). "Life in the fast lane for protein crystallization and X-ray crystallography." *Progress in Biophysics and Molecular Biology* **88**: 359-386.

Racker, E., De la Haba, G. and Leder, I. G. (1953). "Thiamine pyrophosphate, a coenzyme of transketolase." *Journal of the American Chemical Society* **74**: 1010.

Raïs, B., Comin, B., Puigjaner, J., Brandes, J. L., Creppy, E., Saboureau, D., Ennamany, R., Paul Lee, W.-N., Boros, L. G. and Cascante, M. (1999). "Oxythiamine and dehydroepiandrosterone induce a G1 phase cycle arrest in Ehrlich's tumor cells through inhibition of the pentose cycle." *FEBS Letters* **456**: 113-118.

Ramakrishnan, C. and Ramachandran, G. N. (1965). "Stereo-chemical criteria for polypeptide and protein chain conformation." *Biophysical Journal* **5**: 909-933.

Rasor, J. P. and Voss, E. (2001). "Enzyme-catalyzed processes in pharmaceutical industry." *Applied Catalysis A: General* **221**: 145-158.

Rheinwald, J. G., Chakrabarty, A. M. and Gunsalus, I. C. (1973). "A Transmissible Plasmid Controlling Camphor Oxidation in Pseudomonas putida." *Proceedings of the National Academy of Sciences* **70**: 885-889.

Richardson, J. S. (1981). "The anatomy and taxonomy of protein structure." *Advances in Protein Chemistry* **34**: 167-339.

Roberts, S. M. (1998). "Preparative biotransformations: the employment of enzymes and whole-cells in synthetic organic chemistry." *Journal of Chemical Society, Perkin Transactions 1* 157.

Rontgen, W. C. (1896). "On a New Kind of Rays." *Science* **3**: 227-231.

Roseman, S. (2001). "Reflections on Glycobiology." *Journal of Biological Chemistry* **276**: 41527-41542.

Rossmann, M. (1990). "The molecular replacement method." *Acta Crystallographica Section A* **46**: 73-82.

Rossmann, M. G. and Blow, D. M. (1962). "The detection of sub-units within the crystallographic asymmetric unit." *Acta Crystallographica* **15**: 24-31.

Schenk, G., Duggleby, R. G. and Nixon, P. F. (1998). "Properties and functions of the thiamin diphosphate dependent enzyme transketolase." *The International Journal of Biochemistry & Cell Biology* **30**: 1297-1318.

Schmid, A., Dordick, J. S., Hauer, B., Kiener, A., Wubbolts, M. and Witholt, B. (2001). "Industrial biocatalysis today and tomorrow." *Nature* **409**: 258-268.

Schneider, G. and Lindqvist, Y. (1993). "Enzymatic Thiamine Catalysis: Mechanistic Implications from the Three-Dimensional Structure of Transketolase." *Bioorganic Chemistry* **21**: 109-117.

Schneider, G. and Lindqvist, Y. (1998). "Crystallography and mutagenesis of transketolase: mechanistic implications for enzymatic thiamin catalysis." *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1385**: 387-398.

Schneider, G., Sundstrom, M. and Lindqvist, Y. (1989). "Preliminary crystallographic data for transketolase from yeast." *Journal of Biological Chemistry* **264**: 21619-21620.

Schoffers, E., Golebiowski, A. and Johnson, C. R. (1996). "Enantioselective synthesis through enzymatic asymmetrization." *Tetrahedron* **52**: 3769-3826.

Schulze, B. and Wubbolts, M. G. (1999). "Biocatalysis for industrial production of fine chemicals." *Current Opinion in Biotechnology* **10**: 609-615.

Selivanov, V. A., Kovina, M. V., Kochevova, N. V., Meshalkina, L. E. and Kochetov, G. A. (2003). "Studies of thiamin diphosphate binding to the yeast apotransketolase." *Journal of Molecular Catalysis B: Enzymatic* **26**: 33-40.

Sevestre, A., Hélaine, V., Guyot, G., Martin, C. and Hecquet, L. (2003). "A fluorogenic assay for transketolase from Saccharomyces cerevisiae." *Tetrahedron Letters* **44**: 827-830.

Sheldon, R. A., Arends, I. W. and Hanefeld, U. (2007). *Green Chemistry and Catalysis*. Weinheim, Germany, Wiley-VCH-Verlag GmbH & Co. KGaA.

Sheldrick, G. (2008). "A short history of SHELX." *Acta Crystallographica A* **64**: 112-122.

Sheng, D., Ballou, D. P. and Massey, V. (2001). "Mechanistic studies of cyclohexanone monooxygenase: chemical properties of intermediates involved in catalysis." *Biochemistry* **40**: 11156–11167.

Smith, M. E. B., Kaulmann, U., Ward, J. M. and Hailes, H. C. (2006). "A colorimetric assay for screening transketolase activity." *Bioorganic & Medicinal Chemistry* **14**: 7062-7065.

Solovjeva, O. N. and Kochetov, G. A. (1999). "Inhibition of transketolase by p-hydroxyphenylpyruvate." *FEBS Letters* **462**: 246-248.

Sprenger, G. A., Schörken, U., Sprenger, G. and Sahm, H. (1995). "Transketolase a of Escherichia coli K12." *European Journal of Biochemistry* **230**: 525-532.

Storm, D. R. and Koshland, D. E. (1970). "A Source for the Special Catalytic Power of Enzymes: Orbital Steering." *Proceedings of the National Academy of Sciences* **66**: 445-452.

Straathof, A. J. J., Panke, S. and Schmid, A. (2002). "The production of fine chemicals by biotransformations." *Current Opinion in Biotechnology* **13**: 548-556.

Tanner, A. and Hopper, D. (2000). "Conversion of 4-Hydroxyacetophenone to into 4-Phenyl Acetate by Flavin Adenine Dinucleotide containing Baeyer-Villiger type Monooxygenase." *Journal of Bacteriology* **182**: 6565-6569.

Taylor, D. G. and Trudgill, P. W. (1986). "Camphor revisited: studies of 2,5-diketocamphane 1,2-monooxygenase from Pseudomonas putida ATCC 17453." *Journal of Bacteriology* **165**: 489-497.

Thomas, S. M., DiCosimo, R. and Nagarajan, V. (2002). "Biocatalysis: applications and potentials for the chemical industry." *Trends in Biotechnology* **20**: 238-242.

Turfitt, G. E. (1948). "The microbiological degradation of steroids: 4. Fission of the steroid molecule." *Biochemical Journal* **42**: 376-383.

Turner, N. J. (2000). "Applications of transketolases in organic synthesis." *Current Opinion in Biotechnology* **11**: 527-531.

Usmanov, R. A. and Kochetov, G. A. (1983). "Binding of substrates to baker's yeast transketolase. Function of the anionic group of the donor substrate." *Biochemistry. Academy of Sciences of the USSR* **48**: 478-484.

Vagin, A. A. and Teplyakov, J. (1997). "MOLREP: an automated program for molecular replacement." *Journal of Applied Crystallography* **30**: 1022-1025.

van de Sandt, E. J. A. X. and De Vroom, E. (2000). "Innovations in cephalosporin and penicillin production: painting the antibiotics industry green." *Chimica Oggi* **18**: 72-75.

Veitch, N. J., Maugeri, D. A., Cazzulo, J. J. and Lindqvist, Y. (2004). "Transketolase from Leishmania Mexicana has dual subcelluar localization." *Biochemical Journal* **382**: 759.

Villafranca, J. J. and Axelrod, B. (1971). "Heptulose Synthesis from Nonphosphorylated Aldoses and Ketoses by Spinach Transketolase." *Journal of Biological Chemistry* **246**: 3126-3131.

Vinick, F. J. and Jung, S. (1982). "A superior synthesis of aspartame." *Tetrahedron Letters* **23**: 1315-1318.

Walfridsson, M., Hallborn, J., Penttilä, M., Keränen, S. and Hahn-Hägerdahl, B. (1995). "Xylose-metabolizing Saccharomyces cerevisiae strains overexpressing the TKL1 and TAL1 genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase." *Applied Environmental Microbiology* **61**: 4184–4190.

Wandrey, C., Liese, A. and Kihumbu, D. (2000). "Industrial Biocatalysis: Past, Present, and Future." *Organic Process Research & Development* **4**: 286-290.

Wang, J. L., Martin, P. R. and Singleton, C. K. (1997). "Aspartate 155 of human transketolase is essential for thiamine diphosphate-magnesium binding, and cofactor binding is required for dimer formation." *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1341**: 165-172.

Ward, J. M. and French, C. (1996). "Production and Modification of E. coli Transketolase for Large-Scale Biocatalysis." *Annals of the New York Academy of Sciences* **799**: 11 - 18.

Watson, J. D. and Crick, F. H. C. (1953). "Molecular structure of nucleic acids - a structure for deoxyribose nucleic acid." *Nature* **171**: 737-738.

Wikner, C., Meshalkina, L., Nilsson, U., Backstrom, S., lindqvist, Y. and Schneider, G. (1995). "His103 in Yeast Transketolase is Required for Substrate Recognition and Catalysis." *European Journal of Biochemistry* **233**: 750-733.

Wikner, C., Meshalkina, L., Nilsson, U., Nikkola, M., Lindqvist, Y. and Schneider, G. (1994). "Analysis of an invariant cofactor-protein interaction in thiamin diphosphate-dependant enzymes by site-directed mutagenisis. Glutamic acid 418 in transketolase is essential for catalysis." *Journal of Biological Chemistry* **269**: 32144-32150.

Wikner, C., Nilsson, U., Meshalkina, L., Udekwu, C., Lindqvist, Y. and Schneider, G. (1997). "Identification of Catalytically Important Residues in Yeast Transketolase." *Biochemistry* **36**: 15643-15649.

Winter, J., Schneider, B., Meyenburg, S., Strack, D. and Adam, G. (1999). "Monitoring brassinosteroid biosynthetic enzymes by fluorescent tagging and HPLC analysis of their substrates and products." *Phytochemistry* **51**: 237-242.

Wittig, G. and Schöllkopf, U. (1954). "Über Triphenyl-phosphin-methylene als olefinbildende Reagenzien I." *Chemische Berichte* **87**: 1318.

Wright, J. L., Hu, T., Mclachlan, J. L., Needham, J. and Walter, J. A. (1996). "Biosynthesis of Dtx-4 - Confirmation of a Polyketide Pathway, Proof of a

Baeyer-Villiger Oxidation Step, and Evidence for an Unusual Carbon Deletion Process." *Journal of the American Chemical Society* **118**: 8757-8758.

www.ebi.ac.uk/pdbsum/. (2008). "www.ebi.ac.uk/pdbsum."

www.pdb.org. (2008). "RSCB PDB Protein Data Bank."

Yoshida, H., Kamitori, S., Agari, Y., Iino, H., Kanagawa, M., Nakagawa, N., Ebihara, A., Kuramitsu, S. and Yokoyama, S. (2007). "X-ray Structure of *Thermus thermophilus* HB8 TT0505." *TO BE PUBLISHED*.

Zhang, M., Eddy, C., Deanda, K., Finkelstein, M. and Picataggio, S. (1995). "Metabolic Engineering of a Pentose Metabolism Pathway in Ethanologenic Zymomonas mobilis." *Science* **267**: 240-243.