ANALYSIS OF THE TRANSITION STATE OF DIHYDROFOLATE REDUCTASE

A thesis submitted to Cardiff University for the degree of

Doctor of Philosophy by

William Michael Dawson

Supervisor: Prof Rudolf K. Allemann

2014

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ………………………………………… (candidate)

Date …………………………

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Signed ………………………………………… (candidate)

Date …………………………

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ………………………………………… (candidate)

Date …………………………

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ………………………………………… (candidate)

Date…………………………

ABSTRACT

The role for protein dynamics in the transition states (TS) of enzyme reactions has been debated over decades. Dihydrofolate reductase (DHFR) catalyses the NADPH-dependent reduction of dihydrofolate to tetrahydrofolate and has long been considered a paradigm of enzymology. Numerous studies on DHFR have provided strong evidence that there is no coupling between the long-range motions on nanosecond to millisecond timescales to the chemical coordinate. However, the role of femtosecond (fs) bond vibrations in the TS has not been fully investigated. This investigation focused on understanding how these fast protein vibrations affect enzyme catalysis. A thermophilic DHFR from *Geobacillus stearothermophilus* (BsDHFR) was investigated kinetically by complete enzyme isotope substitution. Our studies indicated that, whilst protein vibrations do couple to the reaction coordinate, they do not affect the height or width of barrier crossing. Instead, dynamic coupling enhances the frequency of dynamic recrossing. In line with the other DHFR investigations by enzyme isotope substitution, efficient enzymes tend to reduce dynamic coupling as a mean to maximise enzyme catalytic efficiency. The transition state in DHFR was also characterised by α-secondary hydrogen and heavy atom kinetic isotope effects (KIEs). Secondary KIEs were measured for DHFR isolated from *Escherichia coli*, *Moritella profunda* and *G. stearothermophilus*. The high resemblance in the magnitude and temperature dependence of the measured α-secondary KIE implied that the reaction ready configuration is essentially the same among these enzymes. Carbon isotope effect measurements were measured for EcDHFR. The reacting carbon in NADPH showed a profound isotope effect at low temperature. Further analysis by measuring hydride KIE indicated the hypothetical "promoting" motion is unlikely to act via the C4 atom of NADPH. The carbon isotope effect likely reports on the recrossing events or the reorganisation effect that occur along the transition state dividing surface.

ACKNOWLEDGEMENT

Firstly I would like to thank my supervisor, Professor Rudolf Allemann, for allowing me the opportunity to work on this exciting research project, and for support and guidance throughout my Ph.D. Special thanks are also given to Dr. Joel Loveridge and Dr. Louis Luk for their suggestions and advice throughout my project, putting up with my (numerous) questions, and for proof reading this thesis. I would also like to thank Dr. Verónica González and Dr. Rob Mart for all their help and advice in the lab during my research and Dr. Niek Buurma for helpful discussions regarding my data.

Thanks to all the DHFR members and the rest of the Allemann group that have made my studies so enjoyable. Thanks to Dan for the morning discussions over pastry, and Mel for making our office the envy of the rest. Special mention has to go to Louis, Rob and Sarah (the original Thursday crew) and Vero for the good times in the Pen, I don't think I would have got through to the end without you guys!

I would like to thank my family for all their help and encouragement over the years, even if they still don't understand what my project is about! Finally, a special thank you to Amy, I couldn't have done it without your support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

LIST OF ABBREVIATIONS

6PGDH	6-phosphogluconate dehydrogenase
A	Adenine
ABD	Adenosine binding domain
ADP	Adenosine diphosphate
A$_H$	Arrhenius prefactor
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
a.m.u	Atomic mass units
BAP	Bovine alkaline phosphatase
bPNP	Bovine purine nucleoside phosphorylase
BSA	Bovine serum albumin
BsDHFR	*Geobacillus stearothermophilus* dihydrofolate reductase
C	Cytosine
CD	Circular dichroism
CV	Column volume
DAD	Donor-acceptor distance
DHAP	Dihydroxyacetone phosphate
DHF	Dihydrofolate
DHFR	Dihydrofolate reductase
DHFS	Dihydrofolate synthetase
DHNA	Dihydroneopterin aldolase
DHPS	Dihydropteroate synthetase
DNA	Deoxyribonucleic acid
dTMP	Deoxythymidine monophosphate
E	Enzyme
E	Energy
***E*$_a$**	Activation energy
EA-VTST/MT	Ensemble averaged-Variational transition state theory/multidimensional tunnelling
EcDHFR	*Escherichia coli dihydrofolate* reductase
E$_{F(m)}$	Distance sampling term in Marcus-like tunnelling theories
EIE	Equilibrium isotope effect

ES	Enzyme substrate complex
ES‡	Activated enzyme substrate complex
ESI-MS	Electrospray ionising mass spectroscopy
FAD/FADH$_2$	Flavin adenine dinucleotide
F$_{(m)}$	Franck-Condon term
FMN	Flavin mononucleotide
FPLC	Fast protein liquid chromatography
FRET	Fluorescence resonance energy transfer
fs	Femtosecond
G	Guanine
G6PDH	Glucose-6-phosphate dehydrogenase
GMP	Guanosine monophosphate
GS	Ground state
GTP	Guanosine triphosphate
GTP cyc I/GTPCI	Guanosine triphosphate cyclohydrolase I
GTPCII	Guanosine triphosphate cyclohydrolase II
GTPCIII	Guanosine triphosphate cyclohydrolase III
h	Planck's constant
H$^+$	Proton
H$^-$	Hydride
H$^{\bullet}$	Hydrogen radical
HK	Hexokinase
hPNP	Human purine nucleoside phosphatase
HPLC	High pressure liquid chromatography
HPPK	Hydroxymethyldihydroneopterin pyrophosphokinase
IPTG	isopropyl-β-D-1-thiogalactopyranoside
$K^{\ddagger}$	Psuedo-equilibrium constant of transition state formation
k_B	Boltzmann constant
k_H/k_D	Hydride/deuteride transfer rate constant
kJ	Kilojoules
K_M	Michaelis constant
KIE	Kinetic isotope effect
LAO	L-Aspartate oxidase

LCMS	Liquid chromatography mass spectroscopy
LD	Loop domain
mol	Moles
MR	Morphinone reductase
MpDHFR	*Moritella profunda* dihydrofolate reductase
MptA	Archaea guanosine triphosphate cyclohydrolase
ms	Millisecond
MTEN	Mes tris ethanolamine sodium chloride
N/NADP$^+$	Nicotinamide adenine dinucleotide phosphate
NAAD	Nicotinic acid adenine dinucleotide
NAD$^+$	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
NADPH/NH	Reduced nicotinamide adenine dinucleotide phosphate
NADK	Nicotinamide adenine dinucleotide kinase
NADS	Nicotinamide adenine dinucleotide synthetase
NAMN	Nicotinic acid mononucleotide
NAMNT	Nicotinic acid mononucleotide adenylyltransferase
NMR	Nuclear magnetic resonance
ns	Nanosecond
OMP	Orotidine monophosphate
ps	Picosecond
P	Product
pABA	Para-aminobenzoic acid
PETNR	Pentaerythritol tetranitrate reductase
P$_i$	Phosphate
PNP	Purine nucleoside phosphorylase
PP$_i$	Pyrophosphate
PRPP	Phosphoribose pyrophosphate
PRPPS	Phosphoribose pyrophosphate synthetase
Q	Bell correction factor
QA	Quinolinic acid
QAPRT	Quinolinic acid phosphoribosyltransferase
QS	Quinolinate synthetase

R	Universal gas constant
R5P	Ribose-5-phosphate
R5PI	Ribose-5-phosphate isomerase
RRC	Reaction ready configuration
s	Second
S	Substrate
S‡	Activated substrate
SAM	S-adenosyl methionine
T	Temperature
T	Thymine
THF	Tetrahydrofolate
TmDHFR	*Thermotoga maritima* dihydrofolate reductase
Tris	Tris(hydroxymethyl)aminomethane
TS/TS‡	Transition state
TST	Transition state theory
VTST	Variational transition state theory
XGPRT	Xanthine-guanine phosphoribosyl transferase
ZPE	Zero point energy
2°	Secondary
Å	Angstrom
βME	β-Mercaptoethanol
ΔG^{o}	Gibbs free energy of the reaction
ΔG‡	Gibbs free energy of activation
ΔH‡	Gibbs free enthalpy of activation
ΔS‡	Gibbs free entropy of activation
μs	Microsecond
Γ	Transmission coefficient
κ	Tunnelling coefficient
γ	Recrossing coefficient
λ	DeBroglie wavelength
λ	Marcus reorganisation energy

Through this work, amino acids are expressed as their triple or single letter code:

Amino Acid	3 letter code	1 letter code
L-Alanine	Ala	A
L-Arginine	Arg	R
L-Asparagine	Asn	N
L-Aspartic acid	Asp	D
L-Cysteine	Cys	C
L-Glutamic acid	Glu	E
L-Glutamine	Gln	Q
L-Glycine	Gly	G
L-Histidine	His	H
L-Isoleucine	Ile	I
L-Leucine	Leu	L
L-Lysine	Lys	K
L-Methionine	Met	M
L-Phenylalanine	Phe	F
L-Proline	Pro	P
L-Serine	Ser	S
L-Threonine	Thr	T
L-Tryptophan	Trp	W
L-Tyrosine	Tyr	Y
L-Valine	Val	V

1   INTRODUCTION

1.1   Enzymes - nature's catalysts

Enzymes are capable of enhancing the reaction rate of biological processes by many orders of magnitude (Figure 1.1A). Orotidine monophosphate (OMP) decarboxylase is able to increase the rate of reaction by a factor of 10^{17} compared to the rate of spontaneous OMP decarboxylation (1). Like all catalysts, enzymes lower the activation energy of the reaction (Figure 1.1B) in this case through the formation of the enzyme-substrate (ES) complex. One of the earliest theories to explain this extraordinary activity was proposed by Fischer in 1894, who published the "lock and key" theory in which enzyme active sites are specifically designed to tightly bind the substrate (2). Later, Eyring developed transition state theory (TST), a key component of which is the existence of an activated ES complex ($ES^{\ddagger}$). This was modified by Pauling soon after, who proposed the preferential binding of the transition state as a mechanism for enzyme activity (3, 4).

The increasing experimental results that could not be explained by the "lock and key" model, including examples of non-competitive inhibition, led to Koshland proposing the "induced fit" mechanism (5, 6). This theory suggested that upon binding the substrate(s), the enzyme undergoes conformational changes, forming an ES complex that is closer in structure to the transition state. This is thought to reduce the activation energy by three main methods, including bringing the reacting groups of the substrates closer together, weakening bonds that need to be broken and/or stabilising the transition state itself. Finally, Warshel has shown that enzymes can form a favourable polar environment, termed "electrostatic preorganisation", and this contributes to the increased catalytic power (7).

Subsequent studies, including allosteric control of enzymes and protein NMR investigations, provided evidence that enzymes could adopt different conformations, often with wide ranging catalytic properties (8, 9). However, the intrinsic catalytic power of enzymes is still not well understood. This is highlighted by results from research into the design of artificial enzymes. Despite numerous efforts, the ability to mimic nature and produce synthetic enzymes that are not only able to catalyse reactions, but do so under typical physiological conditions (water,

pH 7, ambient temperature) is still incredibly challenging, with few cases reported to date (see recent reviews for more information (10-13)).

B
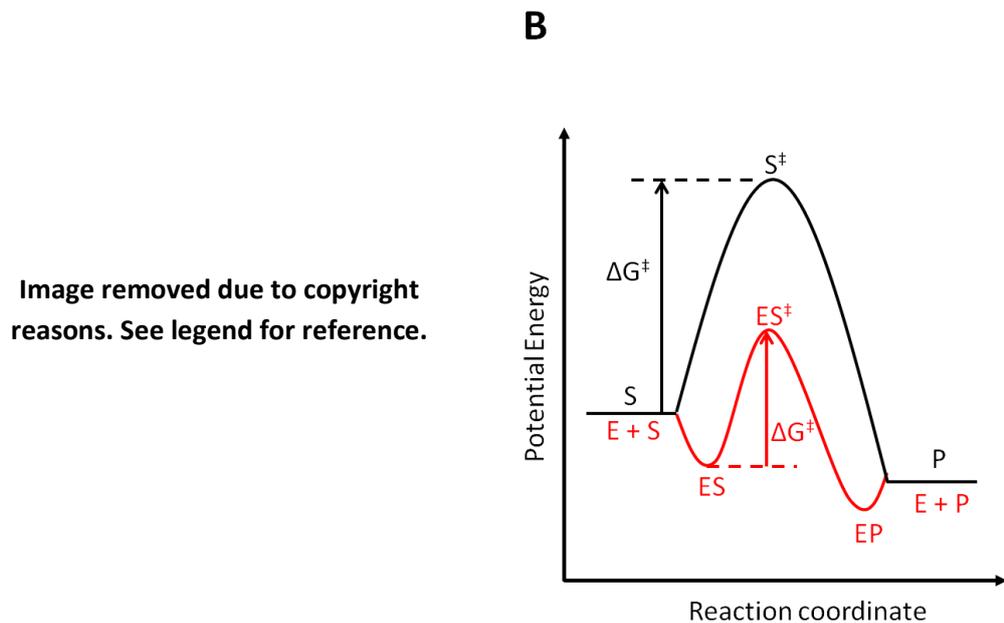
Figure 1.1 A) Comparison of the rate of a typical enzyme catalysed reaction with uncatalysed spontaneous reaction rates. Figure taken from Wolfenden, R., (2003) Thermodynamic and extrathermodynamic requirements of enzyme catalysis, Biophysical Chemistry, 105, 559-572. B) Reaction pathway of an uncatalysed reaction (black) and a simple enzyme-catalysed reaction (red). Formation of the activated enzyme-substrate complex (ES‡) from the ES requires less energy than formation of the uncatalysed activated substrate (S‡).

1.1.1 Enzyme dynamics

Evidence of proteins being able to adopt multiple conformations led to the idea that they were dynamic entities, and it is thought that understanding this phenomenon will shed light on the incredible efficiency of enzymes. Enzyme motions span a wide range of time scales: bond vibrations (femtoseconds), rotation of side-chains at a protein's surface (picoseconds), hinge-bending at the interface between domains (pico-nanoseconds), rotation of side-chains buried in the interior

of a protein (micro-milliseconds) and allosteric transitions (micro-milliseconds) (14-16). Whilst most large scale conformational changes, such as allosteric control, are widely accepted, the contribution of faster motions, typically those that occur after formation of the ES complex, have proved much more controversial (17-19).

Motions are implicit in chemical reactions, with vibrations on an atomistic scale affecting bond lengths, angles and charge distributions, all of which enable (enzymatic) reactions to occur. The term "enzyme dynamics" is used specifically to refer to non-statistical motions. Statistical motions are those in thermal equilibrium with the surrounding environment and encompass the vast majority of enzyme motions. On the other hand, non-statistical motions or enzyme dynamics are not in equilibrium with the surrounding thermal environment. Controversy arises when enzyme dynamics (encompassing a wide range of timescales) are claimed to couple directly to enzyme catalysis, enhancing the rate of reaction. A typical TS exists for 50-100 fs (18), and it is how this short-lived intermediate is affected by motions on a similar timescale (*i.e.* bond vibrations), and those on a much longer time scale (*e.g.* ms loop movements), that has sparked numerous debates in the literature.

Recent enzymology studies, both theoretical and experimental, have focused on determining how enzyme dynamics contribute to enhancing reaction rates. The large number of enzymatic reactions that involve hydrogen transfer (H^+, H^- or $H^{\bullet}$) have provided an excellent framework for these investigations (20-25). Numerous catalytic models have been proposed to explain the results from these investigations, and will be explained below.

1.1.2   Transition state theory and isotope effects

Transition state theory (TST) is often used as a simple, but nonetheless effective, model to explain enzyme catalysis (26). TST relies on two basic assumptions: 1) that the reaction follows a reaction coordinate and must pass through a maximum point, known as the transition state (TS), and 2) that the concentration of the TS is in equilibrium with the substrate(s) of the reaction. The concentration of the TS is governed by the law of thermodynamics and leads to the equation:

$$[TS^\ddagger] = K^\ddagger[A][B]$$

where A and B are the substrates, and $K^\ddagger$ is the psuedo-equilibrium constant of the formation of the TS given by:

$$K^\ddagger = e^{\frac{-\Delta G^\ddagger}{RT}}$$

where $\Delta G^\ddagger$ is the Gibbs free energy of formation of the TS. The concentration of the TS is thus:

$$[TS^\ddagger] = [A][B]e^{\frac{-\Delta G^\ddagger}{RT}} = [A][B]e^{\frac{\Delta S^\ddagger}{R}}e^{\frac{-\Delta H^\ddagger}{RT}}$$

As TST assumes that the TS is in equilibrium with the substrates, the rate of reaction is governed by the concentration of the TS and the frequency with which the TS decomposes to form the products. It is assumed that the frequency of the decomposition of the transition state is proportional to the vibrational frequency of relevance (*v*) and the transmission coefficient (Γ):

$$\frac{-\delta[A][B]}{\delta t} = v\Gamma[TS^\ddagger]$$

where *v* is equal to $k_B T/h$ where k_B and *h* are the Boltzmann constant and Planck's constant respectively. The transmission coefficient is defined as the factor that includes all effects not incorporated in the TST constant (27). In most reactions Γ is assumed to be unity (see below for corrections to this theory). This leads to the Eyring equation:

$$k = \Gamma\left(\frac{k_B T}{h}\right)e^{\frac{\Delta S^\ddagger}{R}}e^{\frac{-\Delta H^\ddagger}{RT}}$$

One of the most effective tools for studying enzyme reactions is the use of kinetic isotope effects (KIE), defined as:

$$KIE = \frac{k_L}{k_H}$$

where k_L is the rate constant of the reaction involving the light isotope and k_H is the rate constant of the reaction involving the heavier isotope. In regards to enzyme

studies the most common isotope effects measured are primary isotope effects involving the isotopes of hydrogen – deuterium and tritium. A primary isotope effect occurs when the isotope is part of the bond being broken during the reaction (28).

In semi-classical TST, KIEs are due solely to the difference in zero point energy (ZPE) between the bonds to light and heavy isotope (Figure 1.2). Each bond has many different energy levels, and these levels are expressed as:

$$e_n = \left(n + \frac{1}{2}\right) h v' \qquad n = 0,1,2...$$

therefore when n = 0,

$$e_0 = \frac{1}{2} h v'$$

where v' is the frequency of the stretching vibration of a bond:

$$v' = \frac{1}{2\pi} \sqrt{\frac{k'}{\mu}}$$

where k' is the force constant of the bond and μ is the reduced mass:

$$\mu = \frac{m_1 m_2}{m_1 + m_2}$$

The presence of the heavier isotope results in a smaller stretching frequency of the bond due to the greater reduced mass. This causes a lower ZPE, hence heavier isotopes have a larger $\Delta H^\ddagger$ to reach the TS compared to the light isotope. At 25 °C the difference in ZPE between a C-H and a C-D bond is 4.8 kJ mol^{-1}, and following substitution into the Eyring equation, results in a k_H/k_D value of 6.9:

$$\frac{k_H}{k_D} = \frac{e^{\frac{-\Delta H^\ddagger}{RT}}}{e^{\frac{-\Delta H^\ddagger + 4800}{RT}}} = e^{\frac{4800}{RT}} = 6.9$$
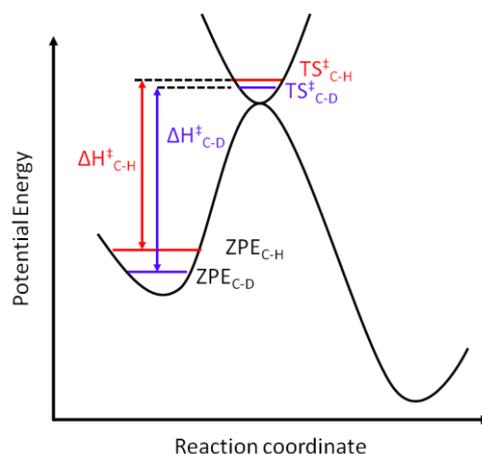
Figure 1.2 Reaction profile showing the difference in ZPE between a C-H and C-D bond (red and blue respectively) and the resulting differences in ΔH‡.

The Arrhenius law can also be used to relate the rate of reaction to the energy barrier that has to be overcome during the reaction:

$$k = A\, e^{\frac{-E_a}{RT}}$$

where A is the pre-exponential factor and E_a is the activation energy. Similarly to the Eyring equation, the difference in ZPE between the light and heavy isotope causes the perceived KIE:

$$\frac{k_H}{k_D} = \frac{A_H}{A_D}\, e^{\frac{1}{2}h\,v_H - \frac{1}{2}h\,v_D}$$

where A_H and A_D are pre-exponential factors for hydrogen and deuterium transfer and 1/2 hv_H and 1/2hv_D are the ZPE of a C-H and C-D bond respectively. In semi-classical theory the A_H/A_D ratio is 1, *i.e.* at infinite temperature the difference in activation energy disappears and so there is no difference between k_H and k_D.

Early KIE theories focused on explaining how smaller than expected values could be obtained. Eyring's theory assumes sole importance of the stretching vibrational frequency becoming translational frequency, a characteristic of the TS. This simplification ignores contributions from bending vibrational frequencies, vibrational modes of all other atoms in the molecule, and new vibrational

frequencies introduced at the TS (29). It was found that these other vibrational modes at the TS could reduce the difference between the ZPE at the GS and TS, resulting in lower KIEs (29). This led to the theory that the magnitude of the KIE would reflect the position of the TS. A symmetric TS results in a maximum KIE value, *i.e.* ~7, whereas a TS resembling either the reactants or products would result in a lower KIE (29). However, with the discovery that hydrogen tunnelling may play a significant role in enzyme catalysis, it was found that models incorporating quantum mechanics were required to explain the observed results.

1.1.3 The Bell tunnelling correction

As stated, the semi-classical models surrounding the Eyring and Arrhenius equations fail to explain larger than expected KIEs, thought to be attributed to tunnelling. Bell proposed a modification of semi-classical theory to incorporate the uncertainty of a particle's position as it passed through the reaction coordinate, allowing the molecule to "tunnel" through the reaction barrier (Figure 1.3) (30). Bell used the theory, first proposed by de Broglie, that a particle's uncertainty of position is due to its wavelength (λ):

$$\lambda = h\sqrt{2mE}$$

where m is the mass and E is the energy of the molecule. As can be seen, the lighter the molecule the greater the de Broglie wavelength and therefore the greater level of uncertainty of a particle's position. This is particularly relevant to the isotopes of hydrogen as the relatively large differences in mass mean significant differences in the respective de Broglie wavelengths (Table 1.1). As a typical reaction coordinate is approximately 1 Å or less, these de Broglie wavelengths signify a high probability of tunnelling occurring through the reaction barrier.
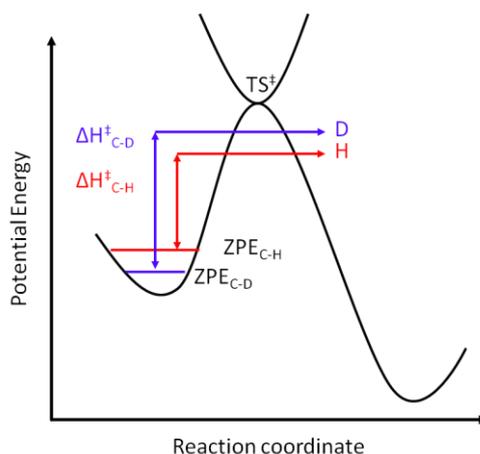
Figure 1.3 Reaction coordinate highlighting the barrier width at which H and D tunnelling can occur.

Table 1.1 Comparison of the de Broglie wavelengths of H, D T and carbon (assuming an energy of 20 kJ mol⁻¹).

	H	D	T	C
Mass (amu)	1	2	3	12
λ (Å)	0.63	0.45	0.36	0.18

Bell's theory focuses on a minor adaption to the Arrhenius equation by the correction factor, Q (30):

$$k = QAe^{\frac{-E_a}{RT}}$$

where Q is:

$$Q = \frac{e^{\frac{E_a}{RT}}}{\beta - \frac{E_a}{RT}} \left(\beta e^{\frac{-E_a}{RT}} - \frac{E_a}{RT} e^{-\beta} \right)$$

and β is:

$$\beta = \frac{a\pi^2 \sqrt{2mE_a}}{h}$$

where a is the width of the barrier, m the mass of the tunnelling particle, and E_a the barrier height. The mass of the particle has a large effect on Q due to the exponential involving β. The large difference in mass between hydrogen and deuterium/tritium can therefore result in large KIEs.

Figure removed due to copyright reasons.

See legend for reference

Figure 1.4 Arrhenius plot of hydrogen transfer (above) and KIE (below) consistent with the Bell tunnelling correction. Figure taken from Stojković, V. and Kohen, A. (2009), Enzymatic H Transfers: Quantum Tunneling and Coupled Motion from Kinetic Isotope Effect Studies. Isr. J. Chem., 49: 163–173.

The temperature dependence of the KIE and the A_H/A_D ratio has also been used to infer a contribution from tunnelling during the reaction (30). TST with a Bell correction predicts that at high temperatures *i.e.* with excess thermal energy, the reaction will proceed over the barrier with very little contribution from tunnelling. This will result in a temperature dependent KIE and $A_H/A_D \approx 1$ (region I, Figure 1.4). At low temperatures, tunnelling becomes the main contribution to the reaction rate, resulting in temperature independent KIEs and an $A_H/A_D > 1$ (region III, Figure 1.4). However, these are the extremes of the scale, with enzyme catalysis typically studied in a narrow temperature range between the two (5 - 80 °C). Over this scale hydrogen is more sensitive to tunnelling than deuterium and this can result in strong temperature dependence of the KIE and $A_H/A_D << 1$ (region II, Figure 1.4) (30).

1.1.4 Recent tunnelling models of enzyme catalysis

Although the Bell tunnelling correction went some way towards explaining results that did not fit the semi-classical TST model, it is widely accepted that this model has a number of flaws with regards to enzyme catalysis. The Bell tunnelling model treats the reaction barrier as static and focuses on the reaction coordinate of the hydrogen atom (or deuterium), ignoring any contribution that heavy atom motions might play in the reaction. A number of experimental results from a range of enzymes couldn't be explained by TST with a simple tunnelling correction (25, 31-35). These violations include large KIEs (> 80) and a greater than unity A_L/A_H ratio, both with temperature dependent and temperature independent KIEs (17, 36). This led to the development of a number of models based on Marcus theory of electron transfer.

Marcus theory of electron transfer states that when electrons transfer between two orbitals, the donor and acceptor, they do not change energy. Therefore, for this transfer to occur both orbitals must assume the same energy level. This requires energy to be put into the system in order to distort the reactant-solvent and/or product-solvent ensembles to achieve this equal energy level (37). In Marcus theory, this energy is termed reorganisation energy (λ, Figure 1.5). Marcus-like models applied to hydrogen tunnelling replace the transferring electron with a hydrogen atom. It is reorganisation of the heavy nuclei of the reactant and product to equal energy levels that causes the barrier to tunnelling, rather than the stretching of the C-H bond in simpler models discussed above (17).
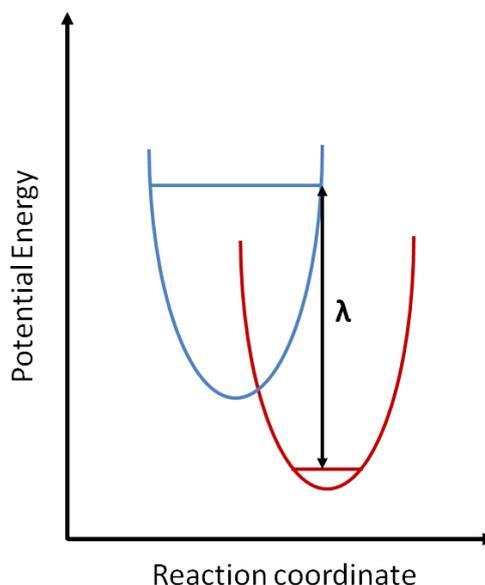
Figure 1.5 Reaction coordinate of electron transfer using Marcus theory. An electron is transferred from the donor orbital (blue) to the acceptor orbital (red). The proposed reorganisation energy is shown by λ.

Marcus-like models can be described using the general equation:

$$k = Ce^{\frac{(\Delta G^o + \lambda)^2}{4\lambda RT}} \int_{r_2}^{r_1} e^{F(m)} e^{\frac{-E_{F(m)}}{k_B T}} \, dDAD$$

where C is an isotope independent pre-exponential value, the first exponential is the Marcus term that relates the reorganisation energy (λ) to the driving force of the reaction ΔG^o, the second exponential is the Franck-Condon term (F(m)) and the final exponential includes a distance sampling term ($-E_{F(m)}$) (36). The Marcus term causes the rate of reaction to be dependent on the reorganisation energy and driving force as these will affect the transfer of the particle through the reaction barrier. This term is temperature dependent but whilst this term is also mass dependent the effect of isotopic change is thought to be minimal in relation to the size of primary isotope effects influenced by the second exponential (36). The Franck-Condon term is temperature independent but, as stated, is greatly affected by the mass of the transferring particle. The third exponential, known as the distance sampling term, accounts for the sensitivity of tunnelling to the donor-

acceptor distance (DAD), and this term is both temperature and isotope sensitive. The Franck-Condon term and the distance sampling term are integrated for all the conformations sampled by the DAD (36).

These Marcus-like models include rate-promoting vibrations (38, 39), vibrationally enhanced tunnelling (25, 40) and environmentally coupled tunnelling (41, 42). Rate promoting vibrations, proposed by Schwartz and coworkers, states that a network of long range motions is coupled to hydrogen tunnelling by modulating the DAD (38, 39). The opposite view is taken by Scrutton *et al.*, who argued that long-range motions are not coupled to the reaction coordinate. They proposed that short-range vibrations promote hydrogen tunnelling by reducing the barrier width *i.e.* vibrationally enhanced tunnelling (25, 40). Klinman's model, environmentally coupled tunnelling, predicts that two sets of motions are linked to enzyme catalysis. The first set of motions, known as preorganisation, allow the enzyme to adopt a reaction ready conformation. The second set of motions then modulate the DAD, known as reorganisation (not to be confused with reorganisation energy in relation to Marcus theory discussed above), and directly couple to the barrier width and thus couple to tunnelling (41, 42).

Truhlar and coworkers also presented a model for enzyme catalysis based on a generalised TST equation. In this case the correcting factor, Γ(T), is split into the different factors affecting the rate of reaction (27):

$$k = \Gamma(T)(C^0)\frac{k_B T}{h}e^{\frac{-\Delta G^\ddagger}{RT}}$$

where (C^0) is the standard-state concentration and Γ(T) comprises γ(T) which accounts for dynamic recrossing of the barrier, κ(T) which accounts for the tunnelling contribution in the reaction and g(T) which accounts for deviations in the equilibrium distribution (27). Recrossing refers to reaction trajectories that, having crossed the transition state dividing surface from reactants to products, return across the dividing surface to the reactant side (43). This equation has been used as the basis for ensemble averaged-variational TST and multidimensional tunnelling (EA-VTST/MT) model (43, 44). VTST assumes a simple reaction will have a transition

state that has a significant rate of recrossing, and thus the position of the transition state can be chosen to minimise this rate. EA-VTST/MT takes into account the complexity of enzyme systems, in that the reaction may proceed by numerous reaction paths, each of which has a transition state that can be chosen to give the lowest rate of recrossing, and thus takes an average of the overall ensemble (45, 46). The multi-dimensional tunnelling calculations take into consideration that tunnelling can occur through multiple pathways, rather than the one-dimensional pathway assumed by the simple tunnelling models discussed earlier (43). EA-VTST/MT couples enzyme motions to the reaction coordinate, with the reasoning that any differences between the enzyme reaction and the non-enzyme catalysed reaction must occur from the added complexity of the environment surrounding the reaction coordinate. This theory has been applied to a number of enzymes (see reviews for more detail (27, 43, 47)) and has shown that TST can still be used to explain the non-classical results observed in numerous enzymes.

The Marcus-like models focus on the theory that enzyme motions (whether long-range or short-range) are directly coupled to the reaction coordinate or promote hydrogen tunnelling. By comparison EA-VTST/MT theory developed by Truhlar does not give a general trend to the role of enzyme motions. It is worth noting though that any coupling of long-range motions or multiple conformations of enzymes are incorporated into the EA-VTST part of the calculations, and DAD fluctuations are incorporated into the MT part of the calculations (44, 46, 48). As a result the coupling of enzyme dynamics to the reaction coordinate are often used to explain the results from EA-VTST/MT calculations.

However, the idea that enzyme dynamics and hydrogen tunnelling play a significant role in the chemical step of enzyme catalysis is strongly disputed by Warshel (49-52). This view is also taken by Mulholland and coworkers, who showed that a simple model involving multiple enzyme conformations could explain many of the non-classical findings without involving enzyme dynamics (53, 54). However, whilst Mulholland and coworkers have fitted data from numerous enzymes to their model, no direct evidence of multiple conformations has been reported in these enzymes. Calculations performed by Warshel and coworkers have shown that

electrostatic effects have the greatest effect on enzyme catalysis. The active site of the enzyme was found to act as a much more efficient "solvent" for the reaction coordinate than the comparable reaction in aqueous solution. This was due to the reorganisation energy of water molecules being much higher, as most of the reorganisation of the enzyme active site towards stabilising the TS‡ is performed during protein folding (49). Warshel and coworkers have also shown that mutations to enzymes that cause decreases in reaction rates and are then used as evidence to suggest promoting motions are actually due to preorganisation effects. It was shown that mutations to enzymes actually change the conformational sampling of an enzyme due to the change in potential energy surface, rather than providing evidence of coupled motions (50, 51). Further evidence that non-statistical motions do not play a key part in enzyme catalysis of the chemical step is the similarity between the transmission coefficients of the catalysed and uncatalysed reaction, as well as the similarity in vibrational modes between the enzyme reaction and the equivalent reaction in solution (55). It is perhaps worth noting that Truhlar, through the use of EA-VTST, stated that the effects contained in the transmission coefficient were very small compared to the overall rate enhancement of enzymes from reducing the free energy of activation (typically 1 or 2 orders of magnitude from the transmission coefficient compared to rate enhancements up to 17 orders of magnitude for OMP decarboxylase) (43).

Overall, enzyme dynamics and the role they play in catalysis is a topic that is hotly debated, with many contrasting viewpoints. More evidence on the role that enzyme motions play in enzyme catalysis is therefore needed to allow a consensus model to be adopted. As explained above, enzymes involved in hydride transfer are ideal targets for investigating this possible relationship between enzyme dynamics and catalysis. Studies involving alcohol dehydrogenase (20, 21), soybean lipoxygenase (22), methylamine dehydrogenase (25) and dihydrofolate reductase (23, 24) have contributed significantly (amongst others), both in support and against, to the proposed tunnelling models. In particular, dihydrofolate reductase has been the focus of a large number of experimental and computational investigations, the specific findings of which will be discussed in more detail below.

1.2 Dihydrofolate reductase

Dihydrofolate reductase (DHFR) catalyses the reduction of 7,8-dihydrofolate (DHF) to 5,6,7,8-tetrahydrofolate (THF), via the oxidation of NADPH (56). The H_{Re} hydride is transferred from the C4 position of NADPH to the C6 position of DHF, and a proton is transferred to the N5 position (Scheme 1.1)(57, 58).
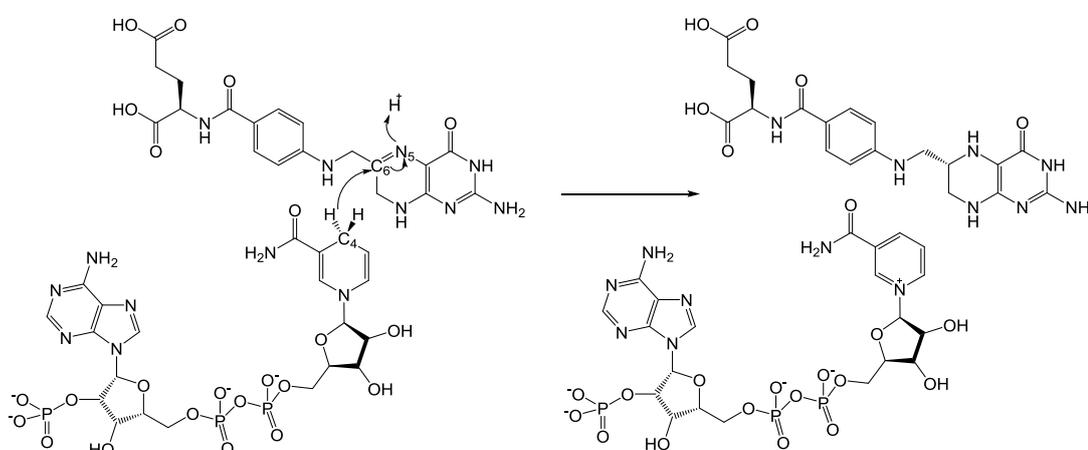
Scheme 1.1 Reaction catalysed by DHFR

The product of this reaction, THF, is involved in a number of reactions within cells due to its capability as a carrier of one-carbon units at the N5, N10 or both positions (Scheme 1.2) (59). These cofactors are used for amino acid biosynthesis including serine, glycine and methionine, cofactor biosynthesis including S-adenosyl methionine and both purine and pyrimidine nucleotides. The majority of these reactions result in the recycling of THF or its derivatives, however, the biosynthesis of thymidine monophosphate (dTMP) produces DHF as a product. THF is converted to 5,10-methylene-THF via serine hydroxymethyl transferase and utilised by thymidylate synthase in the production of dTMP, resulting in the production of DHF. DHF is then recycled via DHFR, allowing the levels of THF in the cell to be maintained (60). The importance of DHFR results in this enzyme being found in almost every organism, although a bifunctional DHFR-thymidylate synthase is found in some primitive protozoa (61). Some bacterial species have evolved without the chromosomal gene encoding DHFR (*folA*) (62-64). Many of these species have

evolved a different thymidylate synthase that utilises $FADH_2$, producing tetrahydrofolate rather than dihydrofolate as the product of the reaction (62, 63).
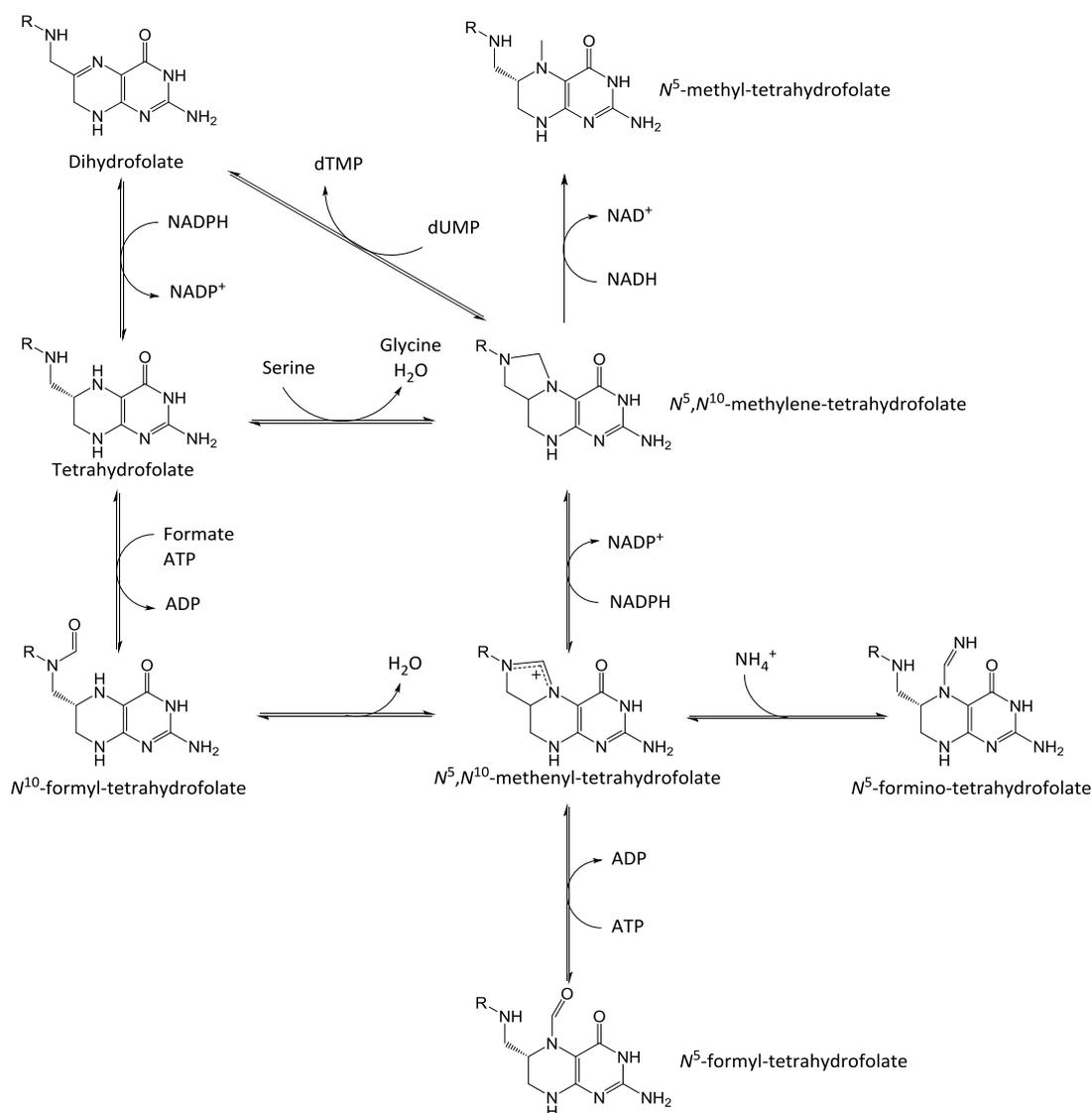
Scheme 1.2 Formation of one-carbon carrier units from tetrahydrofolate

As stated, DHFR is used as a model system for studying the relationship between dynamics, catalysis and hydrogen tunnelling. *Escherichia coli* DHFR (EcDHFR) is one of the most studied DHFRs, but more recently thermophilic DHFRs from *Thermotoga maritima* (TmDHFR) and *Geobacillus stearothermophilus* (BsDHFR)

have also been used to investigate enzyme dynamics in this family of enzymes (see below).

1.2.1 Escherichia coli DHFR (EcDHFR)

1.2.1.1 Structure of EcDHFR

EcDHFR is a monomeric enzyme with a mass of 18000, consisting of 159 amino acids that make up four α-helices (αB, αC, αE and αF) and an eight stranded β-sheet (βA-βH) (Figure 1.6) (65-69). The strands that make up the β-sheet are named sequentially: βA (residues 1-8), βB (39-43), βC (89-95), βD (73-75), βE (89-95), βF (108-116), βG (132-141) and βH (150-159) (65). The β-strands are all parallel with the exception of the anti-parallel βH at the C-terminal of the protein. The α-helices are named by their proximity to the β-sheets with two helices each side of the β-sheet: αB (residues 24-35), αC (43-50), αE (77-86) and αF (96-104) (65).
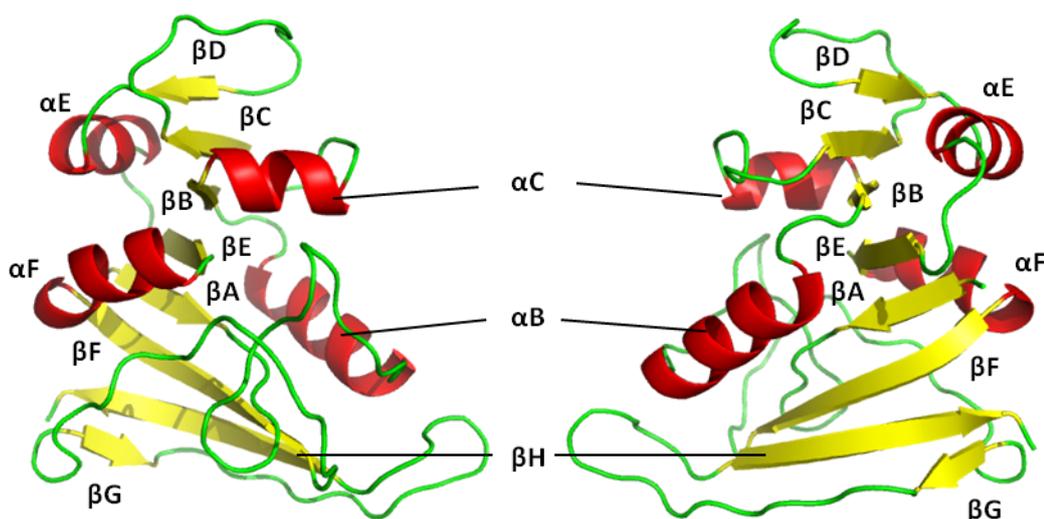
Figure 1.6 Cartoon representation of EcDHFR (1RX2 (69)) in two different orientations. The α-helices (red) and β-strands (yellow) are labelled.

The tertiary structure of EcDHFR was originally split into two main subdomains: the adenosine binding domain (ABD, residues 36-106) and the loop domain (LD, 1-35

and 107-159) (67, 69) (Figure 1.7). The ABD consists of β-strands B-E and α-helices C, E and F. This subdomain is rigid, and as the name suggests, binds the adenine part of the cofactor NADPH. The remaining part of the tertiary structure is deemed the LD due to 45% being made of three important loops: M20 (residues 10-24), FG (116-132) and GH (141-150) (Figure 1.7). The gap between the two subdomains, caused by the hydrogen bonding pattern between αA and αE, allows the nicotinamide ring to span this gap coplanar to the β-sheet. The pteridine ring of DHF occupies the cleft between αB and αC, and the reaction centre, *i.e.* C4 of NADPH and C6 of DHF, are held within sub van der Waals contact. This cleft forms the active site of EcDHFR and is shielded by the M20 loop. The binding of the nicotinamide-ribose moiety of NADPH is stabilised by the presence of the pteridine ring of substrate and it was found that only 75% of the nicotinamide-ribose moiety bound in the absence of substrate (69). More recently, the substrate binding domain (SBD) has been used to describe the areas that bind DHF, primarily consisting of αB (70, 71).
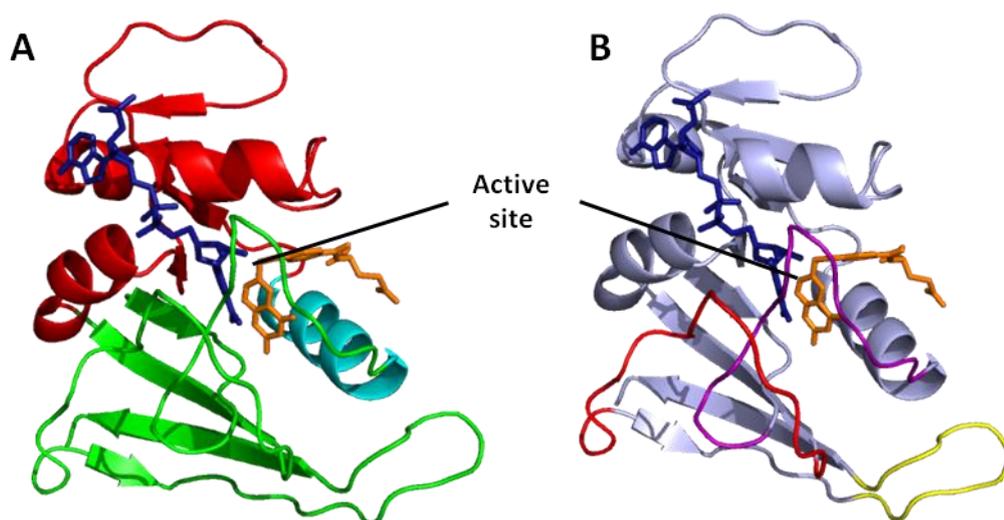
Figure 1.7 Cartoon representation of EcDHFR (1RX2 (69)) showing A) the three subdomains: ABD (red), LD (green) and SBD (cyan) and B) the M20, FG and GH loops (purple, red and yellow respectively). NADP^{+} and folate are shown in blue and orange respectively.

The substrate binds to the active site through hydrogen bonds to Asp27, Trp22, Thr113 and two water molecules (Figure 1.8) (68, 69, 72). Asp27 forms two hydrogen bonds with the C2-amino group and the N3 of the pterin ring. Thr113 also forms a hydrogen bond via one of the water molecules to the C2-amino group and Trp22 forms a hydrogen bond through the second water molecule to the carbonyl group of the pterin ring. This hydrogen bonding network between the three residues has been shown through site directed mutagenesis (SDM) studies to contribute to EcDHFR activity, and a decrease in catalytic activity is observed if one of the hydrogen bonds cannot form (73-75). The substrate was found to have hydrophobic interactions along the molecule between a number of residues in the SBD, in particular the conserved Leu28, Phe31, Ile50 and Leu54 residues. A hydrogen bond between Arg57 and the glutamate moiety is also present (68). A number of active site residues have been identified with regards to cofactor binding. The cofactor was found to hydrogen bond with Ser49 via a water molecule, as well as having hydrophobic interactions with Ile14 and Tyr100. Once again, SDM studies showed reduced reaction rates when these residues were substituted (76, 77).
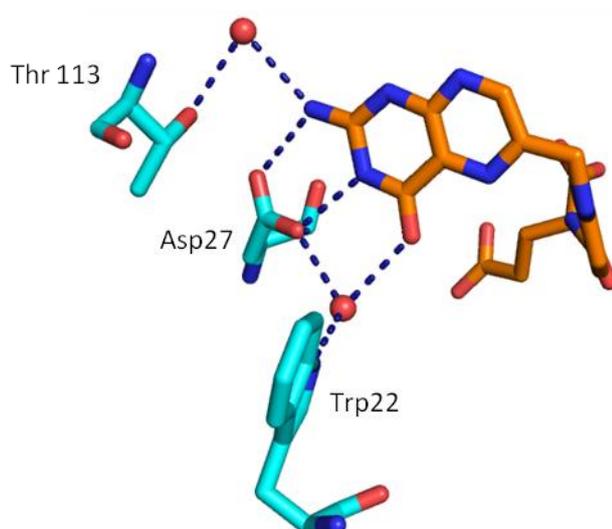
Figure 1.8 Active site of EcDHFR (1RX2 (69)) showing hydrogen bonds (dashed lines) between folate (orange) and the three key residues (cyan). The two water molecules are shown as red spheres.

1.2.1.2 Protonation of DHF

The protonation step is thought to be the route of the pH dependence of the EcDHFR reaction (with an approximate pK_a of 6.5). Asp27 is the only ionisable residue in the active site and, as it is highly conserved between other DHFRs (78), was predicted to be the proton source. However, the large distance between this residue and N5 atom of DHF makes direct protonation impossible. A proposed keto-enol tautomerisation of the substrate following binding (Figure 1.9) was suggested to stabilise a water molecule in the active site, hence allowing protonation of the N5 position (77, 79). However, it was shown by Shrimpton and Allemann that the bond angles and position of the water molecules were not conducive towards this proposed mechanism (80).

Significant evidence now suggests that the proton comes directly from the solvent. Crystal structures have shown a water molecule coordinated between Asp27, Trp22 and the carbonyl at C4 position in the pterin ring and this has been designated site A in the literature (Figure 1.9). Despite this position being too far away from the protonation site at N5, Shrimpton and Allemann calculated that only minor changes were needed in the active site conformation to move water molecules within hydrogen bonding distance of the reaction centre (site B, Figure 1.9 (80). The results from the modelling also showed an increase in the likelihood of water being found in site B as the reaction time progressed to 1 ns (80). Recent data suggests the first evidence of a bound water molecule at the N5 position has been witnessed (S. Benkovic, private communication). Additionally, an isotopic Raman shift was previously seen from the N5=C6 stretch in the Michaelis complex with D_2O as the solvent (81). The presence of a water molecule at site B was also seen by NMR spectroscopy, although in this case it was in human DHFR (82). Modelling has shown that as the reaction progresses the active site becomes more accessible to water, implying that after hydride transfer the conformation of the active site changes to allow protonation via the solvent (72, 80, 83).
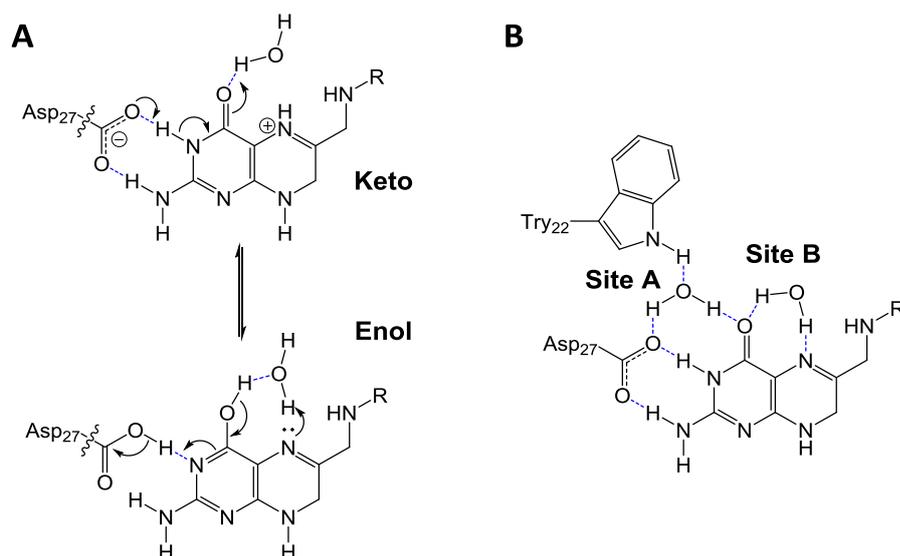
Figure 1.9 A) Proposed enol-keto tautomerisation mechanism of N5 protonation. B) Position of the water molecules in site A and site B of the active site of EcDHFR. Hydrogen bonds in both diagrams are shown by dashed lines.

1.2.1.3 Catalytic cycle of EcDHFR

Early kinetic studies on EcDHFR determined the pH dependence of the catalytic cycle and showed that hydride transfer was not rate limiting at physiological pH (75, 84, 85). However, the complete kinetic cycle of EcDHFR was determined by Benkovic and coworkers in 1987 (86). It was found that EcDHFR cycles through five key intermediates during the kinetic cycle: the holoenzyme (E:NADPH), the Michaelis complex (E:NADPH:DHF), the product ternary complex (E:NADP+:THF), the product binary complex (E:THF) and the product release complex (E:NADPH:THF). Rate constants and dissociation constants were determined for all steps in the cycle (Figure 1.10). Proceeding from the holoenzyme, the first step in the cycle is binding of DHF to form the Michaelis complex. The chemical step from the Michaelis complex to the product ternary complex is pH dependent, as stated earlier, with a pK_a of 6.5 and the maximum rate of 950 s^{-1} was determined at pH 5. At pH 7, the rate of hydride transfer drops to 220 s^{-1} but is still 18 fold faster than the rate limiting step. Following the chemical step the oxidised cofactor is released,

NADPH rebound, and the final step is product release to reform the holoenzyme. The rate limiting step at pH 7 is this release of THF to reform the holoenzyme complex (12 s^{-1}). At high pH (> 9) the chemical step becomes rate limiting due to the lack of protons available to protonate at the N_5 position of DHF. Free EcDHFR is not formed at any stage of the preferred cycle (86). KIE studies have also been performed at pH 7 by Allemann and coworkers using transient state kinetics, and at pH 9 using steady state kinetics by Kohen and coworkers (87, 88). Although the measurements at pH 9 allow the intrinsic KIE to be determined, the pH 7 measurements provide physiologically relevant data. Recently, an investigation has claimed to have measured the intrinsic k_H/k_T value at pH 7 (89). However, very little description of how this was performed is given, and it is unclear how the inherent problems of performing this measurement at pH 7 were overcome, and in fact the authors rely on pre-steady state measurements for analysis of results in the same study (89). At high pH the KIE is temperature independent with a A_H/A_D value of 4.0 ± 1.5 which is above the semi-classical limit (88). At pH 7 the KIE becomes temperature dependent with an inverse A_H/A_D value (0.108 ± 0.04) (87). Both the intrinsic KIE at pH 9 and the observed KIE at pH 7 are unable to be explained by semi-classical theory and infer that hydrogen tunnelling is involved in the EcDHFR reaction.

Image removed due to copyright reasons.

See legend for reference.

Figure 1.10 Kinetic scheme of EcDHFR at 25 °C, pH 7. E = enzyme, NH = NADPH, N = NADP$^+$, DHF = dihydrofolate and THF = tetrahydrofolate. Adapted from Fierke, C. A., Kohnson, K. A. and Benkovic, S. J., (1987), Construction and evaluation of the kinetic scheme associated with dihydrofolate reductase from *Escherichia coli*, J. Am. Chem. Soc., 26, 4085-4092.

Following on from the kinetic characterisation by Benkovic and coworkers, the structures of the five key intermediate complexes were determined by Sawaya and Kraut (69). Crystal structures were obtained of the holoenzyme (E:NADPH), Michaelis complex mimic (E:NADP$^+$:Folate), product ternary complex (E:ATP-ribose:dideazaTHF), product binary complex (E:ddTHF) and product release complex (E:NADPH:ddTHF). An E:NADPH:methotrexate complex was also obtained, proposed to be a transition state analogue. The M20 loop, which shields the active site in the Michaelis complex, was seen to adopt three different conformations: closed, occluded and open. The closed and occluded conformations are adopted during the main catalytic cycle of EcDHFR (Figure 1.11). The findings by Sawaya and Kraut have since been verified via NMR spectroscopy, with conformational differences witnessed in solution by Wright and coworkers, inferring the presence of the open and occluded conformations (70, 90-92).

Figure removed due to copyright reasons.

See legend for reference

Figure 1.11 Main catalytic cycle of EcDHFR showing the closed (C) and occluded (O) conformations adopted at each stage. E = enzyme, NH = NADPH, N = NADP$^+$, DHF = dihydrofolate and THF = tetrahydrofolate. Figure adapted from Venkitakrishnan, R.P., Zaborowski, E., McElheny, D., Benkovic, S.J., Dyson, H.J., and Wright, P.E., (2004), Conformational changes in the active site loops of dihydrofolate reductase during the catalytic cycle, Biochemistry, 43, 16046-16055.

The closed conformation is adopted in both the holoenzyme and the Michaelis complex. In this state the M20 shields the active site by forming a short anti-parallel sheet (residues 16-19) and hairpin turn. This effectively seals the active site by placing Asn18, part of the hairpin, in a position to form a hydrogen bond with a histidine residue in αC (His45). The M20 loop also forms a hydrogen bond with the amide group of the nicotinamide ring through Ile14. This conformation is stabilised through a pair of hydrogen bonds that form between Asp122 of the FG loop and Gly15 and Glu17 on the M20 loop (Figure 1.12) (69). The remaining three complexes (with the product bound) all form the occluded conformation. In this conformation the M20 loop "occludes" the active site and prevents the coenzyme binding in the active site. This is caused by a 180° rotation around the ψ bond of Ile14, resulting in the N-terminal end of the loop blocking the nicotinamide binding pocket. The central part of the M20 loop forms a helix in this conformation, instead of the sheet and hairpin seen in the closed conformation. The rotation at Ile14 also disrupts the hydrogen bonds with FG that stabilised the closed conformation. As a result the occluded conformation is stabilised by a pair of hydrogen bonds between Asn23 in the M20 loop and Ser148 in the GH loop (Figure 1.12) (69). The open conformation shows a mixture of characteristics between the closed and occluded forms (69).
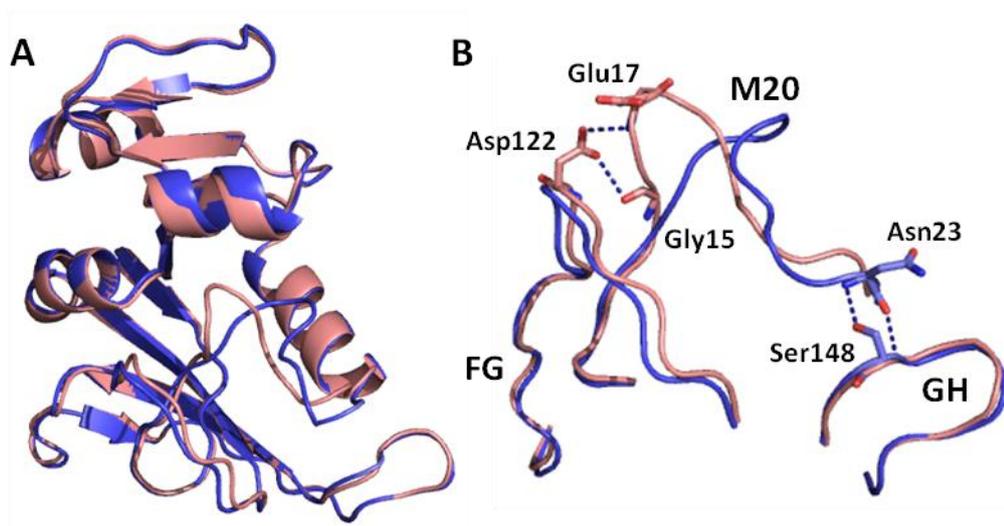
Figure 1.12 A) Overlay of the closed and occluded structures of EcDHFR (pink, 1RX2 and blue, 1RX6 respectively (69)). B) The M20, FG and GH loops in the closed and occluded conformations (red and blue). The hydrogen bonds that stabilise each structure are shown (dashed lines).

1.2.2 EcDHFR catalysis and the role of dynamics

It is generally accepted that the physical steps of catalysis in EcDHFR are controlled by the large loop movements and conformational changes described above. However, this has led to numerous studies and proposals on the coupling of these motions to the hydride transfer step in EcDHFR. This proposed coupling has been debated by a number of groups both experimentally and theoretically, and will be discussed here.

1.2.2.1 Dynamic nature of EcDHFR

Early evidence that suggested the coupling of motions to catalysis in EcDHFR was the effect that mutations distal to the active site had on hydride transfer rate. Model-free ^{15}N relaxation measurements were performed on the binary product complex (E:Fol) by Wright and Benkovic (93). They found a number of residues in the FG loop, M20 loop and ABD that showed dynamic fluctuations on the ps-ns timescale. It was predicted that these residues could potentially be coupled to the

conformational changes or catalysis of EcDHFR (93). One of the residues identified was Gly121 positioned in the FG loop over 12 Å from the reaction centre (C4 of NADPH). Benkovic and coworkers found that the G121V mutant exhibited a drop in hydride transfer rate by around 160 fold (94). Other distal mutants have also been investigated that affect the early stages of the catalytic cycle by a factor of 6 or more, including mutations to Met42, Arg44, His45, Thr46, Leu54, Asp122 and Phe125 (95-101).

Radkiewicz and Brooks performed molecular modelling on the Michaelis complex (E:NADPH:DHF), the product ternary complex (E:NADP$^+$:THF) and the product release complex (E:NADPH:THF) (102). Simulations were performed over 10 ns for each complex in order to capture slow motions, and it was found that there were strong correlated and anti-correlated motions in the Michaelis complex of EcDHFR (Figure 1.13). A high level of correlation was seen between the FG, M20 and αC regions in the Michaelis complex (yellow regions in Figure 1.13). These motions were not present in the product complexes leading to the presumption that these motions are necessary/involved in the chemical step of EcDHFR. It was also seen that many of the mutations that were reported to affect the catalytic rate show anti-correlated motions with other parts of the enzyme (102). Following on from this investigation, Wright and Benkovic repeated model-free ^{15}N relaxation NMR studies, this time on the Michaelis complex (E:NADP+:Fol), the product ternary complex/product release complex (E:DHNADPH:Fol) and the product binary complex (E:Fol) (90). The NMR investigation also found a difference in the dynamics seen in the Michaelis complex and the two product complexes that supported the findings from the modelling. The product complexes showed large motions in the M20, FG loop and ABD (as seen in the previous investigation on the E:Fol complex only). The Michaelis complex showed perturbation in the M20 and FG loop movements but increased flexibility in the ABD. These findings supported the theory that long distance coupling is involved in EcDHFR catalysis (90).

Image removed due to copyright reasons.

See legend for reference.

Figure 1.13 Residue-residue map of correlated and anti-correlated motions (red and blue respectively). Letters on the top and bottom edge show residues that have been shown to affect catalysis when mutated. Reprinted with permission from Radkiewicz, J.L. and Brooks, C.L., (2000), Protein dynamics in enzymatic catalysis: Exploration of dihydrofolate reductase. J. Am. Chem. Soc., 122, 225-231.

Benkovic and Hammes-Schiffer provided a more complete picture of this dynamic network proposed in EcDHFR (103). Genomic analysis highlighted conserved residues that are found in different DHFRs, and kinetic analysis of mutants of these residues showed detrimental effect on catalysis. In addition, mixed quantum/classical molecular dynamics were performed on the hydride transfer step of the reaction. It was suggested that the DAD decreased during the course of the reaction from 3.4 Å in the reactants to 2.7 Å, and the authors insinuated that a number of distal and proximal residues contributed to this change in the reaction coordinate as a network of residues *e.g.* Asp122 and Gly15 (Figure 1.14). Finally, analysis of specific mutants suggested that vibrations occurring over a time scale of femtoseconds to picoseconds were also present during the hydride transfer reaction. It was found that this network of proposed coupled promoting motions was present through to the transition state. However, significantly, it was noted that the study was unable to differentiate motions that play an active role in catalysis and those that occur as a result of it (103).

Image removed due to copyright reasons.

See legend for reference.

Figure 1.14 Network of coupled promoting motions that occur during the hydride transfer reaction in EcDHFR. Figure taken from Agarwal, P.K., Billeter, S.R., Rajagopalan, P.T.R., Benkovic, S.J., and Hammes-Schiffer, S., (2002), Network of coupled promoting motions in enzyme catalysis. Proc. Nat. Acad. Sci. U.S.A., 99, 2794-2799.

These findings all suggested that the EcDHFR catalytic cycle involves a large contribution from enzyme dynamics. However, none of the experimental or theoretical studies showed direct evidence of the dynamics being intrinsically coupled to the hydride transfer step. It is perhaps worth noting that the high level of structural interlinking in EcDHFR, through the many hydrogen bonds between loops and with the substrate, would naturally result in correlated motions whether the distal region is involved in catalysis or not.

1.2.2.2 Role of dynamics in EcDHFR

Whilst the dynamic nature of EcDHFR is clear, and not unexpected from the large number of interactions within the protein itself and with the cofactor/substrate, the potential coupling of slow, large-scale motions to the chemical step, itself several orders of magnitude faster, has been the subject of much discussion. At pH 9 the primary KIE was found to be temperature independent, have a larger than zero ΔE_a and A_H/A_D ratio greater than unity (88). A temperature independent KIE would be expected when the main contribution is from tunnelling and does not require thermal activation, hence, an ΔE_a greater than zero in EcDHFR requires further explanation. It was proposed that a Marcus-like model with a temperature

dependent, isotope independent rearrangement and an isotope dependent, temperature independent tunnelling step explained the results obtained at pH 9 (88).

However, pH dependence measurements of the hydride transfer showed that the temperature independent regime of the KIE appears between pH 7.5 and 8, with a temperature dependent regime below this (104). This implied that it was most likely a conformational change or electrostatic effect that caused the temperature dependence in EcDHFR, ruling out a role for environmentally coupled tunnelling. Furthermore, the effect of numerous solvents on hydride transfer was measured with the hypothesis that any long range coupling should be affected by the viscosity of the solvent. It was found that whilst electrostatic properties of the solvents had an effect on the hydride transfer step there was no correlation with the viscosity of the solvent (105). Whilst these findings provide evidence against long-range coupling, further claims were published, again relying on assumptions that coupling was the reason behind observed results.

One EcDHFR mutant in particular, N23PP/S148A, has been at the centre of numerous statements surrounding the existence of large-scale motions coupled to EcDHFR catalysis (106). This mutant is catalytically compromised, with a much reduced hydride transfer rate supposedly due to a loss of ms motions between the WT and mutant enzyme (106). However, investigations have highlighted the dangers of comparing results from WT enzymes and their mutants without taking into account the specific effect that the mutation has on the potential energy surface (107-109). Atomistic simulations (51) and stopped-flow analysis (110) have shown that a different potential energy surface and subsequent preorganisation of the active site does indeed account for the observed reduction in rate, thus ruling out a role for long-range coupled motions. Perhaps most tellingly, the original authors behind the assumption that coupling exists in EcDHFR-N23PP/S148A have both reversed their opinions. Benkovic states that although ms motions do exist in EcDHFR, a fact that was widely accepted previously, coupled motions do not (111).

Similarly, Wright claims increased conformational sampling of non-productive conformations is the inherent cause behind the slower rates in EcDHFR-N23PP/S148A, rather than the loss of assumed coupled motions (112).

The evidence that long-range motions do not play a role in the chemical step of EcDHFR significantly outweighs the somewhat circumstantial evidence to suggest coupling ever existed. Instead, the ability of the enzyme to undergo conformational sampling and preorganisation of the reaction coordinate are suggested as the main factors for the observed effects in EcDHFR catalysis (51, 110, 113). However, the role of short-range, vibrational motions on the fs timescale, coupled to the chemical step, cannot be ruled out in the experimental investigations above. Recently, isotopically labelled enzymes have been employed as tools to study this problem. Electrons have a mass roughly 1/2000 of that of a proton, the lightest component of the nucleus. This huge mass difference led to the Born-Oppenheimer approximation, which states that at any given point in time nuclei are stationary with regards to the much lighter, fast-moving electrons. Therefore isotopes of the same element have identical potential-energy surfaces (28). This means that any differences in reaction rates arising from isotopic substitution are due to mass-dependent motions rather than any electrostatic properties being altered. These mass-dependent motions can refer to large conformational changes on a millisecond time scale as well as the ultrafast, short-range bond vibrations that occur on the fs time scale.

"Heavy enzymes", labelled with the stable isotopes ^{13}C, ^{15}N and ^{2}H, have been used to investigate HIV protease, purine nucleoside phosphorylase (PNP), alanine racemase (deuterated only) and pentaerythritol tetranitrate reductase (PETNR) (19, 114-116). These studies all found a measurable effect on reaction rates upon isotopic labelling, implying that fs vibrations play a role in catalysis. Additionally, two investigations on isotopically labelled EcDHFR and the isotopically labelled EcDHFR-N23PP/S148A have determined the specific effect that these ultrafast, mass-dependent motions have on catalysis (117, 118). The fs motions were shown to couple to the reaction coordinate, as seen in other heavy enzyme studies, but additional calculations enabled the specific role that these vibrations play on barrier

crossing, in particular the role that they play in both tunnelling and recrossing, to be elucidated (117). In EcDHFR, mass-dependent vibrations had no effect on barrier height or tunnelling probability during the reaction, instead reporting on the frequency of re-crossing events during the reaction (117). This showed that the mass-dependent vibrations do not directly modulate the DAD, but that the protein environment as a whole couples to the reaction coordinate. The heavy EcDHFR mutant investigation found direct coupling of these fs vibrations to the reaction coordinate actually decreases the catalytic efficiency (118), further confirming early conclusions that a reduction in coupled motions was not the cause of the compromised activity. Possibly one of the strongest arguments against the role of enzyme dynamics being coupled to hydride transfer/hydrogen tunnelling in EcDHFR is the fact that at physiological pH the chemical step is approximately 20 fold faster than the rate limiting step. Without selective pressure, therefore, it seems unlikely that the enzyme structure would be optimised with regard to dynamic motions in this way, as a more efficient chemical step offers no evolutionary advantage.

Overall it can be seen that the role of enzyme dynamics in EcDHFR (and enzyme catalysis in general) is still an area of scientific interest. Whilst both long-range and short-range dynamics that influence the catalysis in EcDHFR have seemingly been ruled out, there are still disagreements in the literature. As well as this, although EcDHFR has been used as a paradigm for studying enzyme dynamics, the study of other DHFR homologues might enable a general theory to be postulated.

1.2.3 Extremophilic DHFRs

In the last few years extremophilic DHFRs have also started to be used as model systems to investigate enzyme dynamics. Whilst the information on these DHFRs is limited compared to the volume of information surrounding EcDHFR, they are still potentially useful models.

1.2.3.1   *Thermotoga maritima DHFR (TmDHFR)*

Thermotoga maritima is a hyperthermophilic bacterium that inhabits hot springs and hydrothermal vents. The optimum environmental temperature for this bacterium is 80 °C (119). TmDHFR is therefore a hyperthermophilic enzyme with a melting point of 82 °C, considerably higher than other DHFRs (120). TmDHFR is also unique in the fact that it is the only chromosomally encoded DHFR that forms a dimer quaternary structure (Figure 1.15). The dimer consists of two monomer units that are very similar in tertiary structure to monomeric DHFRs (121). Each monomer consists of the same four α-helices and eight β-strands as seen in EcDHFR but also contains a ninth anti-parallel β-strand. The monomer units dimerise through a hydrophobic core formed through the extra β-strand from each monomer, as well as two salt bridges between Lys129 from one monomer and Glu128 and Glu136 from the other. This hydrophobic core and salt bridges are believed to be the route of the intrinsic thermostability seen in this protein (121-123). The dimerisation region also includes the M20, FG and GH loops, preventing the significant movements that are observed in EcDHFR. As a consequence, the M20 loop of TmDHFR is fixed in an open conformation resulting in the active site being accessible to the solvent. This causes a different electrostatic structure of the active site compared to monomeric DHFRs (124).
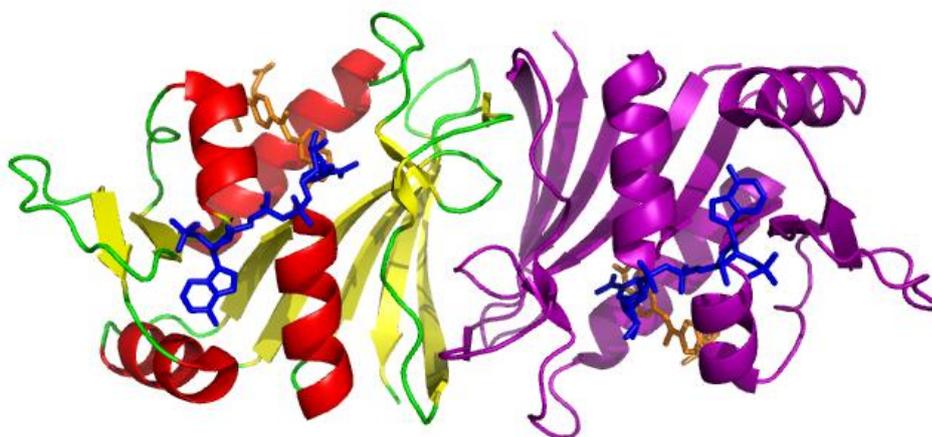
Figure 1.15 Cartoon representation of TmDHFR (1D1G (121)) showing the two monomer units that make up the dimer structure bound to NADPH and MTX (blue and orange respectively).

The catalytic turnover of TmDHFR is much lower than that seen in EcDHFR with k_H values 1000 fold lower (0.14 s^{-1} at 25 °C) and steady state turnover 50 fold slower (0.2 s^{-1} at 25 °C) at pH 7 (120). Unlike in EcDHFR, the hydride transfer is partially rate limiting at all temperatures at pH 7 in TmDHFR. KIE measurements on hydride transfer showed a breakpoint in the primary KIE at 25 °C. Above 25 °C the KIE is temperature independent with A_H/A_D value of 1.56. Below 25 °C the KIE becomes strongly temperature dependent and the A_H/A_D becomes << 1 (0.002). A computational study showed that the reaction in TmDHFR has a significant contribution from hydrogen tunnelling (125). Subsequent work involving numerous co-solvents found that the viscosity of the solvent had no effect on the hydride transfer, ruling out long range motions (or the lack of) as a reason for the breakpoint seen (126). It was suggested that the breakpoint was due to the enzyme being unable to adopt a conformation conducive to hydride transfer at low temperature. An investigation on the V11D mutant, which is able to switch to a monomer in the presence of non-denaturing detergents, found that the k_{cat} rate decreased suggesting the catalytic loops of TmDHFR are unable to function in the same manner as those in EcDHFR (127). More recently a mutant hybrid of EcDHFR and TmDHFR found that whilst the physical steps of catalysis require the loops of EcDHFR to be flexible this is not the case in TmDHFR catalysis (128). This finding shows that the results seen in the dimeric TmDHFR are not always comparable to other monomeric DHFRs.

1.2.3.2 *Geobacillus stearothermophilus DHFR (BsDHFR)*

Geobacillus stearothermophilus (formerly *Bacillus stearothermophilus*) is a moderate thermophile that occupies a wide range of ecological niches (129, 130). The optimal temperature for this organism is 60 °C, although it is able to survive between 30 and 75 °C. BsDHFR (this term is used in previous literature) has a melting temperature of 67 °C, showing moderate thermophilicity (131). As with other monomeric DHFRs the tertiary structure is highly conserved with that seen in EcDHFR. Once again the structure is made up of an eight stranded β-sheet with four α-helices, two either side. It can be seen from the structural overlay with EcDHFR

that there is a difference in the αE region of the ABD, with a 20° tilt seen in BsDHFR compared to the mesophilic enzyme (132). Interestingly for a thermophilic enzyme it was found that BsDHFR is more flexible than EcDHFR over the majority of the enzyme. Neutron scattering and hydrogen/deuterium exchange mass spectroscopy showed that whilst the inner core of the protein is more rigid than EcDHFR, the remaining structure including the loop regions are more flexible in BsDHFR (133, 134). Unlike EcDHFR, which adopts the closed and occluded conformation during the catalytic cycle, there has been no evidence to date of the occluded conformation playing a role in BsDHFR catalysis (132).
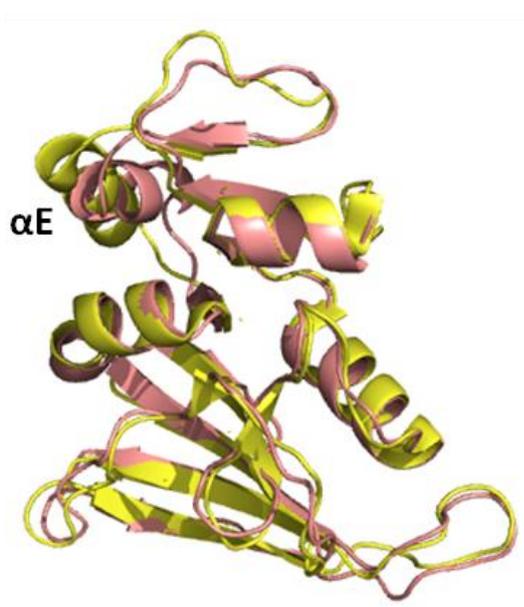
Figure 1.16 Cartoon representation of BsDHFR (yellow, 1ZDR (132)) overlaying EcDHFR (pink, 1RX2 (69)). αE is labelled highlighting the difference between the two enzymes.

The steady state characteristics of BsDHFR were found to be very similar to EcDHFR although the full kinetic cycle, and thus the rate-determining step at pH 7, has not been determined (132, 135). At pH 7, BsDHFR shows a slightly increased rate of turnover compared to the mesophilic enzyme, with a k_{cat} of 21.6 ± 1.6 at 25 °C (136). As with EcDHFR, the hydride transfer step becomes rate limiting at high pH, and KIE measurements at pH 9 showed a weakly temperature dependent KIE and

A_H/A_D value of 0.57 (132). These values are just outside the semi-classical region, therefore suggesting a role for hydride tunnelling. Unlike in EcDHFR, Klinman and coworkers argue that the M20, FG and GH loops are only involved in preorganisation in BsDHFR. Moreover, they argue that temperature dependent motions in the βA and αF regions of BsDHFR contribute to the reorganisation process and thus hydride transfer (131). However, previous work involving cosolvents showed similar findings to that of EcDHFR and TmDHFR (105, 137). It was found that BsDHFR was affected by the dielectric properties but not by the viscosity of the solvent (138), giving a strong indication against long-range coupling to the chemical step in this system.

1.2.3.3 *Moritella profunda DHFR (MpDHFR)*

Moritella profunda is a psychrophilic bacteria that is found in the oceans. At 0.1 MPa the optimal growth temperature is 2 °C, with no growth seen above 12 °C (139). MpDHFR has a lower melting temperature compared to EcDHFR at 37.5 °C, as would be expected for a psychrophilic enzyme (135). The tertiary structure of MpDHFR follows the same trend as the other DHFR enzymes with an eight stranded β-sheet and four α-helices (Figure 1.17) (140). Unlike many other DHFRs, the conserved Asp27 residue (EcDHFR numbering) has been replaced by a glutamate in the active site of MpDHFR (140). This difference has also been found in numerous trimethoprim resistant DHFRs (141, 142), and was thought to be the cause of the lower binding affinity observed in MpDHFR (143). However, investigations of the mutant MpDHFR-E28D have shown that this single mutation is not the cause of the decrease in the binding affinity of trimethoprim (144, 145). *M. profunda* is able to grow at high pressure in deep ocean environment, however, an investigation found that MpDHFR showed no obvious signs of pressure adaptation (140). The kinetic properties of MpDHFR are similar to other DHFRs, in that at pH 7 the rate limiting step is much slower than hydride transfer, and is thought to be product release (146). At pH 9, the hydride transfer becomes the rate limiting step. The k_{cat} at 25 °C, pH 7 is 18.7 ± 0.8, significantly faster than EcDHFR (12 s^{-1}) (135). Similarly the k_H is faster in MpDHFR, with a rate constant of 527 compared to 220 s^{-1} in EcDHFR (113).

KIE measurements at pH 7 found a very similar trend to EcDHFR, with a temperature dependent KIE. The magnitude of the observed KIE is slightly lower than seen in EcDHFR (1.9 compared to 2.7 at 25 °C) (113). MpDHFR has been found to be more flexible, as shown by hydrogen/deuterium exchange measurements using NMR spectroscopy. As with BsDHFR (see above) there is no evidence that the occluded conformation plays a part in the catalytic cycle of MpDHFR. Large-scale dynamics have been ruled out in MpDHFR through analysis of hydride transfer with cosolvents (113), although the role of short-range motions has not been investigated to date in this enzyme.
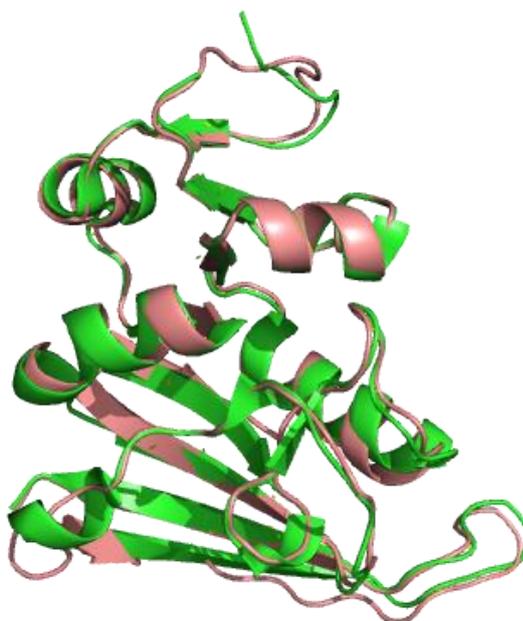
Figure 1.17 Cartoon representation of MpDHFR (green, 3IA4 (140)) overlaying EcDHFR (pink, 1RX2 (69)).

1.3 Secondary KIEs, heavy atom KIEs and TS analogues

Although the use of primary hydrogen KIEs has proved invaluable in the study of enzyme catalysis, many other isotope effects can be measured that elucidate key information regarding reaction mechanisms (147). Isotope labelling of hydrogen atoms that do not directly take part in the reaction allow secondary (2°) KIEs to be measured (28). If the isotope position is bonded to the reaction centre but does not take part in the reaction an α-2° KIE is measured. Any other hydrogen position that

is labelled results in a β-2° KIE (28). Isotope effects involving atoms other than hydrogen are termed heavy atom KIEs (147). Once again, these can be primary or secondary depending on atom position, and are normally restricted to ^{15}N, ^{18}O, ^{13}C and ^{14}C KIEs.
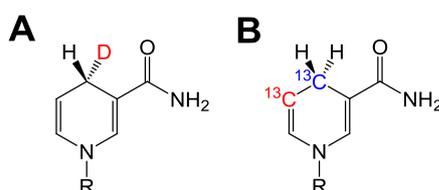
Figure 1.18 A) The α-secondary hydrogen position of NADPH labelled with deuterium (red) with regards to DHFR catalysis. B) The carbon-13 positions of NADPH to give a primary heavy carbon isotope effect (blue) and secondary carbon isotope effect (red) with regards to DHFR catalysis.

1.3.1   α-Secondary hydrogen isotope effects

α-secondary hydrogen (2°) KIEs arise due to rehybridisation of the carbon atom adjacent to the atom that has been substituted, *e.g.* from an sp^3 to sp^2 hybridised carbon (28). As a result of this, only certain bond vibrational modes alter significantly thus affecting the reaction rate. Force constants of stretching modes do alter as a bond undergoes rehybridisation. However, the difference is often too small to create observable isotope effects. Hence, the main contribution to observed 2° KIEs is from differences in bending vibrational modes (37).

In the case of an sp^3 to sp^2 rehybridisation, in-plane bending modes are almost identical and do not significantly contribute to the isotope effect. However, the different steric properties of the two hybridisation states results in a large disparity between out-of-plane bending vibrational modes. An sp^2 hybridised carbon produces very little steric hindrance towards out-of-plane bending of the hydrogen atom. Conversely, an sp^3 hybridised carbon is sterically hindered, resulting in out-

of-plane vibrational frequencies of approximately 1350 and 800 cm^{-1} for an sp^3 and sp^2 hybridised carbon respectively (37). Therefore, large differences in ZPE between the two bonds results in a 2° KIE (Figure 1.19).

As a result of this difference, the mechanism of the reaction can be inferred from the magnitude of the 2° KIE. An sp^3-sp^2 rehybridisation produces normal KIEs, as the light atom reacts faster than the heavier isotope (A, Figure 1.19). In contrast, an sp^2-sp^3 rehybridisation produces an inverse 2° KIE, as the greater force constant results in a larger difference between the light and heavy atom at the TS and a greater energy of activation for the lighter isotope (Figure 1.19) (28, 37). In the vast majority of enzymes the 2° KIE is between unity and the equilibrium isotope effect (EIE). The EIE is defined as K_H/K_D, and is dependent on the equilibrium distribution of isotopes between the reactant and product states, rather than originating from changes between the ground state and TS (148). As full rehybridisation has occurred between the reactant and product, theoretically 2° EIE ≥ 2° KIE providing that no coupling occurs between the 1° and 2° hydrogen (see below).
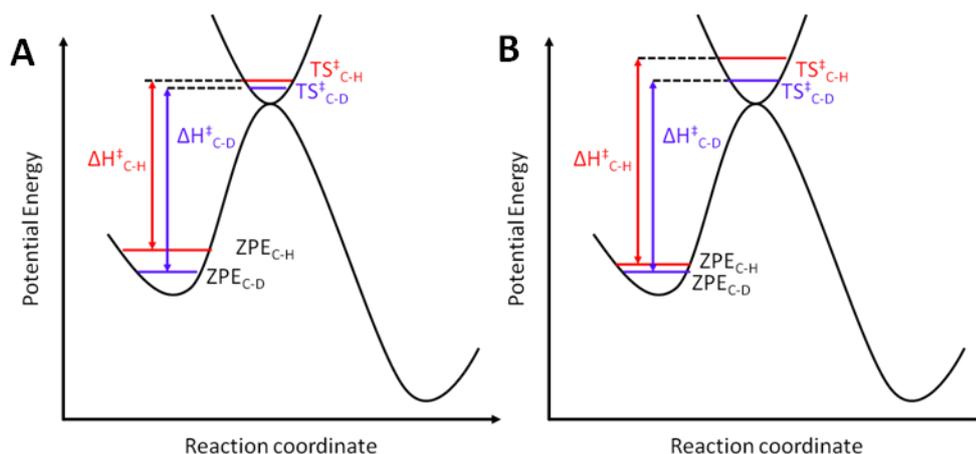
Figure 1.19 Reaction coordinates of A) an sp^3 to sp^2 rehybridisation reaction, and B) an sp^2 to sp^3 rehybridisation reaction. Figure adapted from ref (37).

Early studies focused on the magnitude of the 2° KIE as an indication of the structure of the TS (28), although this approach has been proven to be over-simplified (149). A reactant-like TS would theoretically have a magnitude close to

unity, and a more product-like TS similar to that of the EIE. The appearance of exalted 2° KIEs above the EIE (150, 151), led to the proposal by Huskey and Schowen that this indicated coupling of the 2° hydrogen motion to the 1° hydrogen being transferred, with a concomitant contribution from tunnelling (152).

Scrutton and coworkers measured 2° KIEs for morphinone reductase (MR) and pentaerythritol tetranitrate reductase (PETNR), two enzymes that include the reduction of FMN via NADH in their reaction mechanisms (153). MR exhibits a temperature dependent 1° KIE whilst PETNR is temperature independent (154). However, both enzymes produce temperature independent 2° KIEs with identical magnitude, leading to the suggestion that 2° KIE reports on the reaction ready configuration (RRC) (153, 155). Furthermore, temperature independent KIEs suggest that the arrangement of FMN-NADH does not alter significantly over the temperature range, whilst identical magnitude was said to imply that the arrangement was similar between the two enzymes (153).

1.3.2 Heavy atom isotope effects

A TS lifetime is on the same time scale as a single bond vibration (fs), and thus direct observation is inherently difficult (156). Isotope effects are therefore powerful analytical tools, as they arise from changes in the vibrational state between the reactant state and TS, hence allowing the reaction coordinate to be probed directly (157). The magnitude of the heavy atom KIE reports on the environment of the specific atom in the TS, with a KIE of 1 implying the atom's bonding environment has not changed. A KIE >1 signifies that the labelled atom has become less constrained in the TS thus enabling the lighter isotope to react faster. Conversely, a value <1 is caused by the labelled atom being found in a more constrained environment in the TS, causing a greater difference in ZPE level between the two isotopes at the TS compared to the GS (similar to the effect of rehybridisation observed above). As a result the heavier isotope reacts at a greater rate (158).

As discussed above, numerous tunnelling models suggested that heavy atom motions modulate the DAD, thus directly promoting the transfer of the lighter isotope during hydrogen transfer reactions (25, 38-42). Whilst a role for long-range, slower (ms-ns) motions and fs motions from the enzyme have seemingly been ruled out, there is still uncertainty surrounding the contribution from fast bond vibrations in the substrate. By substituting specific atoms at the reaction centre any ultrafast vibrations that may affect the reaction rate can be assessed, determining the role of these motions in enzymatic reactions. Furthermore, once a number of KIEs have been measured for a specific reaction the data can then be used to predict a TS structure. Current methods involve structure interpolation, resulting in a number of TS produced computationally (159, 160). KIEs are calculated for each theoretical TS and a structure is generated where all calculated KIEs coincide with the experimental data obtained. Calculation of a TS structure opens the potential to synthesise TS analogues, compounds that have shown to be incredibly effective inhibitors. This technique has been used extensively by Schramm and coworkers to investigate a number of enzyme TS, including: purine nucleotide phosphorylase (see below), HIV-1 protease (161), methylthioadenosine phosphorylase (162-164) and nicotinamide phosphoribosyltransferase (165).

In theory, heavy atom isotope effects can be used to probe all enzymatic reactions. However, there are a number of practical limitations surrounding these isotope effects that has restricted their use to date. The first problem is that the magnitude of the isotope effect is small, thus requiring complex measurements to determine the KIE. A typical heavy atom KIE is <1.1, with semi-classical values calculated as: $^{12}C/^{14}C = 1.09$, $^{12}C/^{13}C = 1.04$, $^{14}N/^{15}N = 1.04$ and $^{16}O/^{18}O = 1.07$ at 25 °C (28, 147). This often results in the use of complex techniques, involving remote radiolabelling (166), isotope ratio mass spectrometry (167), whole molecule mass spectroscopy (168) and NMR (169). Furthermore, isolation of the isotopically dependent step is inherently difficult in many enzyme systems, ruling them out as suitable targets for heavy atom KIE analysis. Finally, the synthesis of specific isotopically labelled substrates is often difficult, expensive and time consuming.

1.3.3 Purine nucleoside phosphorylase

One of the best examples of the use of heavy atom isotope effects to investigate the TS is the case of human and bovine purine nucleoside phosphorylase (hPNP and bPNP). PNP catalyses the reversible phosphorolysis of inosine, releasing hypoxanthine and a ribose-phosphate moiety (Scheme 1.3). Early studies focused on bPNP due to the similarity with its human counterpart (87% sequence identity) and its availability. Schramm and coworkers synthesised a number of inosine isotopologues that incorporated ^{3}H-, ^{14}C- and ^{15}N- at specific positions around the molecule from labelled glucose and adenine. Biosynthetic methods were utilised, thus allowing a single route to produce 7 different isotopologues (170-172).
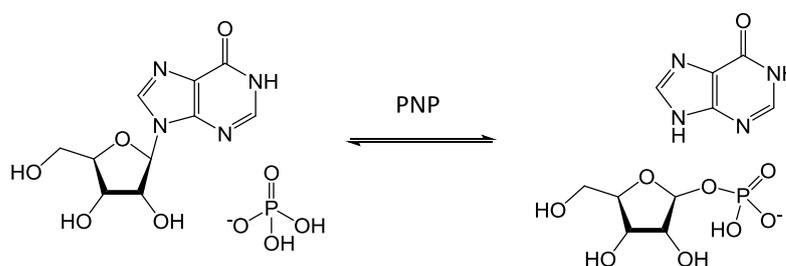
Scheme 1.3 Reaction catalysed by PNP.

Figure removed for copyright reasons.

See legend for reference.

Figure 1.20 The labelled isotopologues, intrinsic KIEs measured and calculated TS structures for bPNP and hPNP. The intrinsic KIE values are given as percentages. Figure taken from Schramm, V.L., (2011), Enzymatic transition states, transition-state analogs, dynamics, thermodynamics, and lifetimes. Annu. Rev. Biochem., 80, 703-732.

Competitive KIE measurements were used to determine intrinsic KIEs for each labelled atom. This data was subsequently used to calculate the TS structure, and applied to the design of TS analogues (Figure 1.20) (173). Binding studies on hPNP with TS analogues based on the TS of bPNP found that despite the high sequence identity between the two enzymes, the TS of hPNP differed significantly (174). Subsequent TS analysis of hPNP led to the development of specific TS analogues towards the human enzyme, with K_d values in the low pM, a 2 million fold increase in affinity than the natural substrate (175, 176).

1.4   DHFR inhibition and TS analysis

The crucial role DHFR plays in numerous metabolic pathways of the cell has led to this enzyme being a key drug target (177, 178). Many inhibitors of DHFR have been produced by the pharmaceutical industry to treat a wide range of illnesses (Figure 1.21). Methotrexate is one of the earliest anti-cancer agents developed, having been used in patients since the 1940's, and is used to treat cancers, such as leukaemia and breast cancer, as well as serious cases of rheumatoid arthritis (179, 180). Methotrexate has a very high binding affinity to DHFR, over a 1000-fold higher affinity to DHFR compared to folate, but is also unspecific in its binding to DHFRs from other species (181). Pemetrexed is also an anti-cancer agent, used to treat lung cancer (182). Trimethoprim and pyrimethamine are DHFR inhibitors that show specificity to non-human DHFRs. Trimethoprim is an antibacterial agent due to its higher affinity for bacterial DHFRs relative to human DHFRs, whilst pyrimethamine is an anti-protozoal agent used in the treatment of malaria (183-186).
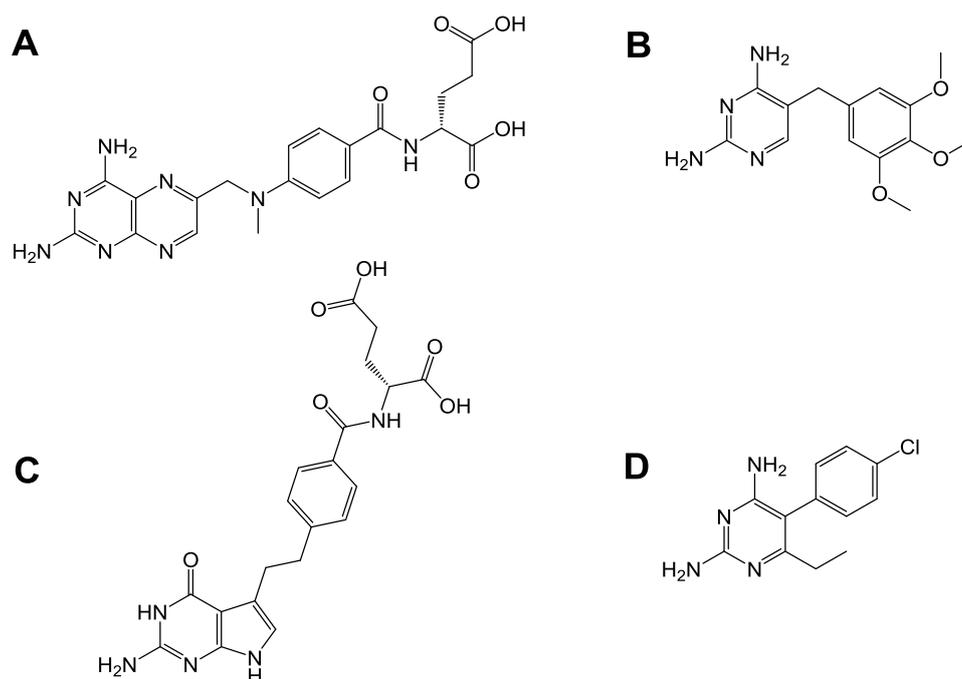
Figure 1.21 Examples of inhibitors of dihydrofolate reductase. A) Methotrexate, B) Trimethoprim, C) Pemetrexed and D) Pyrimethamine.

Despite their success, DHFR inhibitors have been developed through traditional means, *e.g.* fragment based design from structural data (187). TS analysis, through KIE measurements, could potentially lead to a new class of potent DHFR inhibitor. However, as with other enzymes, the ability to measure heavy atom isotope effects depends on the ability to synthesise numerous substrate and cofactor isotopologues, with isotope incorporation specifically around the reaction centre. To date, this approach has not been applied to this important family of enzymes. However, the biosynthetic pathways of NADPH and DHF provide an excellent starting point from which a single method can be developed to obtain the required isotopologues.

1.4.1.1   *NADP$^+$ Biosynthesis*

All organisms are able to synthesise NAD(P)$^+$ *de novo*, and although two different pathways exist in nature, they converge at the key intermediate quinolinic acid (QA). Eukaryotic organisms synthesise QA solely from the precursor tryptophan, through the action of 5 enzymes (Scheme 1.4). The first two enzymes, tryptophan-

2,3-dioxygenase and N-formylkynurenine formamidase, catalyse the cleavage of tryptophan to N-formylkynurenine and the subsequent removal of the formyl group to yield kynurenine (188, 189). Hydroxylation by kyneurine-3-monooxygenase followed by cleavage of the amino acid side chain by kynureninase produces 3-hydroxyanthranilate. The final enzymatic step catalysed by 3-hydroxyanthranilate-3,4-dioxygenase sees the oxidative rearrangement of the product to form QA (190-193). Whilst the enzymes involved in this pathway have been identified, they have proved difficult to study up to this point, thus relatively little is known about their mechanism of action.
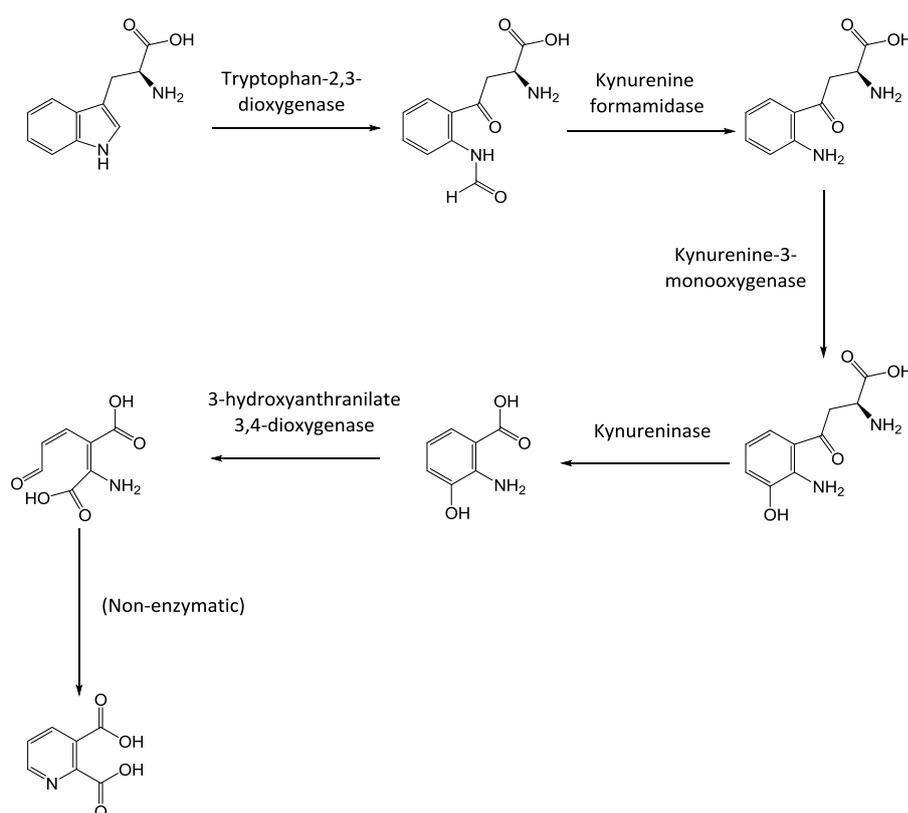
Scheme 1.4 Production of quinolinic acid from tryptophan in eukaryotic cells.

The second route to QA is from L-aspartate and dihydroxyacetone phosphate (DHAP), catalysed through the action of L-aspartate oxidase (LAO) and quinolinate synthase (QS) (Scheme 1.5). This was thought to be the only synthetic route used by prokaryotic organisms (194), however, incubation with labelled tryptophan has shown that a number of bacterial species are able to utilise both *de novo* routes to

QA (195). LAO is an FAD dependent enzyme that oxidises aspartate to iminoaspartate (196). This enzyme is able to utilise both fumarate and oxygen to recycle $FADH_2$ back to FAD, depending on the aerobic/anaerobic nature of the environment (197). It has also been shown that iminoaspartate can be generated chemically through Schiff's base formation, by ammonia and oxalacetate (198), therefore LAO is not essential in the development of the $NADP^+$ biosynthetic pathway.

QS catalyses the condensation of iminoaspartate and DHAP to form QA, resulting in the carbon adjacent to the phosphate in DHAP becoming C4 of QA and subsequently C4 of NAD(P)H (199). Despite numerous studies the reaction mechanism is still unknown, with studies hampered by the presence of an [4Fe-4S] cluster that is essential for catalysis (200, 201). Early investigations were unaware of the presence of this cluster (202), and it was not until *circa* 2005 that evidence of the labile [4Fe-4S] complex was confirmed (200). Sequence analysis has highlighted the fact that only three cysteine residues coordinate to the cluster, and this may influence the stability. A fourth site of coordination is thought to occur with a reaction intermediate (Figure 1.22) (200, 203), and potentially facilitates a dehydration reaction that is common to other Fe-S cluster containing enzymes (204).
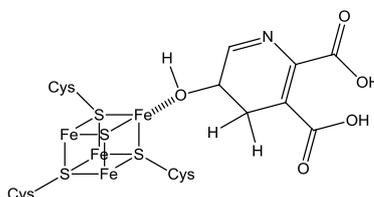
Figure 1.22 Coordination of quinolinic acid intermediate with the [4Fe-4S] cluster of QS.

$NADP^+$ is synthesised from the precursor quinolinic acid through the action of four enzymes (Scheme 1.5): quinolinate phosphoribosyltransferase (QAPRT), nicotinic acid mononucleotide adenylyltransferase (NAMNT), NAD synthetase and NAD kinase (194, 205, 206). QAPRT is classed among the type II phosphoribosyltransferases (PRTs) due to its tertiary structure that incorporates the

active site into a seven stranded α/β fold (207). This enzyme catalyses the addition of a phosphoribose moiety from phosphoribose pyrophosphate (PRPP) to QA, with subsequent decarboxylation, forming nicotinic acid mononucleotide (208). It has been proposed that the pyrophosphate group dissociates from the phosphoribose moiety and stabilises the subsequent oxocarbenium ion (194). Nucleophilic attack by the lone pair of the QA nitrogen results in the formation of quinolinic acid mononucleotide, with decarboxylation at the C2 position the final step of the reaction. However, no direct evidence exists to fully confirm the order of reaction. The decarboxylation step is proposed to occur after the addition of the phosphoribosyl group, due to similarities with the reaction catalysed by OMP decarboxylase (209). Calculations have shown that decarboxylation at the C2 position following formation of the mononucleotide is more energetically favourable, due to the presence of a stabilising positive charge on the adjacent nitrogen atom (210).
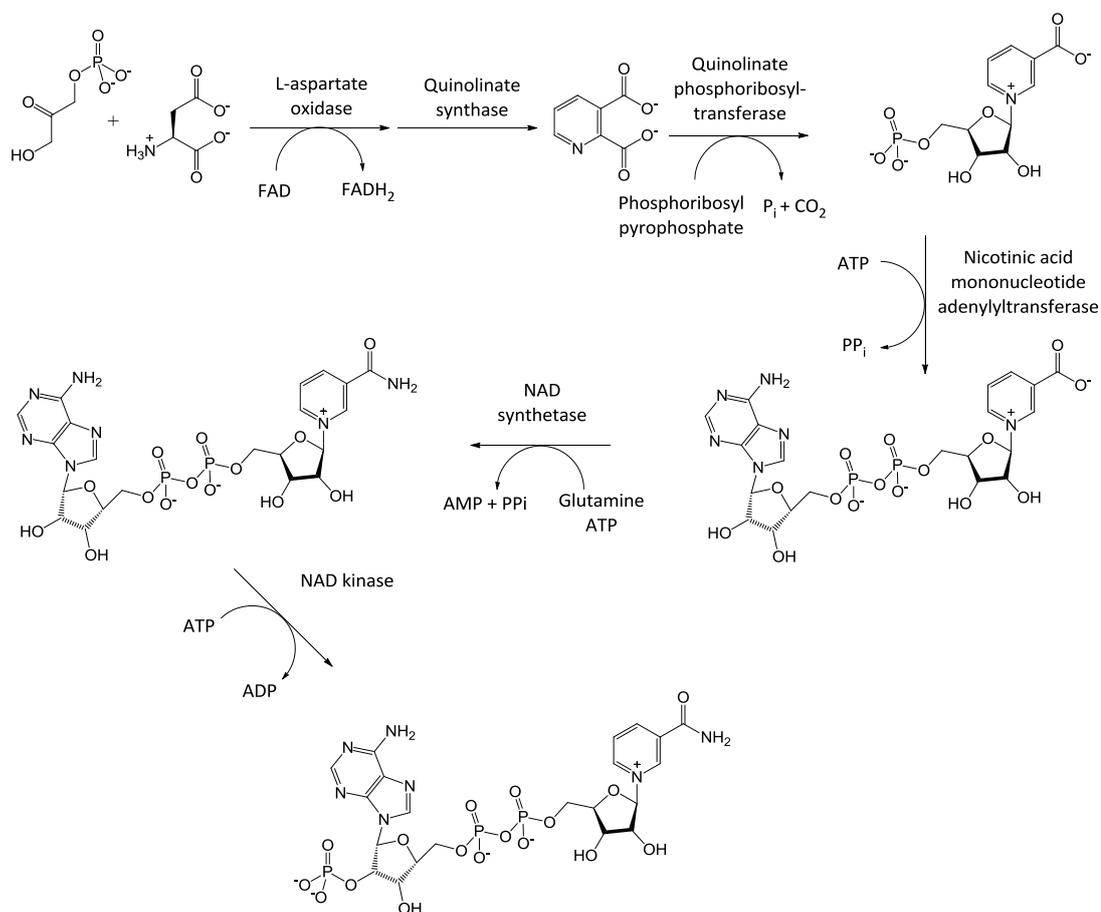
Scheme 1.5 NADP$^+$ biosynthesis in *E. coli.*

NAMNT catalyses the addition of adenosine to nicotinic acid mononucleotide to form nicotinic acid adenine dinucleotide (211). Whilst this enzyme is able to catalyse the addition of adenosine to nicotinamide mononucleotide, the conversion rate is compromised significantly (211). The third enzyme in the pathway, NAD synthetase, is responsible for the ATP dependent conversion of nicotinic acid adenine dinucleotide to nicotinamide adenine dinucleotide (NAD^+) (212). This family of enzymes is split into two distinct classes: glutamine dependent and ammonia dependent (213). In both classes the carboxylate group is first activated by adenylation followed by the nucleophilic attack of ammonia. The glutamine dependent enzymes contain a glutaminase domain/subunit, and following cleavage the ammonia molecule is channelled directly into the amination active site, without dissociation into solvent. As a result, all glutamine-dependent NAD synthetase enzymes (eukaryotic and some bacterial) contain two domains, whereas ammonia-dependent NAD synthetases (bacterial) have a single domain (213).

The final reaction in the pathway is the phosphorylation of NAD^+ to $NADP^+$ by NAD kinase (214). This is an important enzyme due to the distinct roles that NAD(H) and NADP(H) play in catabolism and anabolism respectively (215). NAD kinase catalyses the phosphorylation of the 2' position of the adenine-ribose group to afford $NADP^+$, thus is involved in maintaining cellular levels of each cofactor. As gene sequences, purification methods and activity assays have been reported for each individual enzyme (200, 201, 208, 211, 214, 216, 217), the *E. coli* de novo $NADP^+$ biosynthetic pathway seems an ideal target to reconstitute in the laboratory. This has the potential to enable incorporation of isotopic labels into the nicotinamide ring of $NADP^+$, a process that has never been performed via biosynthetic means.

1.4.1.2   DHF Biosynthesis

DHF and derivatives are essential cofactors for survival, however, vertebrates are unable to synthesise these compounds (218). Instead they are dependent on nutritional sources. Conversely, plants and bacteria are able to synthesise these compounds *de novo*, using a pathway that incorporates 6 steps to produce DHF from GTP (Scheme 1.6) (218-220).

The first committed step in DHF biosynthesis is the cyclisation of GTP to form dihydroneopterin triphosphate (221, 222). This reaction is catalysed by GTP cyclohydrolase I, a Zn^{2+} dependent enzyme, which hydrolyses the guanine base resulting in the loss of the C_8 atom in the form of formic acid. An Amadori rearrangement then occurs involving the ribose moiety to form the dihydroneopterin triphosphate product (223). This enzyme is a member of a family of four GTP cyclohydrolases: I, II, III and MptA. GTP cyclohydrolase II and III are involved in riboflavin and flavo-coenzymes (although the exact role of GTPCHIII is still unclear), whilst MptA is an archaeal enzyme that converts GTP to dihydroneopterin-2,3-cyclic phosphate, an early intermediate in THF biosynthesis in this class of organisms (224).

The mechanism for the subsequent removal of phosphate to yield dihydroneopterin has been an area of discussion for a number of years. Early reports suggested that a non-specific phosphatase and pyrophosphatase were responsible (225), although in recent years the appearance of enzymes that show specific activity towards dihydroneopterin triphosphate have altered this theory (226). In previous studies alkaline phosphatase has successfully been employed to remove the phosphate moieties, removing the need for a specific phosphatase to be produced and purified (220). Dihydroneopterin is cleaved to hydroxymethyl-dihydroneopterin by the action of dihydroneopterin aldolase (227). This enzyme performs a retro-aldol reaction, resulting in the release of glycoaldehyde (220).

The three remaining steps incorporate the addition of the p-aminobenzoic acid and glutamate portions of DHF. Firstly, activation of the hydroxymethyl-dihydroneopterin occurs through the addition of a pyrophosphate moiety catalysed by hydroxymethyl-dihydroneopterin pyrophosphokinase (228, 229). Subsequently, dihydroneopteroate synthetase (DHPS) catalyses the addition of p-aminobenzoic acid (228), a product of the shikimate pathway (230). DHPS is also a well known drug target, and often a combination of drugs that inhibit DHPS and DHFR are used in treatments (231). The final addition of glutamate by dihydrofolate synthetase (DHFS) produces the desired product DHF. In *E. coli*, DHFS also shows

folylpolyglutamate synthetase activity (232), and as such is capable of adding multiple glutamate molecules onto the product.

All these enzymes have been characterised in *E. coli* (227, 228, 232-235). This knowledge from previous studies provides an excellent framework for the development of biosynthetic methods to provide DHF isotopologues.
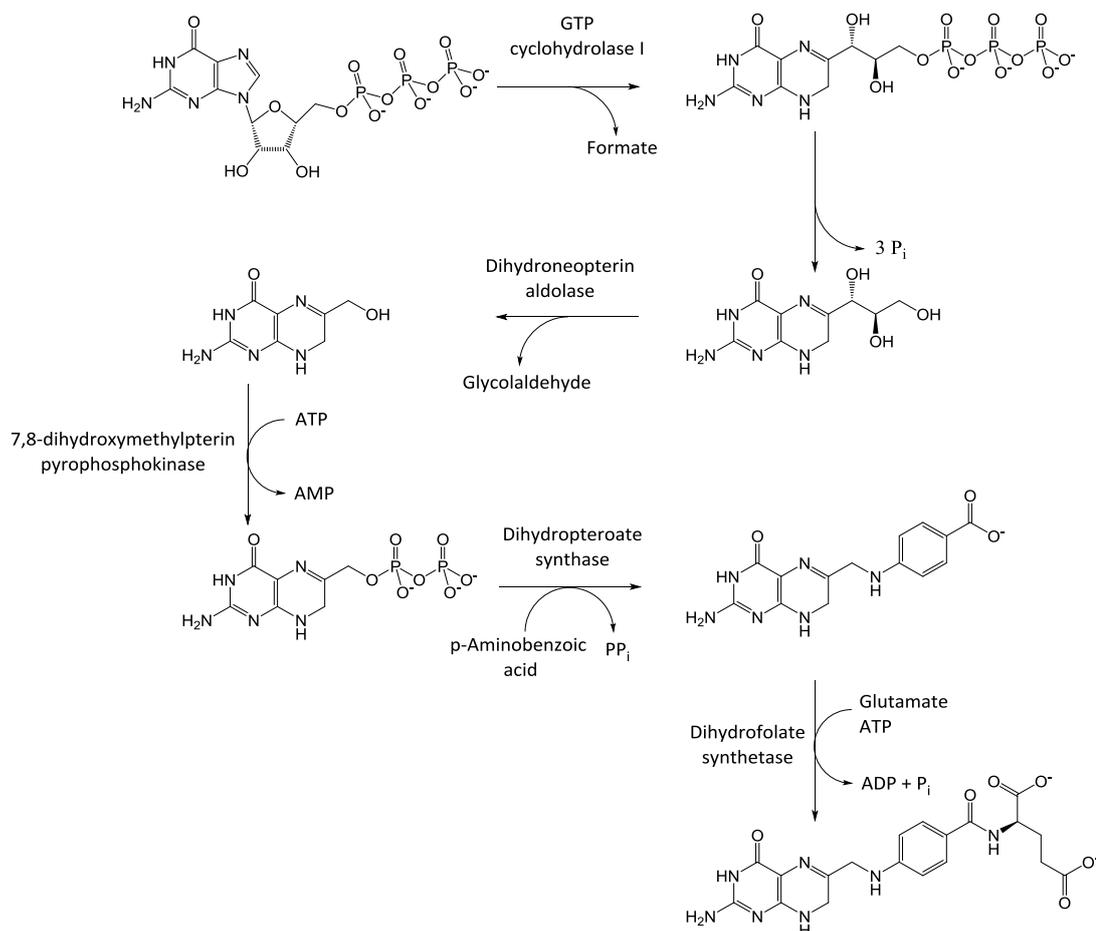
Scheme 1.6 Biosynthesis of DHF from GTP in *E. coli*.

1.5 Aims

The overall aim of this investigation was to explore the nature of catalysis in DHFR, using a number of different isotope effects. This was divided into two separate approaches; investigating the effect that isotopic substitution, particularly around the reaction centre in the cofactor/substrate, has on hydride transfer, and the role

that fs vibrations play on enzyme catalysis, both on the hydride transfer step and catalysis as a whole.

Recently, numerous DHFR enzymes, and variants, have been studied via comparison of the "light", unlabelled and "heavy", $^{13}C^{15}N^2H$ isotopically labelled enzymes (117, 118, 236). "Enzyme KIEs" highlight the role that ultrafast, short-range vibrations of the enzyme play during the catalytic cycle. Whilst long range promoting motions that couple to hydride transfer have been ruled out in all DHFRs studied to date (113, 126, 136), relatively little information has been published on the role of these fs vibrations. Following on from these investigations, the investigation of a "heavy" BsDHFR would enable a more complete picture of any general trend across DHFR enzymes. Steady state and pre-steady state analysis, at pH 7, would enable both the whole catalytic cycle and the chemical step to be analysed.

The contribution of specific atomic motions from the substrate to DHFR catalysis was to be analysed through α-secondary hydrogen (2°) and heavy atom KIEs, with the focus on the magnitude of these effects at pH 7 (rather than at elevated pH often employed for these type of studies (88)). Previous investigations have utilised 2° KIEs in the investigation of other NAD(P)H dependent enzymes (151, 153), as well as EcDHFR at elevated pH (88) and TmDHFR (155). 2° KIEs have been proposed to report on the coupling of the 2° hydrogen motion to the transfer of the primary hydrogen (152), and/or the configuration of the reaction coordinate immediately prior to the transfer of the 1° hydrogen (153). To our knowledge, there are no cases of heavy atom KIEs being utilised in the investigation of DHFR, and in fact there is very little literature surrounding the use of these isotope effects to investigate enzymatic hydride transfer in general. The role of heavy atom motions is key to numerous tunnelling models (see above), as well as the potential production of TS analogues, therefore incorporation of stable isotopes at specific locations in the substrate would potentially enable the role that they play in the formation of the reaction coordinate and transition state to be analysed.

However, to enable heavy atom KIEs to be measured, isotopically labelled cofactor and substrate must be produced in significant amounts that allow kinetic analysis to be undertaken. Thus, a main priority of this investigation was to reconstitute the biosynthetic pathways of NADP$^+$ and DHF from *E. coli in vitro*. This will provide a facile route to isotopically labelled NADPH and DHF isotopologues that are currently extremely difficult and expensive to synthesise chemically. The biosynthetic genes for both the NADP$^+$ and DHF pathway were to be cloned from *E. coli* K-12 strains into expression vectors, allowing large scale production and purification of the enzymes. As previous literature exists for each *E. coli* enzyme (see above), the focus was to incorporate as many steps of the biosynthetic pathways together as possible. There have been no previous reports of large-scale biosynthesis of NADPH or DHF *de novo*. The successful reconstitution of these pathways will allow a single route to a large number of isotopologues, using only simple starting materials that are commercially available.

2   CATALYSIS BY G. STEAROTHERMOPHILUS DHFR

2.1 Preface

Geobacillus stearothermophilus DHFR (named BsDHFR in the literature) is a moderately thermophilic enzyme that shares a very similar tertiary structure with that of the mesophilic EcDHFR (132). Previous studies have shown that the catalytic cycle of BsDHFR differs from that of EcDHFR, a long standing paradigm for enzyme catalysis (see Introduction). Thus the role of protein motions in catalysis may also be different between these two enzymes (132, 138). Here, we aimed to analyse the chemistry of BsDHFR in a detailed manner by measuring the primary kinetic isotope effect (KIE) on hydride transfer by using enzyme isotopic labelling with ^{13}C, ^{15}N and ^{2}H. This allowed a comparison between "light" (natural abundance) and "heavy" (isotopically labelled) BsDHFR. Both isotopically labelled enzymes and substrates were employed in this investigation and mentioned in the discussion. Throughout, "enzyme KIE" refers to k^{LE}/k^{HE} and "hydride KIE" refers to k^{NADPH}/k^{NADPD}.

2.1.1 Kinetic characterisation of enzyme catalysis

2.1.1.1 Steady state kinetics

The simplest enzymatic reaction between an enzyme and substrate can be represented by:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$
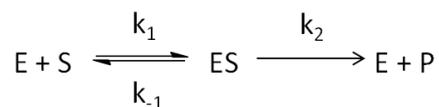
where E is the enzyme, S the substrate, ES the enzyme-substrate complex and P the product of the reaction. This representation works on the assumption that the reverse reaction (k$_{-2}$) is negligible and therefore can be ignored. The initial rate of reaction, v, can be defined as:

$$v = k_2[ES]$$

However, the concentration of ES is also dependent on rate constants other than k_2:

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

$$Rate\ of\ decomposition\ of\ ES = (k_{-1} + k_2)[ES]$$

Steady state kinetics are based on the assumption that the concentration of enzyme-substrate complex does not change over the time-frame of the reaction *i.e.* the rate of formation and decomposition of ES are equal, thus:

$$k_1[E][S] = (k_{-1} + k_2)[ES]$$

If [E] is replaced by [E$_t$]-[ES], *i.e.*, the total amount of enzyme minus the enzyme bound to the substrate, the equation above can be rearranged to give:

$$[ES] = \frac{[E_t][S]}{[S] + (k_2 + \frac{k_{-1}}{k_1})}$$

where (k_2 + k_{-1}/k_1) is the Michaelis constant, K_M, giving:

$$[ES] = \frac{[E_t][S]}{K_M + [S]}$$

By substituting v = k_2[ES] into the equation above we get:

$$v = \frac{k_2[E_t][S]}{K_M + [S]}$$

This equation can be further simplified as the maximum rate, V_{max}, occurs when the enzyme is saturated, *i.e.* V_{max} = k_2[E$_t$], therefore:

$$v = \frac{V_{max}[S]}{K_M + [S]}$$

This is the Michaelis-Menten equation, and gives the relationship between the rate of an enzyme reaction and the substrate and total enzyme concentrations. It also gives the practical definition of K_M:

$$K_M = [S], \text{ when } v = \frac{V_{max}}{2}$$

In practice, the steady-state conditions are maintained by using a much higher concentration of substrate compared to enzyme during the reaction. The DHFR

reaction can be followed by monitoring the decrease in absorption at 340 nm as NADPH is oxidised to NADP$^+$ during the course of the reaction, via UV-Vis spectroscopy. The rate of reaction can be calculated from this change in absorption through the Beer-Lambert law:

$$A = \varepsilon.c.l$$

where A is the change in absorption, l is the pathlength and c and ε are the concentration and extinction coefficient of absorbing species respectively (ε = 11,800 M^{-1} cm^{-1}) (84).

2.1.1.2   Pre-steady state kinetics

Previously, it has been shown that at pH 7, *i.e.* a physiologically relevant pH, the hydride transfer step in BsDHFR is only partially rate limiting (132). Hydride transfer in BsDHFR catalysis therefore could not be characterised by steady-state turnover (k_{cat}) measurements, as it is masked by kinetic complexity. Consequently, pre-steady state or transient state kinetics are needed to monitor the rate of hydride transfer (k_H). Stopped-flow spectroscopy has been used widely to investigate hydride transfer in DHFR catalysis at pH 7 (86, 120), and uses an excess of enzyme and substrate compared to the cofactor concentration to maintain single turnover conditions.

The stopped-flow apparatus consists of two syringes, the first containing enzyme and cofactor and the second containing substrate (Figure 2.1). The enzyme is preincubated with NADPH to avoid hysteresis (87). The syringes are injected using a single syringe drive plate to ensure equal volumes from each syringe are injected, and the solutions are rapidly mixed before passing through the detector. This method allows for very fast measurements to be recorded, as the machine has a dead time of only a few milliseconds during which the rapid mixing of solutions occurs. The stopped-flow apparatus allows the chemical step to be monitored via Förster resonance energy transfer (FRET) and absorption.
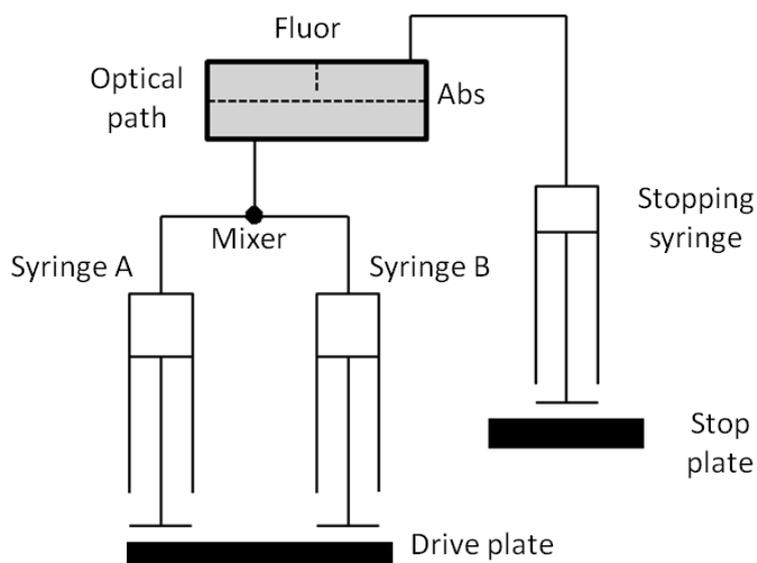
Figure 2.1 The mechanism of action of a stopped-flow spectrometer. Syringe A contains enzyme and cofactor and syringe B contains substrate. The reaction can be followed by absorption (abs) or fluorescence (fluor).

2.2 Results and discussion

2.2.1 Purification and structural characterisation of light and $^{13}C^{15}N^{2}H$ labelled DHFR

The plasmid containing the BsDHFR gene was expressed using *E. coli* BL21 Star^{TM} (DE3) cells in M9 minimal media. The characteristics of the target enzyme allowed the crude lysate to be heat-shocked, enabling many of the native *E. coli* proteins to precipitate whilst the thermophilic BsDHFR stayed in solution. Further purification was performed using anion-exchange chromatography followed by size-exclusion chromatography (Figure 2.2). For the ^{2}H-, ^{13}C- and ^{15}N-labelled "heavy" enzyme, the host cells were grown in M9 minimal media that contained isotopically labelled glucose, ammonium chloride and D_2O as the only sources of carbon, nitrogen and hydrogen. The same purification procedure was used, and they gave identical elution traces in DEAE and size-exclusion chromatographic purifications.
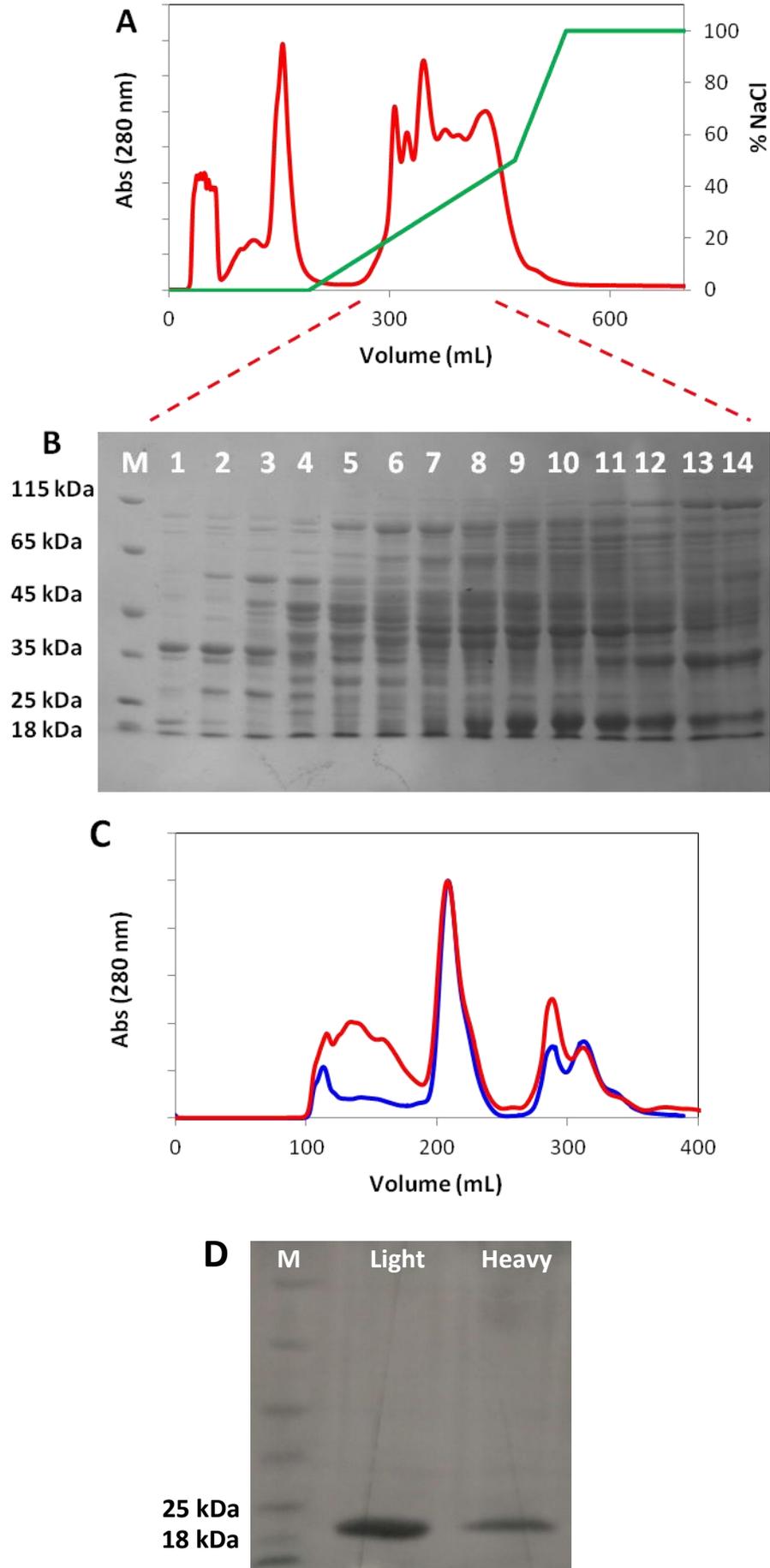
Figure 2.2 (previous page) A) Chromatogram of the elution trace (red) of light BsDHFR from the DEAE column and the corresponding NaCl concentration (green). The fractions containing BsDHFR (1-14, M = marker) were analysed via SDS-PAGE before purification via size exclusion (B). C) The elution traces of light (red) and heavy (blue) BsDHFR from the Superdex75 size-exclusion column. D) Final purifed fractions of light and heavy BsDHFR (light = light BsDHFR, heavy = $^{13}C^{15}N^{2}H$ BsDHFR, M = marker).

ESI-MS was performed to confirm the purification for both the light and heavy BsDHFR (performed by Robin Hicks and Dr. Louis Luk). The deconvoluted mass spectrum of BsDHFR that is composed of naturally abundant isotopes showed a peak at 18696.2, which matches well with the theoretical mass of 18694.7 (Figure 2.3). On the other hand, the 'heavy' BsDHFR mass spectrum showed a peak of 20740.0, indicating a ~10% molecular weight increase. This can be accounted by 99.8% of the carbon, nitrogen and non-exchangeable hydrogen atoms being labelled by their heavier counterparts.
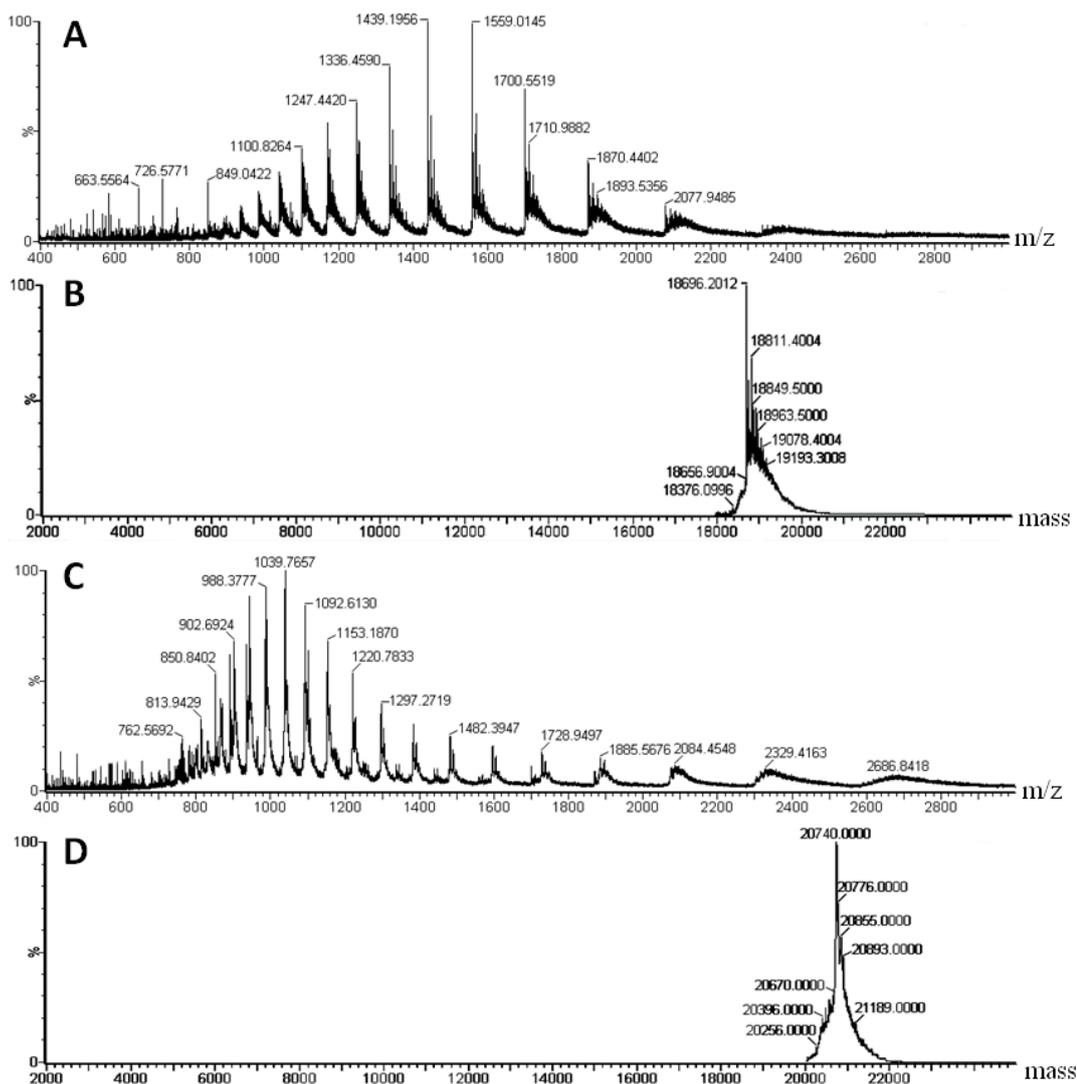
Figure 2.3 ESI-MS chromatogram of light (A) and heavy (C) BsDHFR. The deconvoluted spectra are also shown (B and D).

Circular dichrosim (CD) spectra of the light and heavy enzymes were obtained to verify if heavy isotope substitutions have an effect on protein structural organisation. The light enzyme (Figure 2.4) exhibited characteristic signals of a folded protein, with two minima at 222 and 206 nm and a maximum at 195 nm, signifying the existence of α-helices and β-sheets. At all the examined temperatures (10, 20 and 35 °C), BsDHFR remained properly folded and showed no sign of unfolding. Most importantly, the MRE of the light and heavy BsDHFR were super-imposable, confirming that there was no disparity between the isotopologues (Figure 2.4D). Hence, any reactivity difference between the light and heavy BsDHFR

will be caused by a difference in mass-dependent vibrations rather than structural difference.
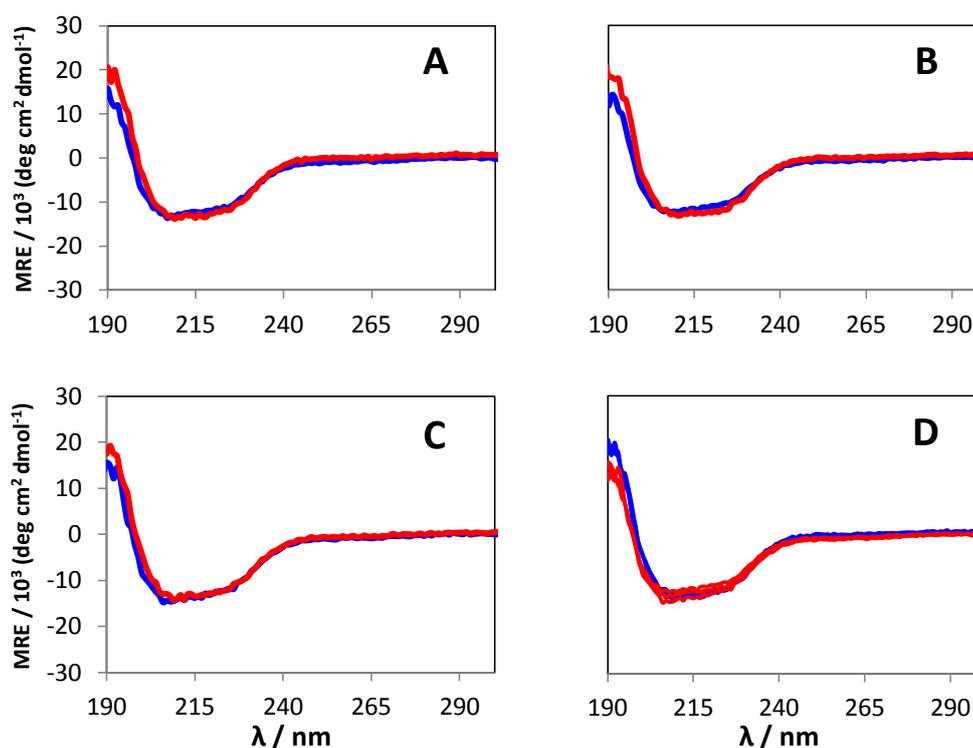
Figure 2.4 CD spectra of light BsDHFR (red) and heavy BsDHFR (blue) at 10 (A), 20 (B) and 35 (C) °C and at all temperatures (D). Each spectrum is an average of at least three measurements per temperature per enzyme.

2.2.2 Steady state kinetic analysis of light and heavy BsDHFR

Kinetic measurements were carried out to investigate the effect of heavy isotope substitution on the catalysis of BsDHFR (data measured with Dr. Louis Luk). The temperature dependence of k_{cat} was measured from 10 to 45 °C (due to the technical difficulty in lowering the temperature of the cell-holder, we were unable to accurately measure kinetics at 5 °C). The measured enzyme KIE on catalytic turnover (k_{cat}) was ~2.6, appearing to be temperature independent (Table 2.1, Figure 2.5), and is considerably larger than that measured with other DHFRs. EcDHFR, whose catalytic cycle was thoroughly characterised previously (86), showed a considerably smaller enzyme KIE that increases steadily from unity at 10 °C to 1.15 ± 0.02 at 40 °C (117). This was attributed to the large conformational

change involved in the rate-limiting product release (69, 86). On the other hand, the rate of catalytic turnover of the catalytically compromised mutant EcDHFR-N23PP/S148A is not isotopically sensitive, giving an enzyme KIE of unity across the temperature range measured (118). This observation agrees well with the fact that this mutant is unable to form the occluded conformation and lacks the large scale loop movement needed for cofactor $NADP^+$ release, the rate limiting step in this mutant (117, 118). A similar isotope effect investigation has been conducted on another thermophilic homologue TmDHFR, which is extremely rigid and locked in the open conformation (236). This enzyme produced a temperature independent enzyme KIE of approximately 1.3 across the measured temperature range. TmDHFR is unique among DHFRs in that it forms a dimer in its native state, and it is argued that the enzyme KIE of 1.3 is as a result of intersubunit interactions that are isotopically sensitive (236).

Table 2.1 Temperature dependence of k_{cat} for unlabelled and labelled BsDHFR, and the enzyme KIE, at pH 7.

T (°C)	k_{cat} (s⁻¹)		Enzyme KIE
	Unlabelled BsDHFR	$^{13}C^{15}N^2H$ BsDHFR	
10	4.71 ± 0.23	2.22 ± 0.15	2.13 ± 0.22
15	8.44 ± 0.37	3.42 ± 0.14	2.47 ± 0.15
20	15.64 ± 0.01	6.41 ± 0.44	2.44 ± 0.24
25	24.94 ± 0.56	9.14 ± 0.99	2.73 ± 0.19
30	41.32 ± 3.82	13.87 ± 2.76	2.98 ± 0.35
35	53.45 ± 0.53	19.79 ± 3.63	2.70 ± 0.18
40	76.67 ± 0.99	30.93 ± 2.93	2.48 ± 0.18
45	99.34 ± 1.72	37.42 ± 1.79	2.65 ± 0.33
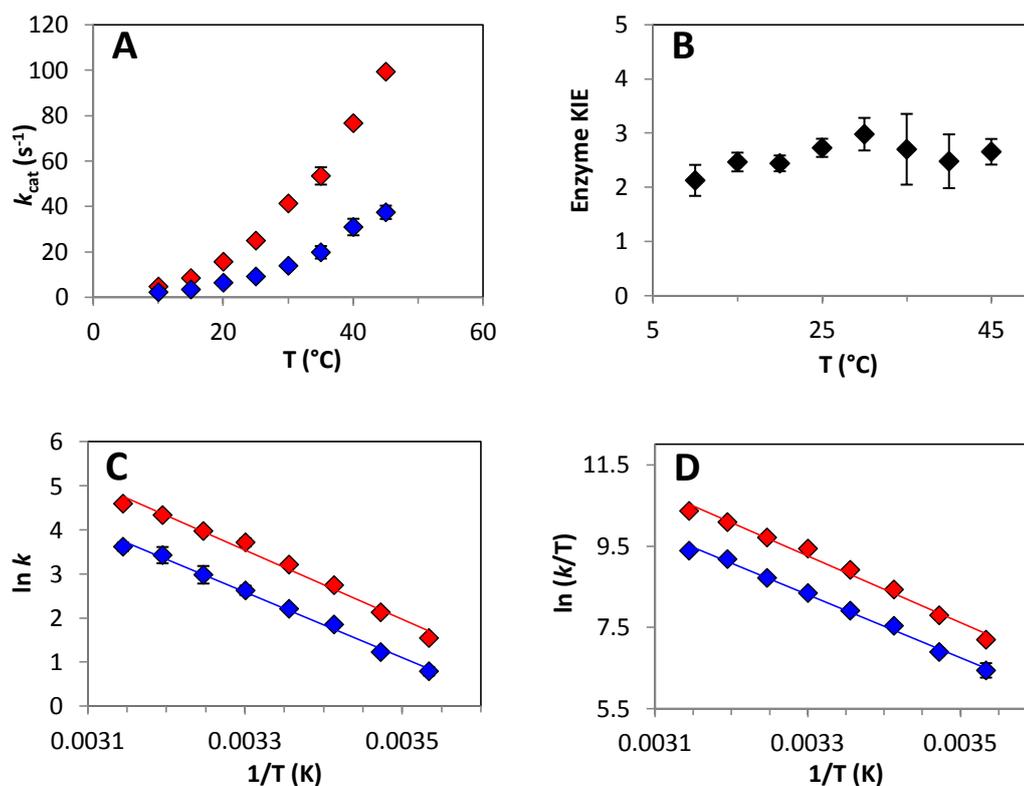
Figure 2.5 A) Temperature dependence of k_{cat} for light (red) and heavy (blue) BsDHFR at pH 7. B) Temperature dependence of the enzyme KIE on k_{cat} for light and heavy BsDHFR at pH 7. C) Corresponding Arrhenius plot of the temperature dependence of k_{cat} of light and heavy BsDHFR (red and blue respectively) at pH 7. D) Corresponding Eyring plot of the temperature dependence of k_{cat} of light (red) and heavy (blue) BsDHFR at pH 7.

The temperature independent enzyme KIE on k_{cat} for BsDHFR suggests that a large loop movement or conformational change is not involved in the catalytic cycle of this enzyme. This is supported by the fact that evidence of the occluded conformation has never been documented in the case of BsDHFR, and thus large conformational changes are absent from the catalytic cycle in much the same way as with the triple mutant EcDHFR and TmDHFR (enzymes that show similar temperature independent enzyme KIEs).

The magnitude of the enzyme KIE in BsDHFR is likely due to a number of factors. BsDHFR has shown to be inherently more flexible than EcDHFR (133, 134), and as such atomic motions on a global scale would be expected to show a greater isotope dependence. It has also been shown that hydride transfer is partially rate determining at pH 7 (unlike in EcDHFR) as a KIE of ~2 was measured when NADPD

was used under steady state conditions (132). This kinetic complexity surrounding the catalytic cycle suggests a number of steps may be affected by isotopic substitution, rather than a single large conformational change as witnessed in EcDHFR, and may contribute to the larger enzyme KIE seen in BsDHFR. Higher pH measurements would enable the enzyme KIE solely on the chemical step to be determined, however, the very slow rate at high pH made a comparison between the two isotopologues impractical due to significant experimental error.

Table 2.2 Activation parameters for light and heavy BsDHFR at pH 7 under steady state conditions.

	E_a (kJ mol^{-1})	A_H (s^{-1})	$\Delta H^{\ddagger}$ (kJ mol^{-1})	$\Delta S^{\ddagger}$ (J K^{-1} mol^{-1})	$\Delta G^{\ddagger}$ (kJ mol^{-1}) (at 25 °C)
Light BsDHFR	65.4 ± 3.1	(6.57 ± 0.3) x 10^{12}	62.9 ± 3.1	-7.9 ± 0.4	65.3 ± 3.2
Heavy BsDHFR	62.0 ± 1.5	(6.57 ± 0.2) x 10^{11}	59.6 ± 2.0	-27.0 ± 1.0	67.6 ± 2.2

The difference in catalytic turnover (k_{cat}) was further analysed by deducing the activation parameters of the light and heavy BsDHFR (Table 2.2). Because the observed enzyme KIE is temperature independent, activation energies (E_a) as well as the enthalpies of activation $\Delta H^{\ddagger}$ are identical between the light and heavy enzymes. Interestingly, the entropy of activation $\Delta S^{\ddagger}$ (and consequently the A_H value) underlies the difference in reactivity. The transition state of the heavy BsDHFR is more ordered, and the calculated $\Delta S^{\ddagger}$ is approximately 20 J K^{-1} mol^{-1} lower than that of the light enzyme (-7.91 ± 0.4 and -27.0 ± 1.0 J K^{-1} mol^{-1} for the light and heavy enzyme respectively). This implies that the enzyme has become more rigid upon heavy isotope substitution, and energetically less favourable for conformational sampling. Therefore, a slightly higher reorganisation energy is required (51, 118, 237), and a slower reaction rate was observed. BsDHFR is known to be more flexible than EcDHFR (134), and this could explain the greater

magnitude of enzyme KIE effect as the inherent flexibility results in a greater entropic burden on the thermophilic enzyme when a rigid conformation is adopted.
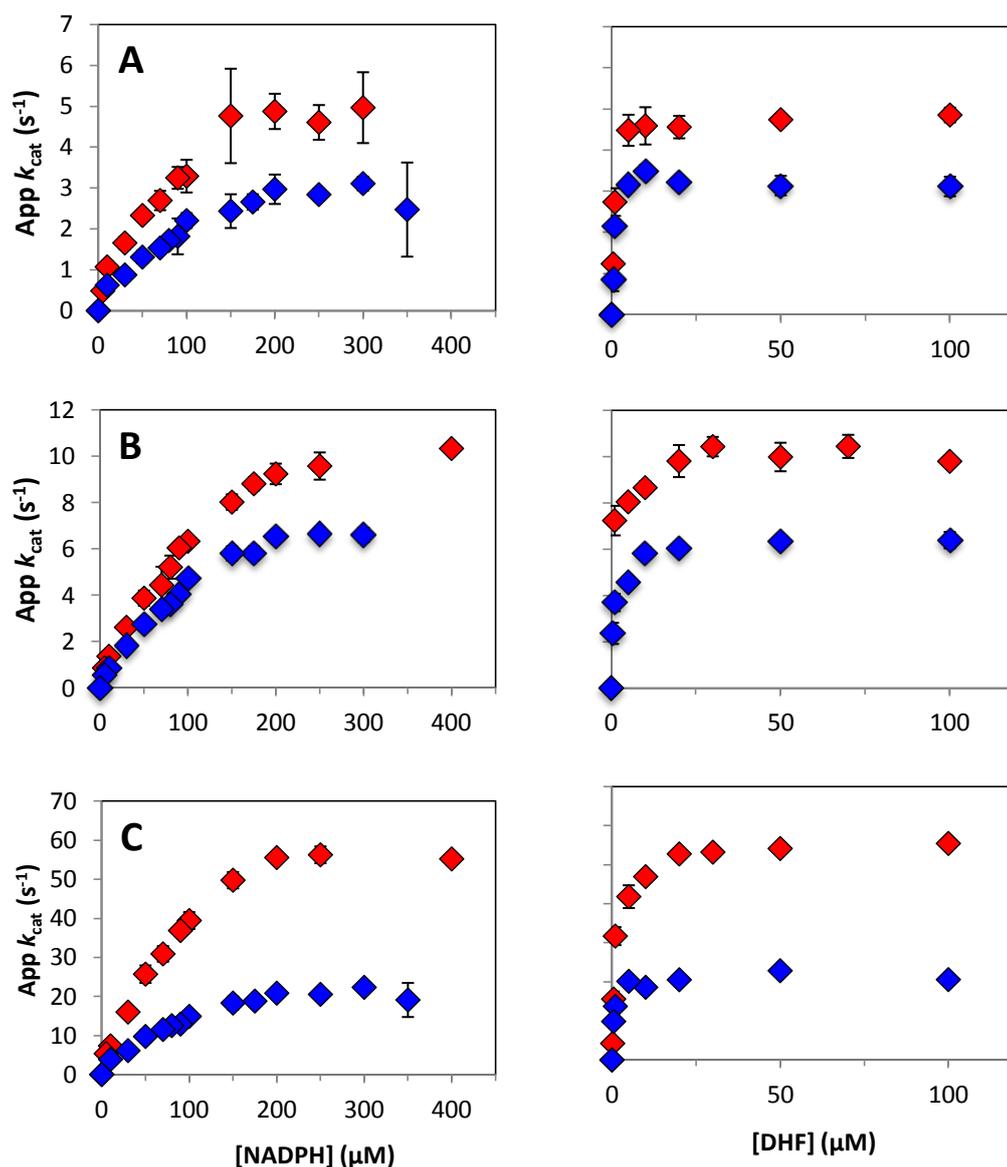
Figure 2.6 Effect of cofactor (left) and substrate (right) concentration at 10 (A), 20 (B) and 35 (C) °C on the apparent (app) k_{cat} value for light and heavy BsDHFR (red and blue respectively) at pH 7.

The Michaelis constants (K_M) for the light and heavy enzyme were determined at 10, 20 and 35 °C for both the cofactor NADPH and substrate DHF (Figure 2.6). The

cofactor K_M for the light and heavy BsDHFR at 20 °C was 127 ± 13.8 and 124 ± 14.9 µM, respectively, showing no difference between the two enzymes. Similarly, at 10 °C and 35 °C the K_M^{NH} values for the light and heavy BsDHFR were identical within error, suggesting no significant variation caused by heavy isotope substitutions (Table 2.3). Therefore, the observed difference on the rate of catalytic turnover likely occurs after formation of the Michaelis complex.

It should be noted that the Michaelis constant for NADPH is noticeably higher than that of other characterised DHFRs. This is due to a conserved glutamine residue normally present in the ABD being replaced with alanine, a more thermostable residue. Previous studies have shown the A104Q mutant of BsDHFR has a fourfold higher affinity to NADPH (138). For DHF, the nature of tight binding is similar to other DHFRs and remains unchanged upon isotope labelling, with K_M values measured to be approximately 1 µM (Table 2.3).

Table 2.3 Apparent k_{cat} and K_M values for light and heavy BsDHFR at 10, 20 and 35 °C at pH 7.

°C	BsDHFR	*App k_{cat}* (s^{-1})	K_M^{NADPH} (µM)	K_M^{DHF} (µM)
10	Light	4.7 ± 0.2	86 ± 15.8	0.98 ± 0.18
	Heavy	2.2 ± 0.2	112.5 ± 15.7	0.79 ± 0.23
20	Light	15.6 ± 0.1	127 ± 13.8	0.56 ± 0.14
	Heavy	6.4 ± 0.4	124 ± 14.88	0.84 ± 0.15
35	Light	53.5 ± 0.5	88.2 ± 12.5	1.01 ± 0.17
	Heavy	19.8 ± 3.6	112 ± 11.52	0.57 ± 0.1

2.2.3   Pre-steady state kinetic analysis of light and heavy BsDHFR

2.2.3.1   *Hydride transfer and kinetic isotope effect of BsDHFR and heavy BsDHFR*

The effect of heavy isotope substitution on the BsDHFR catalysed hydride transfer was measured in single-turnover experiments. Previously, the rate constant of

hydride transfer (k_H) was measured by fluorescence resonance energy transfer (FRET), in which a tryptophan residue inside the active site serves as the donor chromophore, while the nicotinamide ring in NADPH serves as the acceptor chromophore. However, this technique often produced data with a low signal to noise ratio, especially at low temperature (Figure 2.7). In fact, the rate constant of hydride transfer could not be measured accurately at 5 °C. To improve the quality of data, absorbance was measured at 340 nm. The measured k_H values obtained by absorbance are in good agreement with that by FRET, and data can be obtained accurately even at 5 °C (Table 2.4). Hydride transfer rate constants were determined between 5 and 45 °C for both light and heavy BsDHFR (Table 2.5).

Table 2.4 Comparison of k_H values obtained via absorbance and FRET measurements.

T (°C)	k_H (s^{-1})	
	Absorbance	**FRET**[A]
5	46.7 ± 3.2	-
10	67.1 ± 1.2	68.3 ± 8.3
15	84.7 ± 3.9	79.2 ± 3.7
20	112.3 ± 5.0	102 ± 5
25	139.5 ± 2.4	131 ± 2
30	154.2 ± 5.9	160 ± 8
35	188.3 ± 8.9	184 ± 6
40	213.8 ± 21.0	218 ± 7
45	256.3 ± 26.2	249 ± 12
[A] Data measured by J. Guo (136)		
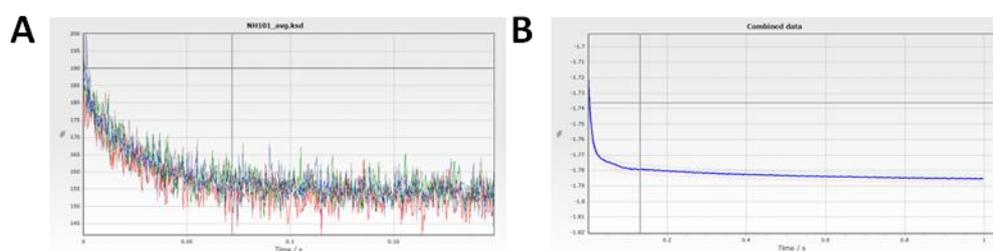
Figure 2.7 Comparison of traces obtained via FRET (A) and absorption (B) measurements.

Table 2.5 k_H and k_D rate constants for light and heavy BsDHFR and the calculated KIE for k_H/k_D

T (°C)	Light BsDHFR			Heavy BsDHFR		
	k_H	k_D	KIE	k_H	k_D	KIE
5	46.7 ± 3.2	7.7 ± 1.5	6.1 ± 0.3	28.3 ± 2.5	3.3 ± 0.3	8.7 ± 0.6
10	67.1 ± 1.2	13.2 ± 0.7	5.1 ± 0.3	49.4 ± 2.6	6.4 ± 1.3	7.7 ± 0.8
15	84.7 ± 3.9	19.7 ± 0.6	4.3 ± 0.3	71.4 ± 3.3	19.1 ± 1.9	3.7 ± 0.2
20	112.3 ± 5.0	32.2 ± 2.5	3.5 ± 0.4	91.4 ± 5.0	32.5 ± 0.8	2.8 ± 0.1
25	139.5 ± 2.4	43.9 ± 4.5	3.2 ± 0.3	122.2 ± 4.0	45.9 ± 2.7	2.7 ± 0.1
30	154.2 ± 5.9	60.7 ± 11.4	2.5 ± 0.5	139.8 ± 8.8	61.1 ± 3.6	2.3 ± 0.1
35	188.3 ± 8.9	69.1 ± 8.8	2.7 ± 0.6	186.1 ± 4.6	75.0 ± 3.1	2.5 ± 0.1
40	213.8 ± 21.0	76.6 ± 6.4	2.8 ± 0.6	215.5 ± 10.8	94.4 ± 8.3	2.3 ± 0.1
45	256.3 ± 26.2	97.2 ± 15.5	2.6 ± 0.5	235.3 ± 7.2	123.1 ± 21.9	1.9 ± 0.2
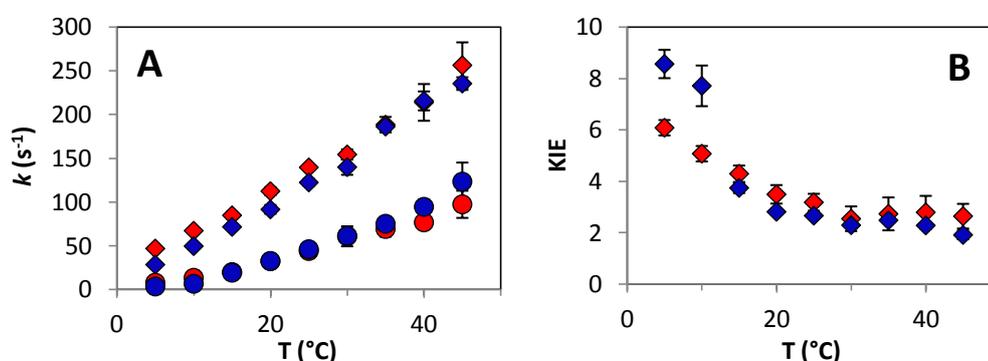
Figure 2.8 A) Graph showing the k_H (diamonds) and k_D (circles) rate constants for light and heavy BsDHFR (red and blue respectively).

Intriguing observations were seen in the hydride transfer rate constant measurements. Between 5 and 45 °C, hydride transfer (k_H) of the light enzyme ranges from 46.7 ± 3.2 to 256.3 ± 26.2 s^{-1}. The rate constant dropped significantly when labelled substrate NADPD was used, with k_D spanning from 7.7 ± 2.5 to 97.2 ± 7.2 s^{-1} at the corresponding temperatures. The calculated "hydride KIE" (k_H/k_D) is approximately 2.5 between 20-45 °C, but increases steeply at low temperature, giving a KIE of 6.1 ± 0.3 at 5 °C. The strong temperature dependence resembles that

discovered in the studies of the hyperthermophilic homologue TmDHFR (120) and a moderately thermophilic alcohol dehydrogenase BsADH (21). Interestingly, this biphasic behaviour is further magnified when isotopically labelled enzyme was employed. k_H ranges from 28.3 ± 1.5 to 235.3 ± 15.5 s^{-1}, whereas k_D starts from 3.3 ± 0.3 to 123.1 ± 21.9 s^{-1} between 5 to 45 °C for the heavy enzyme (Figure 2.8). Hence, a similar breakpoint at ~20 °C was observed, giving a calculated KIE that increases from 2.7 ± 0.1 at 25 °C to 8.7 ± 0.6 at 10 °C (Figure 2.10).

Table 2.6 Enzyme KIE values on hydride and deuteride transfer

T (°C)	Enzyme KIE^H	Enzyme KIE^D
5	1.65 ± 0.09	2.32 ± 0.25
10	1.36 ± 0.04	2.07 ± 0.21
15	1.18 ± 0.04	1.03 ± 0.05
20	1.23 ± 0.04	0.99 ± 0.04
25	1.14 ± 0.02	0.96 ± 0.06
30	1.10 ± 0.04	0.99 ± 0.10
35	1.01 ± 0.03	0.92 ± 0.06
40	0.99 ± 0.05	0.81 ± 0.05
45	1.09 ± 0.06	0.79 ± 0.09
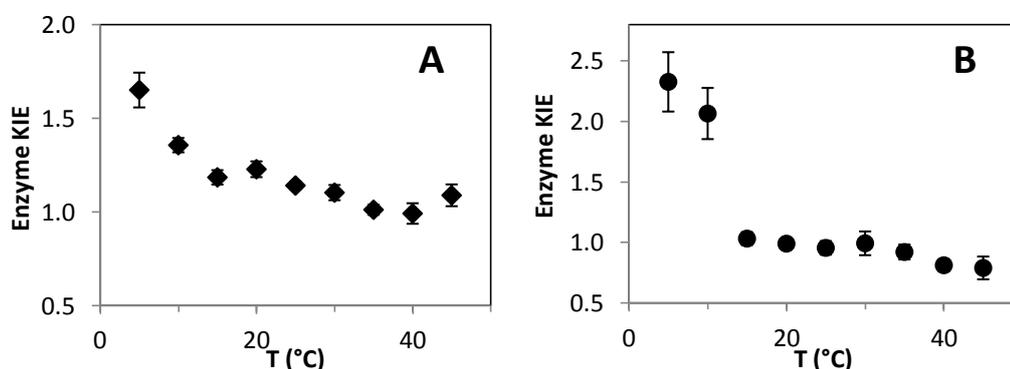
Figure 2.9 Enzyme KIE on hydride (A) and deuteride (B) transfer at pH 7.

The effect of enzymatic isotopic substitution was assessed by calculating the enzyme KIE^H (k_H^{LE}/k_H^{HE}). Similarly, this parameter showed a breakpoint of

temperature dependence. As temperature increases from 5 to 20 °C, the enzyme KIE^H drops sharply from 1.65 to 1.23. However, at 25 °C and above, the enzyme KIE^H becomes significantly less temperature dependent, and eventually reaches unity at 40 °C, indicating that enzymatic isotope substitution causes no effect on the rate constant of hydride transfer. For the labelled cofactor NADPD, enzyme KIE^D (k_D^{LE}/k_D^{HE}) also decreases from 2.33 at 5 °C to 0.99 to 20 °C. Interestingly, at high temperature the rate constant of deuteride transfer in the heavy BsDHFR is higher than that in the light counterpart. The enzyme KIE^D was calculated to be 0.81 at 40 °C, implying that the efficiency of deuteride transfer was actually enhanced when the mass-dependent vibrations of the enzyme were reduced.

Table 2.7 pK_a values for the hydride transfer reaction of light and heavy BsDHFR at 10, 20 and 35 °C

	pK_a		
	10 °C	**20 °C**	**35 °C**
Light BsDHFR	6.64 ± 0.1	6.41 ± 0.1	6.75 ± 0.2
Heavy BsDHFR	6.33 ± 0.2	6.33 ± 0.1	6.92 ± 0.1

The pK_a was measured at 10, 20 and 35 °C for light and heavy BsDHFR to rule out a change between the two enzymes as the cause of the observed hydride/enzyme KIE. Measurements were performed using MTEN buffer between pH 4.5 and 9 (Table 2.7). There was no difference between the two enzymes at each temperature, indicating that a change in pK_a is not the cause of the interesting results.
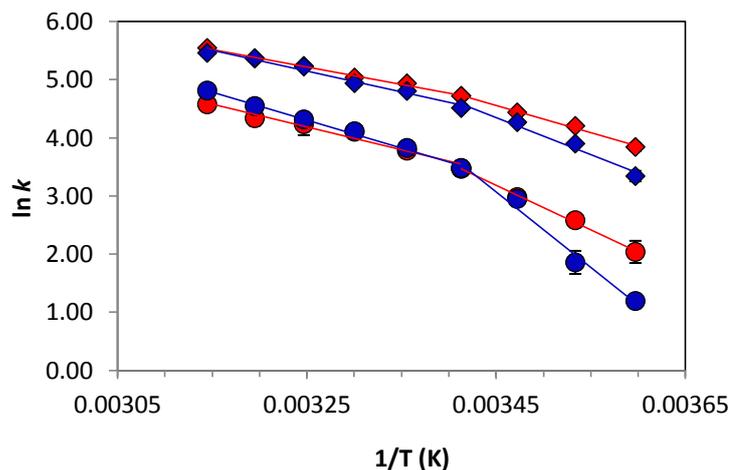
Figure 2.10 Arrhenius plot of the k_H and k_D (diamonds and circles) for the light and heavy BsDHFR (red and blue).

Table 2.8 Activation energy and Arrhenius prefactor values of light and heavy BsDHFR under pre-steady state conditions at pH 7 between 5 and 45 °C

T (°C)		Light BsDHFR		Heavy BsDHFR	
		NADPH	NADPD	NADPH	NADPD
5-20	E_a (kJ mol^{-1})	38.4 ± 2.1	63.6 ± 2.3	52.8 ± 7.0	107.7 ± 9.6
	$A_{H(D)}$ (s^{-1})	9.5 (± 0.4) x10^8	7.0 (± 0.2) x10^{12}	2.5 (± 0.3) x10^{11}	5.6 (± 0.5) x10^{20}
	ΔE_a (kJ mol^{-1})	24.8 ± 3.2		54.9 ± 11.4	
	A_H/A_D	1.8 (± 0.1) x10^{-4}		4.5 (± 0.6) x10^{-10}	
20-45	E_a (kJ mol^{-1})	24.8 ± 1.0	32.6 ± 3.0	29.8 ± 2.2	40.0 ± 1.4
	$A_{H(D)}$ (s^{-1})	3.0 (± 0.08) x10^6	2.2 (± 0.2) x10^7	2.0 (± 0.1) x10^7	4.6 (± 1.3) x10^8
	ΔE_a (kJ mol^{-1})	7.8 ± 3.1		10.2 ± 2.6	
	A_H/A_D	0.14 ± 0.01		0.04 ± 0.002	

The biphasic behaviour observed in the hydride transfer rate constant measurement was characterised by calculating two sets of activation parameters (Table 2.8). At temperatures above 20 °C, the hydride KIE is weakly temperature dependent and showed a small difference in activation energies (ΔE_a = 7.8 ± 3.1 and 10.2 ± 3.1 kJ mol^{-1} for light and heavy BsDHFR respectively). However, because the KIE is strongly temperature dependent at 5-20 °C, the magnitude of ΔE_a greatly increased. Moreover, the change in activation energy for heavy enzyme (54.9 ± 11.4

kJ mol^{-1}) is almost twice that in the unlabelled enzyme (24.8 ± 3.2 kJ mol^{-1}). Both the light and heavy enzyme also display strongly inverse A_H/A_D values (1.8 (± 0.1) x10^{-4} and 4.5 (± 0.6) x10^{-10} respectively) in the low temperature region.

Similar kinetic behaviour has been seen in BsADH (21) and TmDHFR (120). It was originally proposed that this dramatic change in hydride KIE was caused by a contribution from "promoting motions", which function at a physiologically relevant temperature by modulating hydride transfer through the barrier height and/or width (21, 41). These hypothetical promoting motions range from femtosecond (fs) bond vibrations to millisecond (ms) structural changes. Several experiments have been designed to probe for the existence of these "promoting motions". A role for long range dynamic coupling has been ruled out in the studies of EcDHFR, MpDHFR, TmDHFR and BsDHFR, since increasing viscosity of the reaction medium by varying solvent compositions does not affect the nature of hydride transfer. Instead, the electrostatic properties of the reaction medium appear to correlate to the efficiency of hydride transfer (105, 113, 126, 138). Furthermore, fast protein vibrations do not seem to promote catalysis by coupling to the chemical coordinate of hydride transfer. A computational study of heavy EcDHFR and its catalytically compromised mutant N23PP/S148A indicated that the tunnelling coefficient (*i.e.* the probability of hydride tunnelling) remains unaffected upon heavy isotope substitution of the enzyme (117, 118). Additionally, isotopic labelling of TmDHFR caused no observable effect on the rate constant of hydride transfer (236), again indicating the absence of vibrational coupling to the reaction coordinate. Together these results strongly challenge the existence of promoting motions with regards to DHFR catalysis. Accordingly, the reactivity difference in BsDHFR between NADPH and NADPD, and that caused by enzymatic isotope substitution, is unlikely due to a change in the contribution of promoting motions.

Instead, there are a number of factors that could affect the efficiency of hydride transfer in BsDHFR. During catalysis most enzymes undergo conformational sampling, by which an ideal electrostatic environment is preorganised for the transition state of the chemical transformation. In DHFR, this process greatly affects the rate constant of hydride transfer. This is demonstrated by the studies of

EcDHFR and its catalytically compromised mutant N23PP/S148A (110). Reduction of the hydride transfer rate constant in the EcDHFR-N23PP/S148A catalysed reaction is mainly caused by poor electrostatic preorganisation and consequently an increase of reorganisation energy (110). Furthermore, conformational sampling likely relies on structural flexibility of the enzyme and is consequently sequence-specific. In the thermophilic homologue TmDHFR, conformational sampling is thought to be limited because it is locked in the open conformation (121), thus contributing to the relatively low rate constant of hydride transfer observed (120). Also, the frequency of dynamic recrossing, the conversion of product back to reactant through the transition state dividing surface, could be a factor that slows the hydride transfer rate in the BsDHFR reaction (27). Previous computational studies revealed that heavy isotope replacement of EcDHFR disturbed the detailed thermodynamic balance of the reaction coordinate by slowing mass-dependent protein environmental motions (117). Recrossing events in the heavy EcDHFR therefore increased, and the corresponding coefficient deviates further from unity than that of the light enzyme (117). Even though the exact mechanism has not been elucidated, further analysis indicated that dynamic recrossing is actually enhanced when coupling of femtosecond vibrations of the enzyme to the reaction centre is enhanced, compromising catalytic efficiency (118). Results obtained in the heavy enzyme studies of TmDHFR agree well with this conclusion; the rate constant of hydride transfer was shown to be unaffected by enzymatic isotopic substitutions (236).

While the exact cause of the strong temperature dependence in both the hydride and enzyme KIE has not been fully solved, some explanations are postulated in this thesis. The large increase in hydride KIE below the breakpoint suggests that formation of an active site conducive to deuteride transfer is more difficult. The heavy enzyme is affected in the same manner, with an identical hydride KIE trend observed. To begin with, the reactivity difference between NADPH and NADPD, resulting in a biphasic hydride KIE plot, could arise from conformational sampling (110, 113). Previous studies have highlighted that EcDHFR samples multiple substrates during the catalytic cycle, with high energy conformations adopted

before proceeding to the subsequent step (71). At low environmental temperatures, conformational sampling of active sites complementary to the TS would be hindered in BsDHFR, highlighted by the substantial increase in E_a below the breakpoint (Table 2.8).

It has also been hypothesised that certain enzymes can adopt multiple conformers that are favourable for chemical transformation and lead to an abnormal hydride KIE trend (238). The biphasic Arrhenius plot discovered in BsDHFR (Figure 2.10) can be interpreted as evidence of a switch in equilibrium between these conformers (239). This hypothetical change in conformational equilibrium would account for the biphasic hydride KIE at low temperature, with one conformer favouring hydride transfer over the other. Conformational sampling and changes to recrossing frequency (see below) would also be relevant in this case, with different conformers affected to different degrees across the temperature range. However, there is no direct evidence to date of multiple conformers existing in any DHFR homologue.

Additionally, the hydride KIE can be explained using ensemble-averaged variational transition state theory that is incorporated with multidimensional tunnelling (EA-VTST/MT). This theoretical framework has been used to explain the temperature dependence of the hydride KIE in EcDHFR (240). It was postulated that the reaction pathway changes with temperature, leading to different tunnelling and recrossing coefficients. Accordingly, as BsDHFR is considerably more flexible, a greater change in reaction coordinate at low temperature could result in the observed hydride KIE. In all likelihood, a contribution from all the above factors to varying extents causes the weakly temperature dependent KIE above the breakpoint and strongly temperature dependent trend below it.

At physiologically relevant temperatures (>25 °C), the magnitude of the average enzyme KIE[H] is approximately 1.07 and is analogous to that observed in EcDHFR. Accordingly, the lower hydride transfer rate constant in the heavy BsDHFR reaction can be reasoned by an increase in the frequency of dynamic recrossing (117). On the other hand, the sudden increase of the enzyme KIE[H] at low temperature (1.65

at 5 °C) has never been observed in previous heavy enzyme studies. Since BsDHFR is the most flexible thermophilic homologue that has been studied by enzymatic isotope substitutions, a greater entropic burden is expected for conformational sampling, thus a greater disparity of enzyme KIE^H was expected. This could contribute to the large magnitude in enzyme KIE^H at low temperature, in much the same manner as the observed trend in hydride KIE. The possible existence of multiple conformers would also affect the heavy enzyme, to a greater extent than the lighter enzyme, contributing to the increase in enzyme KIE^H. As the hydride KIE is also sensitive to enzyme labelling (Table 2.6), it implies that the reaction coordinate is affected by the change in protein mass-dependent vibrations. Nevertheless, all of these hypotheses require verification, and additional computational studies are needed to pinpoint the exact effect of enzymatic isotopic substitution in the catalysis by BsDHFR (see below).

At 20 °C and above, the rate of deuteride transfer is faster in heavy BsDHFR, a trend also observed in heavy EcDHFR (unpublished) and PETNR (19). This phenomenon is particularly perplexing, since many current theoretical models assume that slowing mass-dependent motions would decrease the reactivity of an enzyme (54). One possible explanation is that slower fs vibrations along the reaction coordinate and/or higher temperatures actually result in favourable enzyme-TS interactions. This tighter interaction would result in a more constrained TS compared to the GS, thus inverse secondary KIEs would result (241). A role for fs motions towards this interesting finding cannot be ruled out, or proven at this time. Further EA-VTST calculations of the deuteride transfer will provide a clearer explanation to the observed inverse enzyme KIE^D.

2.2.3.2 *Solvent isotope effect of light and heavy BsDHFR*

Hydride transfer rate constants in light and heavy BsDHFR were measured in reaction medium composed of D$_2$O. At pH 7 in 100 mM K$_i$PO$_4$, 100 mM NaCl, 10 mM βME buffer, a normal SKIE (k^{H2O}/k^{D2O}) was observed at 5 and 10 °C, and an inverse SKIE between 15 and 45 °C.

Table 2.9 k_H values for light and heavy BsDHFR in H_2O and D_2O and the subsequent SKIE.

T (°C)	Light BsDHFR			Heavy BsDHFR		
	k_H (H_2O)	k_H (D_2O)	SKIE	k_H (H_2O)	k_H (D_2O)	SKIE
5	46.7 ± 3.2	41.7 ± 0.9	1.17 ± 0.08	28.3 ± 2.5	29.0 ± 0.5	1.12 ± 0.07
10	67.1 ± 1.2	62.0 ± 1.4	1.07 ± 0.08	49.4 ± 2.6	47.0 ± 0.9	1.05 ± 0.09
15	84.7 ± 3.9	87.4 ± 1.1	0.90 ± 0.01	71.4 ± 3.3	76.0 ± 2.3	0.84 ± 0.09
20	112.3 ± 5.0	129.8 ± 1.7	0.87 ± 0.05	91.4 ± 5.0	116.3 ± 4.8	0.75 ± 0.04
25	139.5 ± 2.4	157.5 ± 1.9	0.93 ± 0.03	122.2 ± 4.0	161.7 ± 9.0	0.73 ± 0.04
30	154.2 ± 5.9	186.3 ± 3.0	0.84 ± 0.02	139.8 ± 8.8	185.7 ± 11.0	0.65 ± 0.06
35	188.3 ± 8.9	222.7 ± 5.4	0.84 ± 0.03	186.1 ± 4.6	264.2 ± 5.3	0.71 ± 0.05
40	213.8 ± 21.0	244.8 ± 5.6	0.84 ± 0.02	215.5 ± 10.8	297.8 ± 7.2	0.66 ± 0.04
45	256.3 ± 26.2	279.3 ± 5.5	0.91 ± 0.06	235.3 ± 7.2	318.1 ± 19.2	0.66 ± 0.05
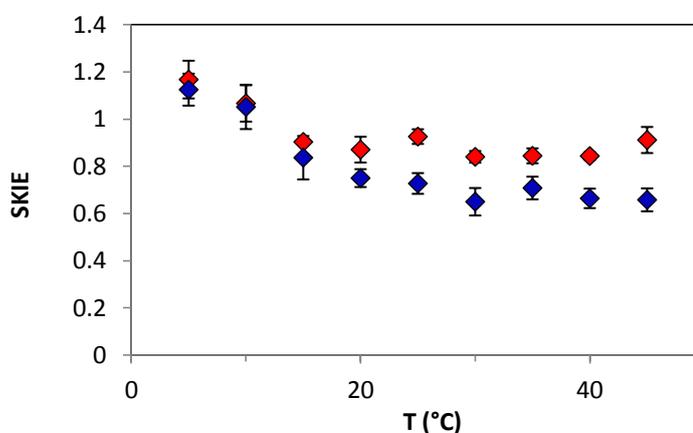
Figure 2.11 SKIE on k_H of light and heavy BsDHFR (red and blue respectively) between H_2O and D_2O

An inverse SKIE indicated that the reaction rate is higher in the deuterated solvent (Figure 2.11), agreeing with trends seen in EcDHFR, TmDHFR and MpDHFR (113, 124). The light enzyme has an average SKIE of 0.88 ± 0.04 between 15 and 45 °C, whilst heavy BsDHFR has a greater inverse SKIE of 0.71 ± 0.07 over the same temperature range. At low temperatures the SKIE increases to above unity for both enzymes. The SKIE was measured to be 1.17 ± 0.08 and 1.12 ± 0.07 for the light and heavy enzyme respectively. The pK_a was measured in D_2O at 10, 20 and 35 °C to determine whether the observed SKIE was due to a significant change between the

light and heavy enzyme at different temperatures (Table 2.10). The results show no significant differences across the temperature range that would explain the sudden change in SKIE that is observed at low temperature.

Table 2.10 pK_a values for the hydride transfer reaction of light and heavy BsDHFR at 10, 20 and 35 °C in D$_2$O.

	pK_a		
	10 °C	20 °C	35 °C
Light BsDHFR	6.77 ± 0.0	6.71 ± 0.1	6.76 ± 0.1
Heavy BsDHFR	6.48 ± 0.1	6.58 ± 0.2	6.62 ± 0.1

The cause for this increase in SKIE for both enzymes is unclear. The "normal" SKIE below 15 °C could support the idea of two conformers in equilibrium with each other. However, without significant evidence of multiple conformers, this cannot be confirmed. In addition to the theories postulated above, a "super heavy" enzyme is created when D$_2$O was used as reaction medium, as all the exchangeable hydrogen are replaced with deuterium. This increases the mass by a further 1.6%, and could contribute to the skewed kinetics observed. Overall further experiments are needed to enable full interpretation of the SKIE results that are observed at low temperatures for light and heavy BsDHFR.

2.2.4 Current progress on the computational studies of BsDHFR

Preliminary results accomplished by Moliner and coworkers (unpublished) have revealed some insights into the mechanism of the BsDHFR catalysed reaction. These preliminary results suggest that temperature dependent deviations in the reaction pathway are not the sole cause of the observed hydride KIE trend. There is little difference in recrossing coefficient (γ) between 5 and 25 °C, (0.619 and 0.611 respectively) for the light enzyme. In fact, this difference increases above the breakpoint with γ= 0.611 and 0.527 at 25 and 45 °C respectively. This implies that the large increase in magnitude of hydride KIE at low temperature stems from formation of the reaction coordinate, rather than significant deviations in reaction

pathway. This would be expected, as fs vibrations on a global scale, previously shown to affect recrossing events (117), are not altered in the measurement of the hydride KIE. Future calculations involving transfer of deuteride will provide further insights into the observed affects at low temperature, and elucidate if tunnelling coefficients differ significantly above or below the breakpoint.

The contribution of tunnelling in hydride transfer drops with increasing temperature, since the corresponding coefficient (κ) decreases from 4.2 ± 0.5 at 298 K to 2.9 ± 0.5 at 318 K (Table 2.11). This agrees well with the phenomenon that tunnelling is temperature independent, while barrier crossing increases at high temperature. Moreover, this parameter remains unaffected by enzymatic isotope substitutions and gives essentially identical tunnelling coefficients between light and heavy BsDHFR at each examined temperature. The height and/or width of barrier crossing is therefore independent of mass-dependent motions of the enzyme.

Table 2.11 Preliminary tunnelling (κ) and recrossing (γ) coefficients for hydride transfer in light and heavy BsDHFR at 278, 298 and 318 K.

	278 K		298 K		318 K	
	κ	γ	κ	γ	κ	γ
Light BsDHFR	4.2 ± 0.5	0.619	3.5 ± 0.3	0.611	2.9 ± 0.5	0.527
Heavy BsDHFR	4.2 ± 0.5	0.402	3.5 ± 0.5	0.559	2.9 ± 0.3	0.534
Calculations performed by Moliner and coworkers (unpublished results)						

The reactivity difference between the light and heavy BsDHFR originates from a difference in the calculated recrossing coefficients (Table 2.11). At 5 °C, dynamic recrossing is enhanced in the heavy enzyme, with γ deviated further from unity than that of the light enzyme (γ^{LE} = 0.619 *vs.* γ^{HE} = 0.402). However, this difference in the recrossing coefficients reduces at 25 °C (γ^{LE} = 0.611 *vs.* γ^{HE} = 0.559) and essentially diminishes at 45 °C (γ^{LE} = 0.527 *vs.* γ^{HE} = 0.534). Studies on isotopically labelled EcDHFR-N23PP/S148A have indicated that enhanced femtosecond motions

compromised the catalytic efficiency of the enzyme (118). Accordingly, in BsDHFR combined experimental and computational results indicate that the coupling between fast protein vibrations and the reaction coordinate are enhanced at low temperature. However, the exact cause for this interaction has not been identified.

2.3   Conclusion

^{13}C, ^{15}N, ^{2}H isotopically labelled "heavy" BsDHFR was prepared. The ESI-MS analysis showed a mass increase of 10.93%, corresponding to 99.8%% of the non-exchangeable atoms having been isotopically labelled. Analysis of the tertiary structure via CD spectroscopy showed no structural change at 10, 20 and 35 °C, and no significant difference between the light and heavy isotopologues at these temperatures was observed. This allowed any subsequent conclusions to be based on the mass increase of the protein rather than a structural variation between the two enzymes.

Steady state kinetic analysis of light and heavy BsDHFR showed a large enzyme KIE, particularly when compared to the previously measured EcDHFR equivalent. This can be explained by the greater flexibility of BsDHFR in comparison to its mesophilic counterpart. Thus conformational changes that occur on a global scale in the protein during the catalytic cycle would cause a large enzyme KIE. The Michaelis constants show no effect from the isotope labelling, implying that the enzyme KIE is not due to binding interactions. The temperature independence of the enzyme KIE suggests that a significant loop movement is not rate determining. It is likely that release of DHF or NADP^{+} is rate limiting, as with other DHFRs, but that a large conformational change does not accompany this step.

The chemical step in light BsDHFR was shown to be biphasic, with a hydride KIE that is weakly temperature dependent above 20 °C but becomes strongly temperature dependent below this point. This phenomenon has been seen in other thermophilic enzymes. This was viewed to be as a result of restricted conformational sampling at low temperature, plus the possibility of multiple conformers, the equilibrium of which is switched as a consequence of temperature. SKIE results support a potential

reaction coordinate change at low temperature, or the existence of multiple conformers in equilibrium. The heavy enzyme shows a similar trend in hydride transfer and KIE values, implying that the reaction coordinate is also affected by mass-dependent protein vibrations. Hence, the observed hydride KIE trend is likely affected by numerous factors. Future NMR analysis may divulge evidence, if any, of multiple conformers that could exist across the experimental temperature range.

Preliminary calculations, performed by our collaborators, have highlighted differences in the frequency of recrossing between the two isotopologues. A significant increase in recrossing events is predicted at low temperature in the labelled enzyme, causing the large enzyme KIE[H]. At high temperatures, the recrossing coefficient of light BsDHFR actually deviates further from unity compared to the heavy enzyme, explaining the enzyme KIE[H] trend. As with other labelled enzymes, isotopic labelling of BsDHFR has no effect on tunnelling contribution. Further computational studies are required to fully determine the specific role that these mass-dependent motions play in the case of BsDHFR, particularly during deuteride transfer.

3   PRODUCTION OF LABELLED NADPH

3.1 Preface

The proposed KIE experiments (see Chapter 1) require the preparation of labelled NADPH compounds with heavy isotope atoms incorporated at specific positions of the nicotinamide ring. These compounds can be generated via a synthetic pathway of the isotopically labelled nicotinamide precursor, followed by an enzymatic exchange with thio-NAD$^+$ using NADase (Scheme 3.1) (242-244).
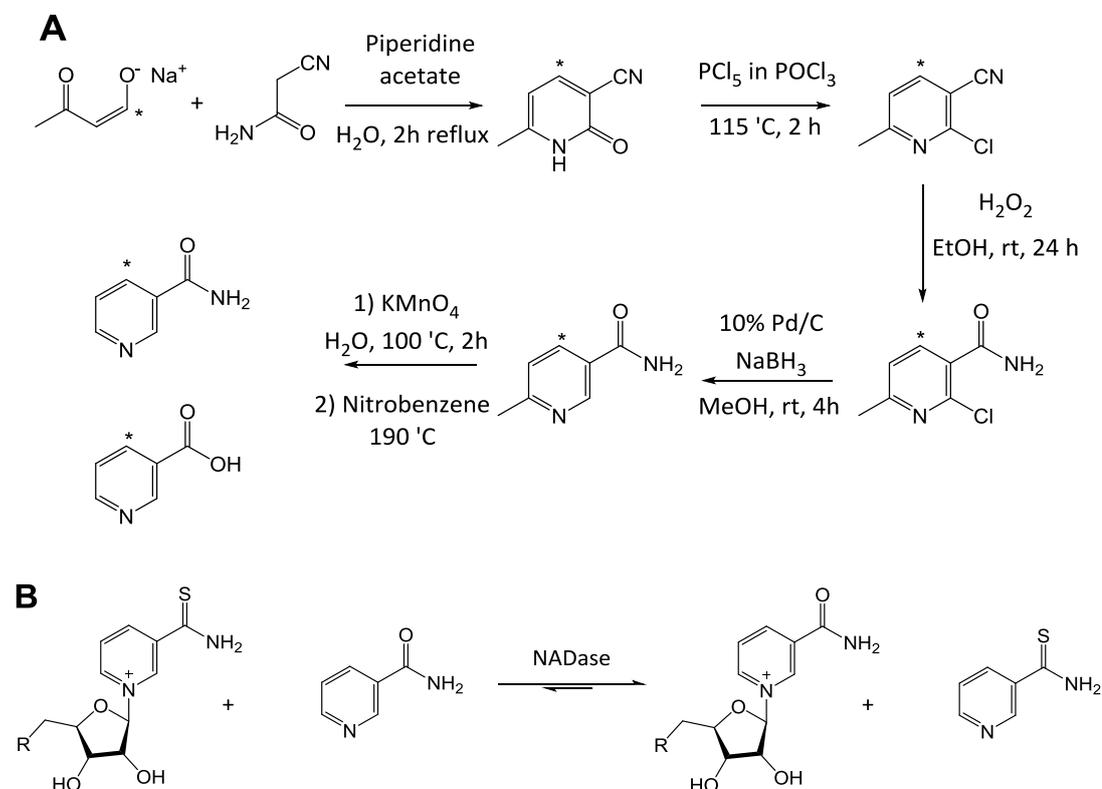
Scheme 3.1 A) Chemical synthesis of [4-^{13}C]-nicotinamide (243). Labelled atom indicated by *. B) Exchange of nicotinamide with thio-NAD$^+$ to give NAD$^+$, catalysed by NADase.

The synthetic methods for labelling the C2, C4, C5 and C6 positions of the nicotinamide ring have been reported (243-247), but these methods are all low yielding, thus are prohibitive when expensive, isotopically labelled starting materials are used. For example, the most recent reported chemo-enzymatic method of [4-^{13}C]-NAD$^+$ synthesis gave a final yield of 0.8%, from a 2%-yield pathway of [4-^{13}C]-nicotinamide (243). Moreover, it is rather difficult to label other positions of the ring due to the symmetry of the chemical reactions involved and

the difficulty in preparing the corresponding labelled starting materials (243). Perhaps most importantly, this synthetic method is rather unpractical, as a large proportion of the labelled starting material is lost. Hence, a method that can produce these isotopologues in high yield, purity and regiospecificity is highly desired.

The biosynthetic pathway of NAD(P)$^+$ is an ideal platform to prepare various isotopically labelled compounds for kinetic and NMR studies of the DHFR reaction (Scheme 3.2). Previously, the pentose phosphate pathway and NAD$^+$ salvage pathways were employed to incorporate heavy isotopes into the ribose and adenosine moieties by feeding ^{3}H and/or ^{14}C-labelled glucose and ATP as starting materials (Figure 3.1) (248). The nicotinamide moiety however has never been labelled, with the exception of commercially available [1-^{15}N]-nicotinamide, because the *de novo* NADP$^+$ biosynthetic pathway was omitted in the reported method (248). Enzymes, including quinolinate synthetase, L-aspartate oxidase, quinolinate phosphoribosyltransferase and NAD kinase, are needed in order to label the nicotinamide ring. Given that all of these enzymes have been biochemically characterised (200, 201, 208, 211, 216, 217, 249, 250), it presented an excellent opportunity to develop a synthetic biology approach to synthesise NADPH containing heavy isotope labels at specific positions in the nicotinamide ring.
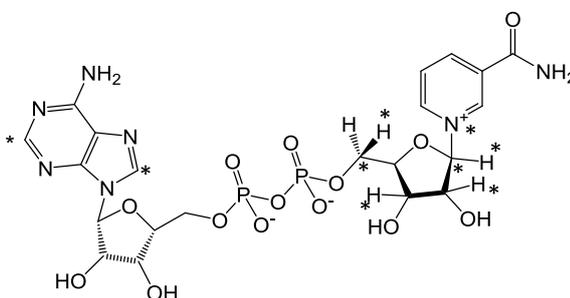
Figure 3.1 Labelled NAD$^+$ produced from (248). Labelled atom indicated by *.
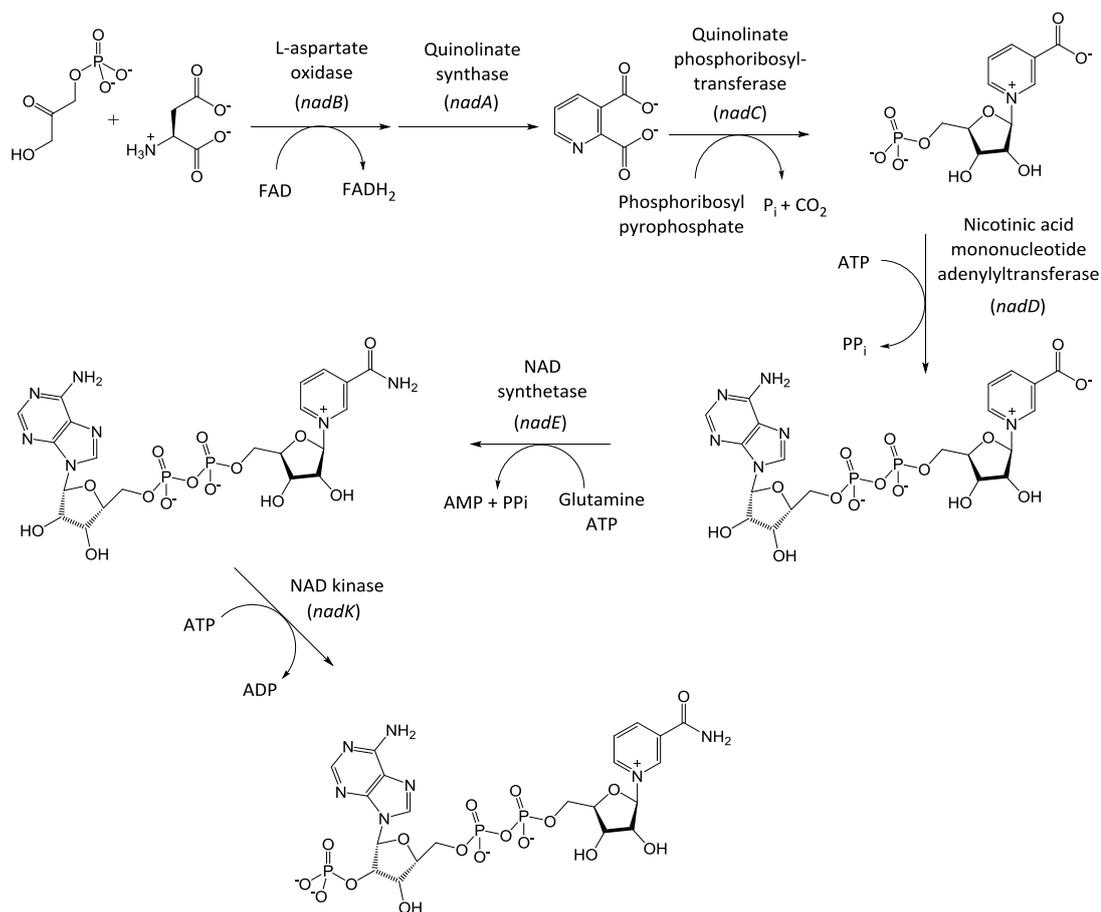
Scheme 3.2 *de novo* **NADP⁺ biosynthesis pathway from** *E. coli*.

3.2 Production and purification of NADP⁺ biosynthesis enzymes

3.2.1 PCR amplification and expression vector production

The six *de novo* NADP⁺ biosynthesis genes involved in the NADP⁺ salvage pathway were amplified from chromosomal *E. coli* DNA by PCR. Primers were designed to be complementary to gene sequences found in the Genbank® database, and in all cases the sequence from *E. coli* K-12 substrains were used due to availability of *E. coli* K-12 substrain XL1-Blue in the laboratory. The primers were designed to be used with Phusion® high-fidelity polymerase, which necessitated primer melting temperatures below 72 °C, and incorporated restriction sites at the beginning and end of the target gene to enable further DNA manipulation (Figure 3.2).

```
5'-catatgccgcctcgccgctataa-3'
        atgccgcctcgccgctataaccctgacacccgacgtgacgagctgctgga
acgcattaatctcgatatccccggcgcggtggcccaggcgctgcgggaagatttag
gcggaacag...//...tgctggataacttcgaaacagaacagatgcgcgaagccg
tcaaacgcaccaacggcaaggcgctactggaagtgtctggcaacgtcactgacaaa
acactgcgtgaatttgccgaaacgggcgtggactttatctccgtcggtgcgctaac
taaacacgtacaagcactcgacctttcaatgcgttttcgctaa
                    3'-gctggaaagttacgcaaaagcgatt**att**cgaa-5'
```

Figure 3.2 The start and end of the *nadC* sequence encoding quinolinic acid phosphoribosyltransferase (black) and the forward and reverse primers designed for PCR amplification (blue). The primers were engineered with restriction sites (underlined) and an extra stop codon at the end of the sequence (bold).

The PCR conditions were optimised for each individual gene by performing six reactions over a 10 °C temperature gradient that spanned the calculated melting temperature. In the case of *nadA*, *nadB*, *nadC*, *nadE*, and *nadK* the recommended conditions for the Phusion® polymerase were successful (Table 3.1), with the exception that initial denaturation was increased to 3 minutes. This was due to the fact that chromosomal DNA from a single colony was used as the template, rather than purified DNA. Agarose gel electrophoresis was used to visualise the DNA after PCR amplification to determine the reaction conditions that produced the highest yield of DNA (Figure 3.3). The *nadD* gene required more reaction cycles, 40 compared to 35, to produce approximately the same yield as the other genes.

Table 3.1 PCR amplification conditions for the NADP⁺ biosynthesis genes.

Step	*nadA*	*nadB, nadC*	*nadD, nadE*	*nadK*
Denaturation	10 s, 98 °C			
Annealing	60 s, 69 °C	60 s, 67 °C	30s, 44.5 °C	30 s, 65 °C
Extension	45 s, 72 °C	90s, 72 °C	45 s, 72 °C	30 s, 72 °C
The denaturation, annealing and elongation step was cycled 35 times, except for *nadD* which was cycled 40 times.				
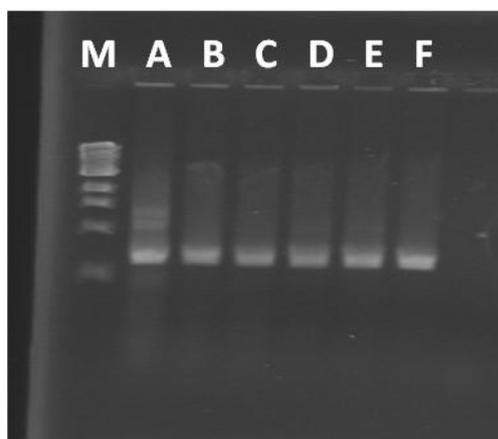
Figure 3.3 Agarose gel showing the PCR product from the amplification of the *nadD* gene. M = marker, A) - F) *nadD* PCR amplification reactions at 40 - 50 °C.

The PCR products were ligated into cloning vectors to avoid the need to repeat the PCR reaction in the future. The StrataClone PCR cloning vector (Agilent) or the pPCR-Script cloning vector (Agilent) were used, as the PCR product could be used directly in the ligation reaction without the need for purification. The ligated vector was transformed into super-competent *E. coli* XL1-Blue cells, enabling plasmid to be acquired and sequenced (performed by Eurofins) to confirm that the correct gene sequence had been amplified during PCR. The cloned genes were then digested from the holding vector using the restriction enzymes that corresponded to the sequences inserted in the primers, and the target expression vector was also digested using the same restriction enzymes. The digests were typically left for 4 hours at 37 °C per enzyme before the cut vector and gene insert were visualised via agarose gel electrophoresis (Figure 3.4) and the digested DNA subsequently

purified. The purified vector and gene insert were ligated together using T4 DNA ligase over a period of 2 hours at 37 °C before transformation into super-competent *E. coli* XL1-Blue cells. In all cases three ligation reactions were performed per gene with varying ratios of vector to insert (1:3, 1:5 and 1:10 vector to gene insert respectively).
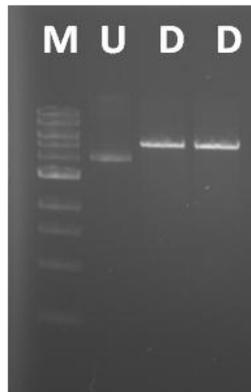
Figure 3.4 Agarose gel of the undigested pET28a(+) vector. M = marker, U = undigested pET-28a vector and D = vector digested with *Nde*I and *Hind*III restriction enzymes

The *nadA*, *nadB*, *nadC*, *nadD*, and *nadE* genes were ligated into the pET-28a expression vector, which contains a T7 promoter for controlling gene expression. *E. coli* BL21 (DE3) strains are an ideal host for expressing these genes, because the chromosome contains the gene for T7 RNA polymerase, whose expression is regulated by the lac operon and thus allows gene expression to be controlled by IPTG. Also, this vector encodes an N-terminal His tag that allows purification of the expression products via nickel affinity chromatography. The expression of *nadK* however is controlled by a T5 promoter following ligation into a pNCO expression vector, therefore *E. coli* strain BL21 (DE3) was not needed. This plasmid also encodes a C-terminal hisactophilin fusion protein that also allows purification via nickel affinity chromatography. To ensure that no mutations had occurred during the ligation reactions the ligated vectors were sequenced again (Eurofins) to confirm the correct gene sequence before enzyme production and purification.

3.2.2 Enzyme production and purification

The expression vectors containing the *nadA, B, C, D* and *E* genes were transformed into *E. coli* BL21 (DE3) and BL21 Star™(DE3) cells for initial test expressions of the target genes. The cells were grown at 37 °C until a OD_{600} 0.6 when IPTG was added to give a final concentration of 0.5 mM. Samples were taken after 4 hours, 5 hours and overnight at 37 °C before analysis via SDS-PAGE, which showed that *E.coli* BL21 Star™(DE3) cells produced greater amounts of protein than *E. coli* BL21(DE3) cells. All the proteins showed clear expression bands after 4 hours, with the exception of *nadE* where the protein band after expression at 37 °C overnight was clearer (Figure 3.5). These successful conditions were scaled up for the large scale production (3 L) of the target enzymes, with the exception of quinolinate synthetase (section 3.2.2.1). *nadK* was cloned into a pNCO expression vector and so *E. coli* JM101 cells were used for protein production. The test expression *E. coli* JM101 cells were left overnight before analysis was performed using SDS-PAGE (Figure 3.5). A clear protein band can be seen on the poly-acrylamide gel, showing that the T5 promoter that controls the expression of the *nadK* gene can be used by native *E. coli* polymerases to produce NAD kinase on a larger scale.
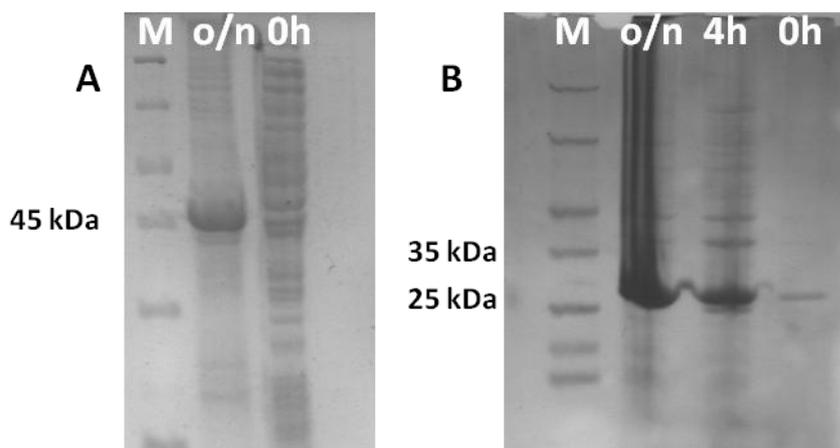
Figure 3.5 A) SDS-PAGE showing the test expression of *nadK* encoding NAD kinase overnight (o/n) compared to the point of induction. B) SDS-PAGE showing the test expression of *nadE* encoding NAD synthetase at the point of induction, after 4 hours and overnight. M = marker, 0h = time of induction, 4h = four hours after induction and o/n = overnight after induction.

As all the *nad* genes were cloned into expression vectors containing a his-tag (N-terminal his-tag and C-terminal hisactophilin fusion protein for pET-28a and pNCO respectively) the enzymes resulting from the over-expression were purified via Ni^{2+}-affinity chromatography. All of the proteins produced were purified under the same conditions, with the exception of quinolinate synthetase (section 3.2.2.1). The pelleted cells were resuspended in 50 mM Tris, 300 mM NaCl and 20 mM imidazole at pH 8, sonicated and clarified via centrifugation before the supernatant was applied to Ni^{2+}-resin equilibrated with the same buffer. The column was washed with the same buffer to remove any proteins that were weakly bound before the target protein was eluted with 50 mM Tris, 50 mM NaCl and 250 mM imidazole at pH 8. The supernatant, flow through and elution were collected and samples were visualised via SDS-PAGE (Figure 3.6). In all cases the purified protein was shown to be >90 % pure, allowing further characterisation of the $NADP^+$ biosynthesis enzymes.
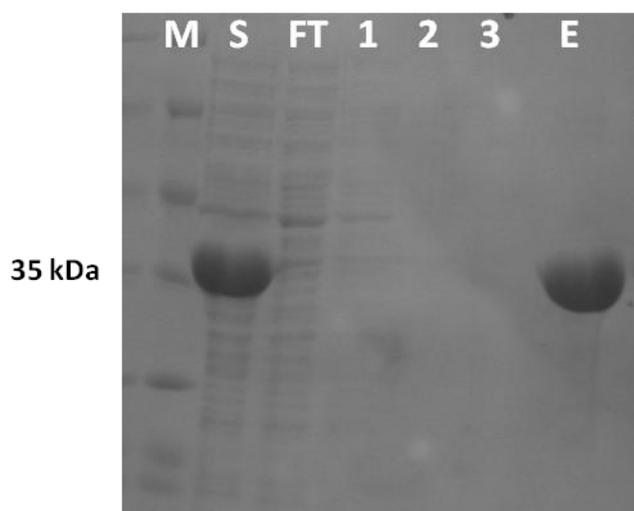
Figure 3.6 SDS-PAGE of NAD synthetase purified via Ni^{2+}-affinity chromatography. M = marker, S = supernatant, FT = flow through, 1, 2 and 3 = washes and E = elution

3.2.2.1 *Quinolinate synthetase production and purification*

Since quinolinate synthetase (QS) contains an oxygen-sensitive iron-sulfur cluster in the active site, and is particularly prone to oxidative degradation, a special expression and purification method was needed. As with the other genes, *nadA* was

expressed in LB medium in BL21 Star™(DE3) cells, but these cells also contained the pB1282 vector. pB1282 encodes part of the *isc* operon responsible for Fe-S cluster formation in *Azotobacter vinelandii* (Figure 3.7), a kind gift from Prof. Squire Booker at Pennsylvania State University (200). Each of these genes encodes a protein that has an important role in Fe-S cluster biosynthesis: *iscS* encodes a cysteine desulfurase, *iscU* and *iscA* are scaffold proteins, hscB and hscA are molecular chaperones and *fdx* a ferrodoxin protein. The *isc* operon components are under the control of an arabinose promoter. The plasmid also confers ampicillin resistance, enabling selection of colonies containing both the pET-28a and pB1282 from agar plates containing both kanamycin and ampicillin. The expression cells were grown until an OD_{600} of 0.3 was reached, when arabinose was added to induce the expression of the Fe-S cluster biosynthesis genes followed by the addition of IPTG and $FeCl_3$ to induce *nadA* expression at OD_{600} 0.6. QS was produced on a 6 L scale, and cells were frozen before enzyme purification.
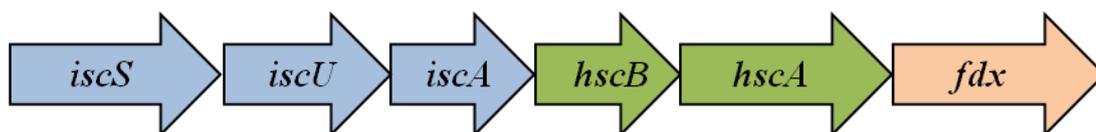
Figure 3.7 A) *Isc* **operon contained in the pB1282 plasmid.**

While the gene was expressed under aerobic conditions, the enzyme needed to be isolated in an environment that had low levels of oxygen. A completely anaerobic purification protocol was performed in the laboratory of Professor N. Scrutton at the Manchester Institute of Biotechnology, using oxygen-free glove boxes for all the purification steps. These glove boxes allowed the enzyme to be purified in an inert N_2 atmosphere, with levels of O_2 lower than 8 ppm at all times, thus the integrity of the [4Fe-4S] cluster remains undisturbed. The purification method was essentially the same as those of other $NADP^+$ biosynthesis enzymes, with the exception that HEPES was used to replace Tris, as the primary amine may interfere with the Schiff

base formation of iminoaspartate (see below). A desalting column was used instead of dialysis to speed up the removal of imidazole from the enzyme solution.

We also developed a purification protocol in Cardiff University, without the use of a glove box. The purified QS was shown to be as active as the one that was purified from a completely anaerobic environment (see below). The key to this success was to perform the purification in a timely manner. Firstly, the enzyme was purified within 60 minutes after the cells were sonicated. A desalting chromatographic method, was also used in this process to remove the need for dialysis. To avoid degradation of the cofactor, the buffers were thoroughly purged with inert gas, by first bubbling N_2 for 4 hours and argon for a further 30 minutes. βME was added to the buffers to maintain a reducing atmosphere during purification (stronger reducing agents, such as DTT and TCEP, cannot be used, as they interfere with the binding between the his-tagged protein and the Ni^{2+}-NTA resin). In both purification protocols, the resulting enzyme solutions were brown, a general characteristic of proteins carrying an intact [4Fe-4S] cofactor (Figure 3.8).
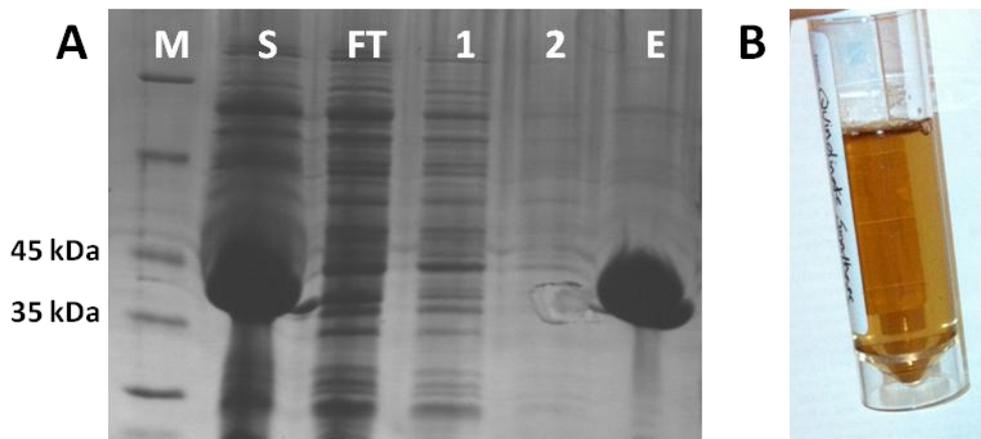
Figure 3.8 Aerobically purified QS. A) SDS-PAGE of quinolinate synthetase purified via Ni^{2+}-affinity chromatography. B) Pure quinolinate synthetase following purification. M = marker, S = supernatant, FT = flow through, 1, 2 and 3 = washes and E = elution

3.3 Characterisation of NADP⁺ biosynthesis enzymes

The NADP^{+} biosynthesis enzymes were characterised by various analytical methods. Whereas SDS-PAGE analysis allows the characterisation of the denatured monomer for each enzyme, size exclusion chromatography was used to verify the native state of the protein. CD spectroscopy was employed to determine if all these proteins were correctly folded in their native states. Additionally, UV-Vis spectroscopy was used to characterise the iron-cluster cofactor of quinolinate synthetase. Activities of these enzymes were assayed by UV-Vis spectroscopy where appropriate.

3.3.1 Size-exclusion chromatography

Size-exclusion chromatography was used to confirm that the native state of each of the seven proteins involved in NADP^{+} biosynthesis was correct. Purified protein was loaded onto a Superdex 200 column and compared to a set of known protein standards (Figure 3.9), allowing each of the enzymes' quaternary structures to be determined. This was done by calculation of the K_{av} value for each protein using the equation below and plotting it against log MW of each standard,

$$K_{av} = \frac{V_e - V_o}{V_c - V_o}$$

where V_o is the void volume of the column, V_e is the elution volume from the column and V_c is the geometric column volume. L-aspartate oxidase and NAMNT were both purified as monomers, while QAPRT was shown to be a dimer in the native form (Figure 3.9). These findings all support previous studies, providing further evidence of the successful production and purification of NADP^{+} biosynthesis enzymes from *E. coli* (208, 211, 216, 250). QS has previously been reported as a mixture of dimer and monomer (201), and in this investigation the monomer was the main form present. Contrary to previous studies, NAD synthetase appears as a monomer rather than a dimer (217, 249). A possible explanation is that insertion of a His-tag at the N-terminal end disrupts the stability of the dimer. Nevertheless, since the enzyme was active in the monomeric form, no attempt was made to remove the his-tag. Likewise, NAD kinase was purified as a

dimer, instead of the hexameric form as previously reported (214), probably because of the insertion of the C-terminal hisactophilin fusion protein. Fortunately, the purified NAD kinase was shown to be capable of catalysing the phosphorylation of NAD^+ to $NADP^+$ successfully (section 3.3.4), thus the fusion protein was not removed from the kinase.
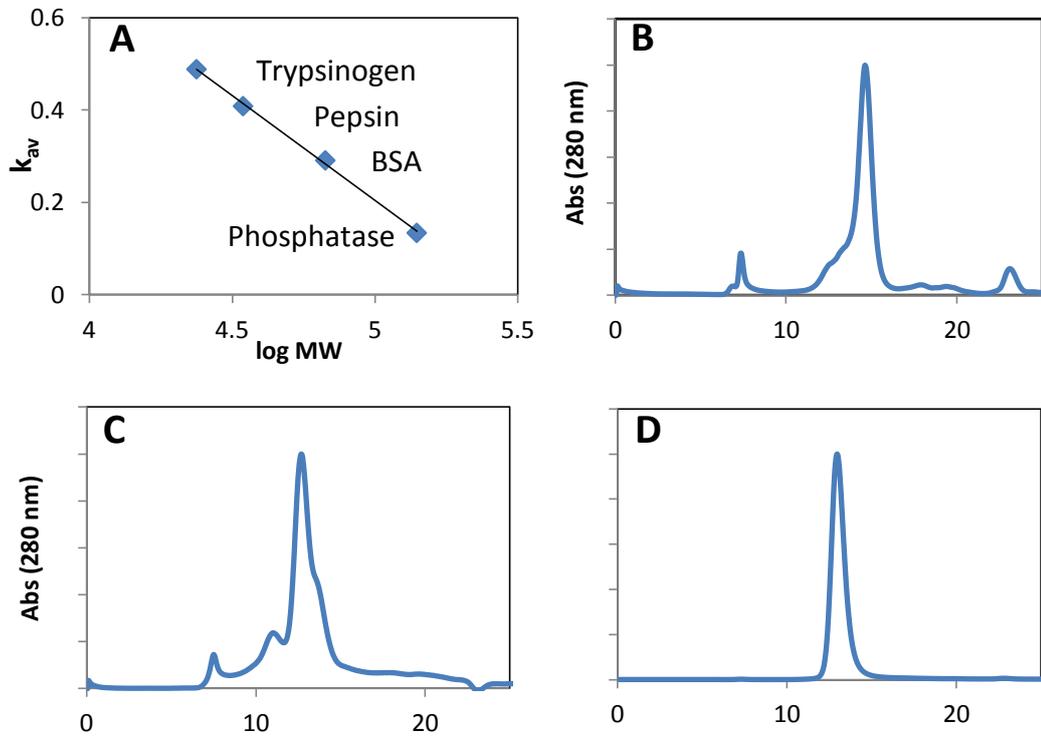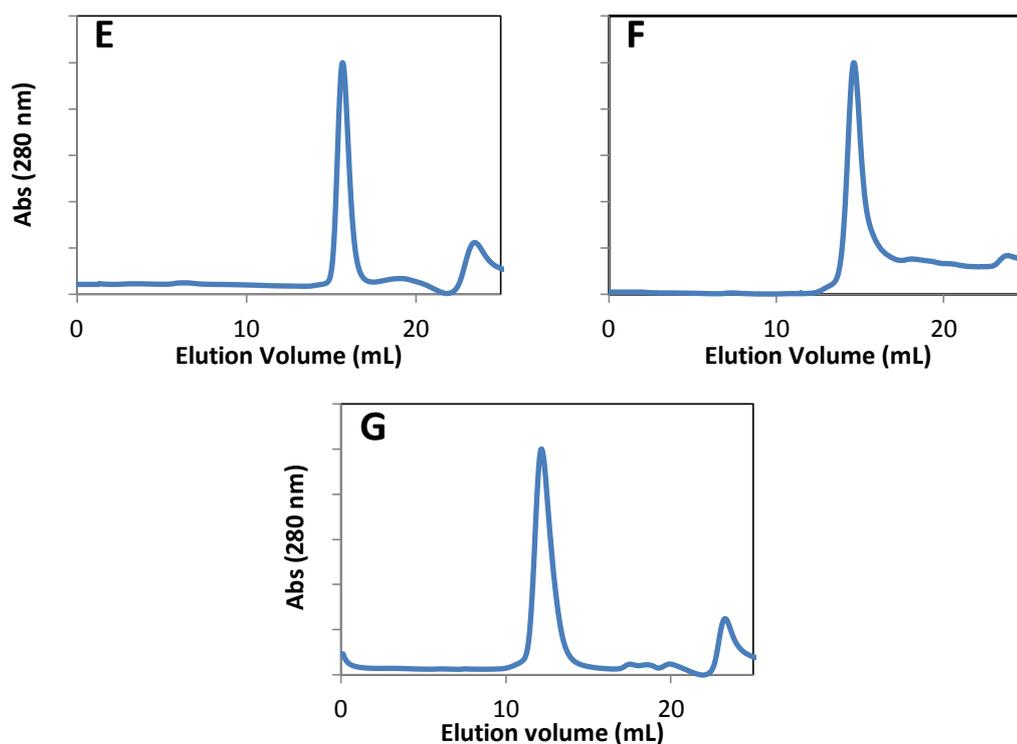
Figure 3.9 A) Standard proteins and the K$_{av}$ values used to calibrate the size-exclusion column. B) - G) Chromatograms of QS (B), LAO (C), QAPRT (D), NAMNT (E), NAD synthetase (F), and NAD kinase (G).

3.3.2 Circular dichroism spectroscopy

CD spectroscopy was performed on each protein to confirm their native folded state. CD spectra were measured from 190 to 400 nm in 10 mM potassium phosphate buffer at pH 7 and 20 °C. Most of the spectra showed characteristics of folded proteins; all of them contained minima at 206 and 222 nm, indicating that the proteins are a mixture of α-helices and β-sheets (Figure 3.10). Unfolded enzymes have very distinct CD signals with signals close to zero in the far UV region, and it is clear that none of the CD spectra obtained show unfolded protein.

Little information was revealed from the CD spectrum of the QS enzyme. Since the [4Fe-4S] cluster absorbs at around 420 nm (see below), CD signal were collected from 190 to 500 nm. However, no signal was seen between 400 and 500 nm . This is most likely due to the weak molar absorption of the cluster (see below) and the low concentration of enzyme required for CD samples. It has also been reported that

magnetic CD and UV-Vis spectroscopy are potentially more powerful tools for characterising specific Fe-S clusters due to variation exhibited among proteins with the same type cluster when using normal CD spectroscopy (251, 252). Moreover, signal around 200 nm is generally unreadable because of the presence of the reducing agent βME, which is needed to prevent the degradation of the iron-sulfur cluster. Thus far, there are no reported CD spectroscopic studies of QS. In order to obtain useful information from CD spectroscopy, the sample preparation and measurement may need to be performed in a glove box to remove the need of βME, potentially allowing a more concentrated sample to be analysed.

For NAD synthetase and NAD kinase, although their quaternary structures are different from those in previous reports (214, 217, 249), the CD spectra of both showed characteristics of a folded protein. This further confirmed that it is unnecessary to remove the purification tags to enable both enzymes to be active.
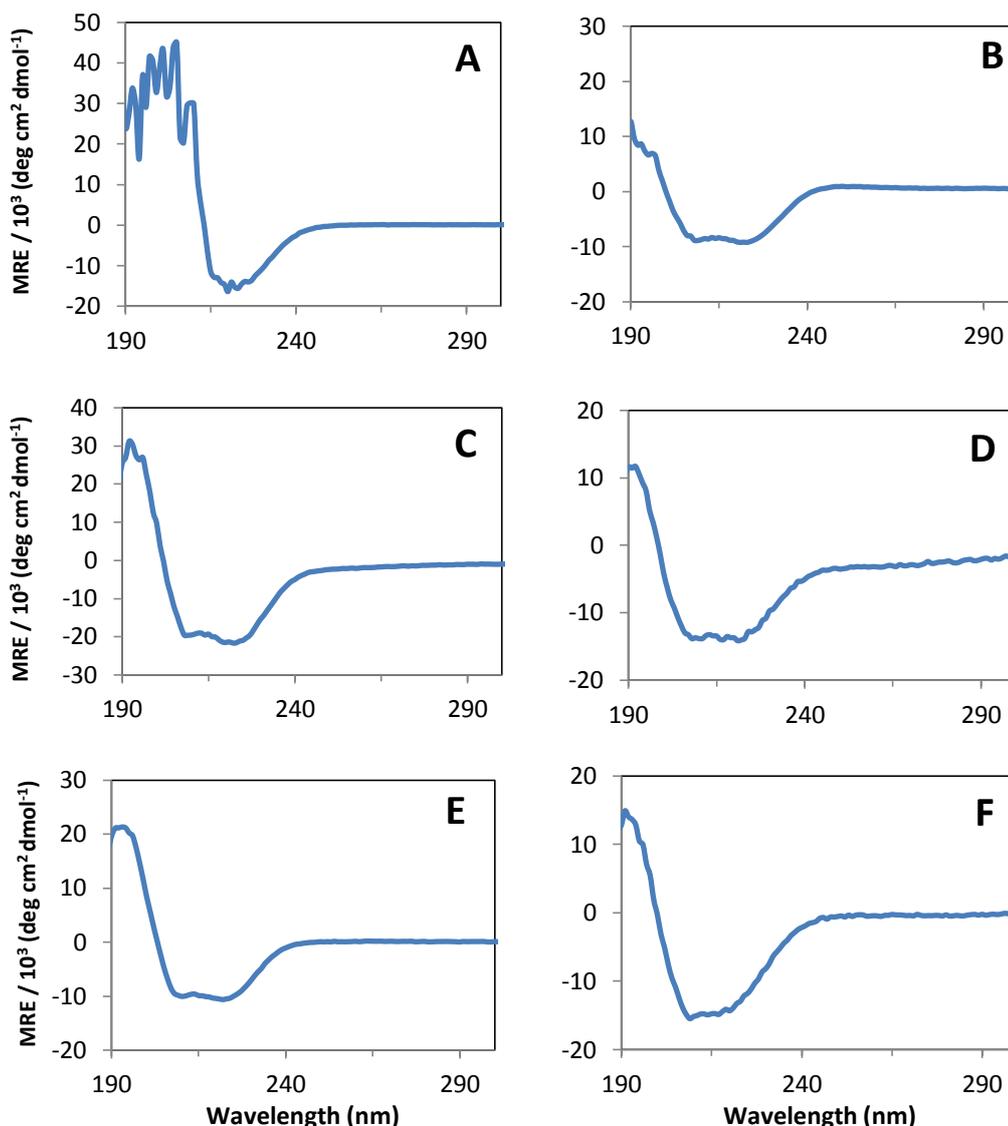
Figure 3.10 CD spectra of the NADP$^+$ biosynthesis enzymes. A) Quinolinate synthetase, B) L-aspartate oxidase, C) quinolinate phosphoribosyl transferase, D) nicotinic acid mononucleotide adenylyltransferase, E) NAD synthetase and F) NAD kinase

3.3.3 UV-Vis spectroscopy

UV-Vis spectroscopy was used to assess the presence of the [4Fe-4S] cluster in QS. Unlike in the CD spectroscopic data, the UV-Vis spectrum indicated the presence of the [4Fe-4S] cluster, with a moderate absorbance peak at 420 nm (Figure 3.11) (252). Since the estimated molar absorptivity (ε_{420} = 4000 M^{-1} cm^{-1}) was relatively weak, this may account for the absence of CD signal, which was derived from the small difference between the absorbance readings of left and right polarised light.
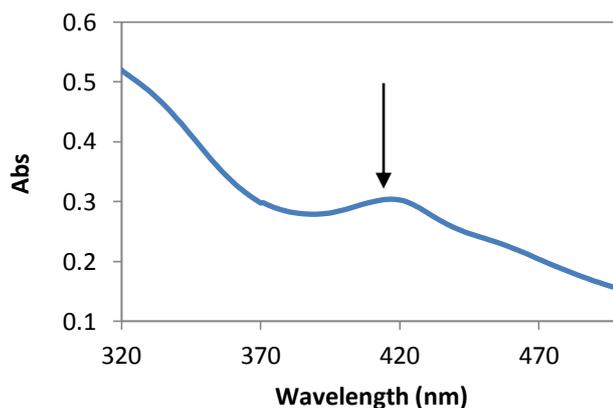
Figure 3.11 UV-Vis spectra of QS showing the absorption at 420 nm of the [4Fe-4S] cluster.

3.3.4 Activity assays

Certain enzymes in the NADP$^+$ biosynthesis pathway can be assayed by UV-Vis spectroscopy. The LAO reaction can be followed directly by monitoring the formation of FADH$_2$, or indirectly by a coupled assay that consumes the product. On the other hand, the NAD kinase reaction can only be followed spectrophotometrically by an enzyme-coupled assay. Other enzymes in the NADP$^+$ biosynthesis pathway were unable to be assayed by UV-Vis spectroscopy, as assays for these reactions are lacking or rely on measuring the release of phosphate moieties, a technique that is to prone to false positives. Hence, their activities were assessed by purifying and characterising the reaction products via NMR spectroscopy and mass spectrometry.

3.3.4.1 *L-aspartate oxidase*

The activity of LAO was monitored either directly or indirectly in an enzyme-coupled assay. In the direct enzyme assay, the reduction of FAD to FADH$_2$ was followed by a decrease in absorbance at 450 nm. There is a clear decrease in signal upon the addition of LAO to a cuvette containing L-aspartate and FAD in 50 mM HEPES at pH 8. However, it was found that oxygen could diffuse into the reaction buffer, despite being thoroughly degassed and purged with inert gases, and recycle

FADH2 to FAD. The signal therefore returned to the starting absorbance over time (Figure 3.12). To circumvent the problem, the reaction was monitored using a coupled assay, which uses hydrogen peroxide to oxidise o-dianisidine by the action of horseradish peroxidase (Scheme 3.3) (197). The oxidised product in the coupled assay can be monitored by an increase in absorbance at 436 nm. When adding LAO to a cuvette containing L-aspartate, FAD, horseradish peroxidase and o-dianisidine, there was clear increase in absorbance at 436 nm (Figure 3.12). These assays together showed that the purified LAO is active.
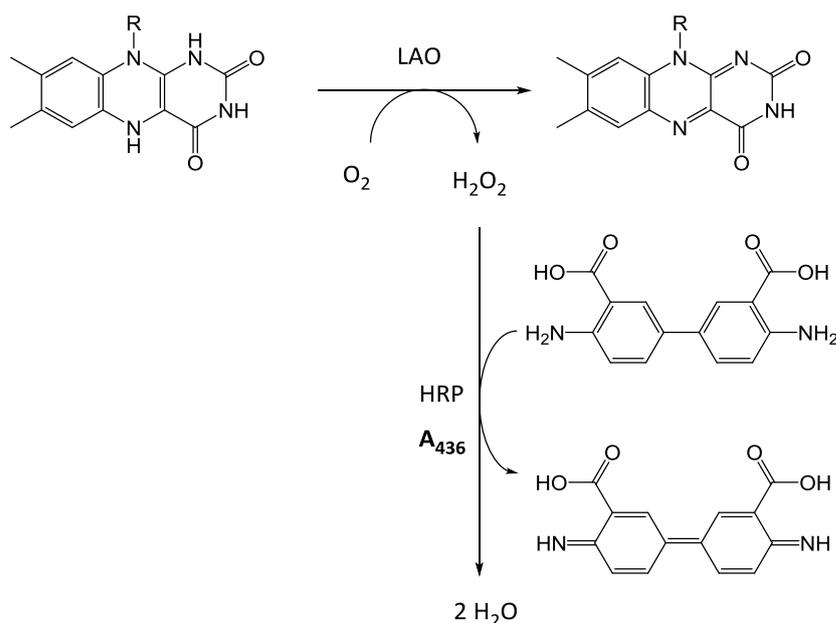
Scheme 3.3 HRP coupled assay to monitor LAO activity.
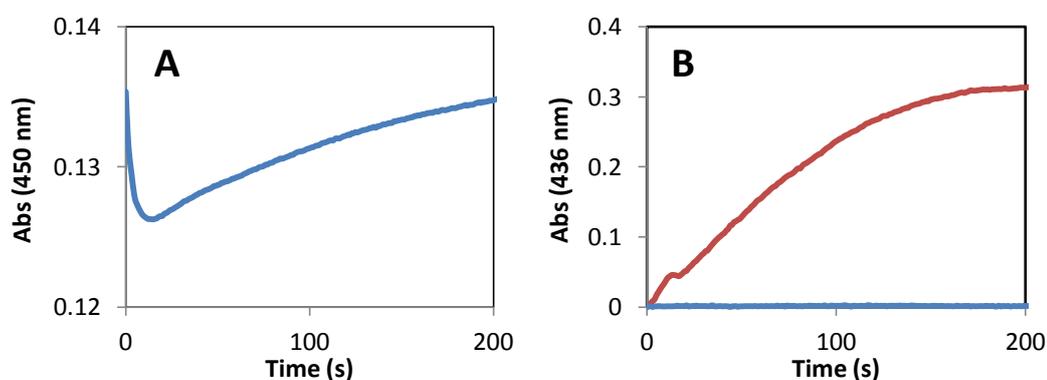
Figure 3.12 A) UV-Vis spectrum of direct LAO activity assay. B) UV-Vis spectrum of the HRP coupled assay when LAO is added (red) compared to no enzyme (blue).

3.3.4.2   NAD kinase

The activity of NAD kinase was monitored by another photometric coupling assay (Scheme 3.4), which converts $NADP^+$ to NADPH by *Thermoanaerobacter brockii* alcohol dehydrogenase (TbADH) (253). In a cuvette containing NAD^+, ATP, TbADH and isopropanol, the enzyme NAD kinase was added and absorbance at 340 nm was monitored by a UV-Vis spectrometer. The production of NADPH was indicated by the increase of absorbance and verified that the purified kinase was active (Figure 3.13)
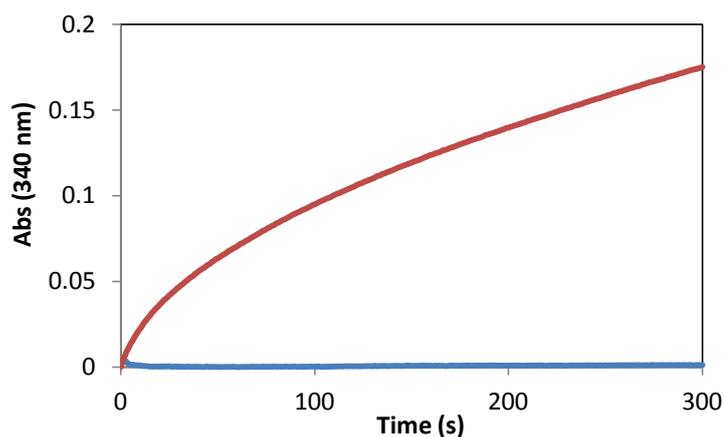
Figure 3.13 Chromatogram showing the conversion of $NADP^+$ to NADPH upon addition of NAD kinase (red), and when no enzyme (blue) is added to the assay.
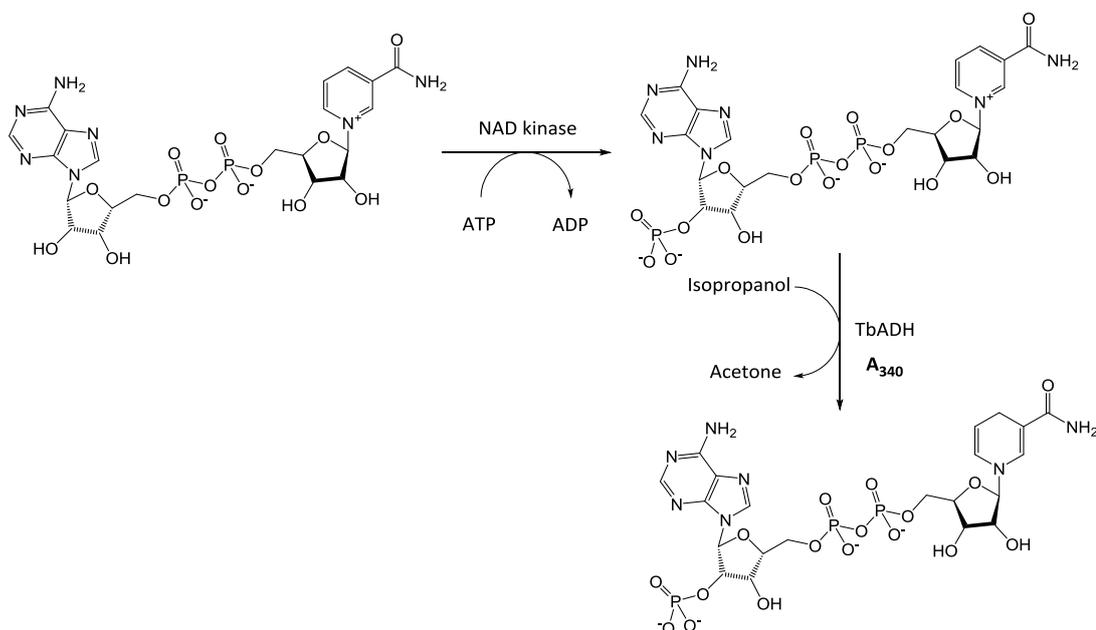
Scheme 3.4 NAD kinase/TbADH coupled assay. The production of NADPH can be monitored at 340 nm.

3.4 Unlabelled NADP⁺ biosynthesis

To verify if the biosynthetic pathway of NADP$^+$ can be reconstituted in the laboratory, the reaction conditions for each of the recombinant enzymes was systematically characterised. This optimised these enzymes for maximal activity. Starting from unlabelled glucose, we managed to obtain NADP$^+$, and subsequently the reduced form NADPH (~2% yield).

3.4.1 Quinolinic acid biosynthesis

In *E. coli* L-aspartate is oxidised by the action of LAO to iminoaspartate, which then condenses with dihydroxyacetone phosphate (DHAP) in the QS reaction to generate quinolinate (QA). In the laboratory the LAO reaction was actually omitted (section 3.3.4, Table 3.2), though it was demonstrated to be active. This is because formation of iminoaspartate can be achieved by condensing oxaloacetate and ammonium chloride in a simple Schiff's base forming reaction (Scheme 3.5) (198). In turn, to obtain DHAP commercially available aldolase and triose phosphate isomerase (TIM) were used, with fructose-1,6-bisphosphate as substrate.
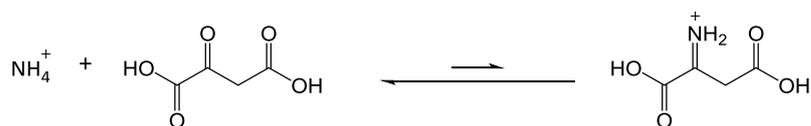
Scheme 3.5 *In situ* production of iminoaspartate from oxaloacetate and ammonia

Initially, much effort was made to optimise the synthesis of QA. Numerous reaction conditions were tested by varying different concentrations of substrate, enzymes, and pH as well as the reaction buffers (Table 3.2). It was only when a relatively large amount of QS (80 µM) was used that the product was able to be purified from the reaction. The yields were determined by comparing the area of the product peak to known concentrations of QA (Figure 3.14). QS purified with or without the use of a glovebox gave essentially identical yield (46% and 48% respectively). Hence, we concluded that this enzyme could be purified without a glove box. As a matter of fact, this finding supports earlier literature, in which aerobic techniques were used to purify active QS (196, 198, 202, 254). Furthermore, production of iminoaspartate by either the LAO reaction or the Schiff's base condensation gave identical yield of QA (45%). This therefore implies that the QS reaction is the limiting factor.

Table 3.2 Conditions tried during the optimisation of QA biosynthesis.

[QS] µM	Anaerobic/ Aerobic	[LAO] µM	[Aspartate] µM	[OAA/NH$_4$Cl] µM	pH	Time	Yield %
5	Aerobic	20		-	8	6 h	-
10	Aerobic	20		-	8	24 h	-
10	Aerobic	-		20	8	24 h	-
25	Anaerobic	-		20	8	48 h	8
50	Anaerobic	-		20	8	48 h	22
80	Anaerobic	-		20	8	48 h	46
80	Aerobic	20		-	8	48 h	38
80	Aerobic	-		20	8	48 h	48

Next, we expanded the biosynthetic pathway of QA by feeding unlabelled glucose as the starting material (Scheme 3.6). Glucose is desired, because the

101

corresponding ^{13}C and/or [2]H-labelled isotopologues are readily available at reasonable prices; thus, isotope labelling of quinolinate at the C4, C5 and C6 positions becomes accessible. Three commercially available enzymes, hexokinase, phosphoglucoisomerase and phosphofructokinase, are required to convert glucose to fructose-1,6-bisphosphate.
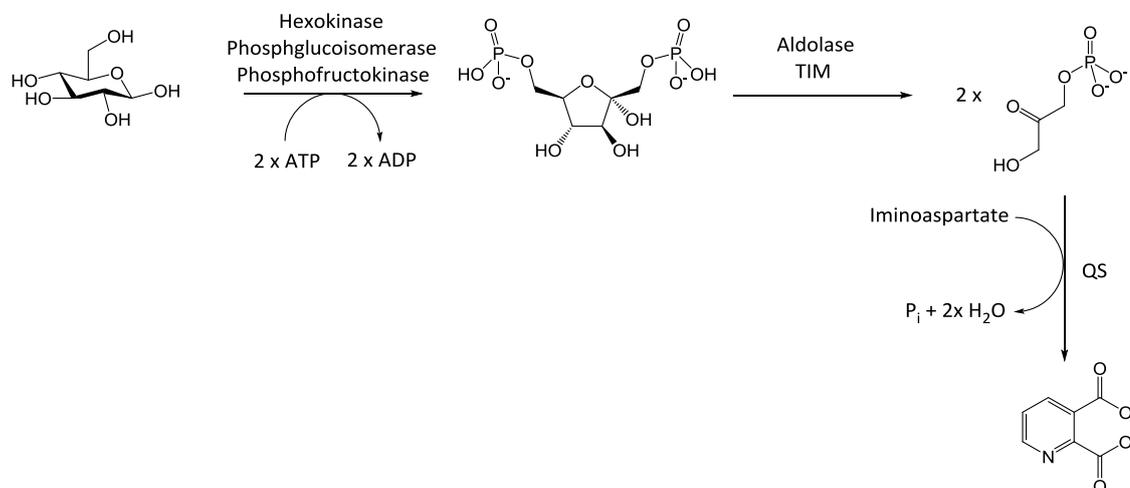
Scheme 3.6 Biosynthesis of QA from glucose using the glycolysis enzymes.

Following an overnight incubation of these enzymes with glucose, QA was generated by adding aldolase, TIM, QS, oxaloacetate and ammonium chloride into the mixture. Although the reaction was successful (Figure 3.14) the additional steps resulted in a decreased yield (5% yield from glucose vs. 45% from fructose-1,6-bisphosphate). The addition of glycolysis enzymes appeared to decrease the reaction yield extensively. For future studies, it might be ideal to obtain these enzymes by recombinant techniques and optimise each of them in order to increase the yield of QA. The purified QA was analysed via ^{1}H-NMR spectroscopy (Figure 3.15), with two doublets corresponding to the C4 and C6 protons at 8.1 and 8.5 ppm respectively, and a triplet (C5 proton, 7.5 ppm) in the aromatic region.
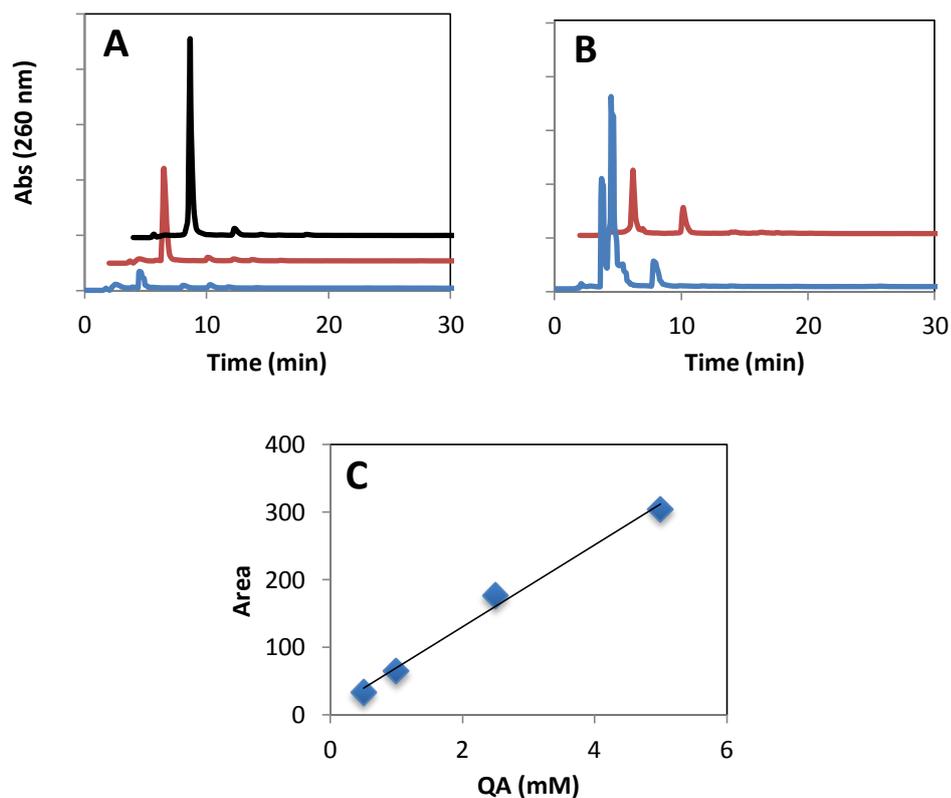
Figure 3.14 A) Chromatogram showing purification of QA from anaerobic QS reactions using enzyme concentrations of 25 μM (blue), 50 μM (red) and 80 μM (black). B) Chromatogram showing purification of QA from aerobic QS reactions starting with glucose (red) and fructose-1,6-bisphosphate (blue). C) Standards used to calculate yield of QS reactions.
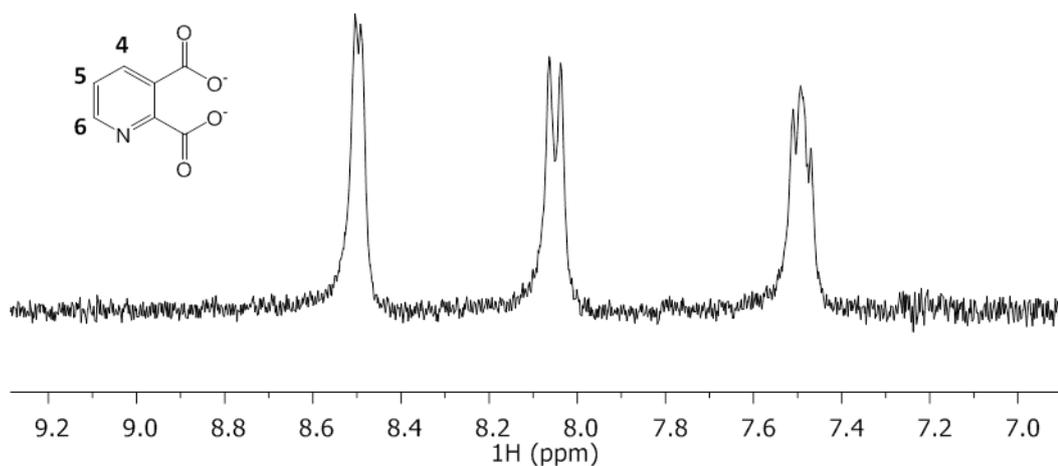
Figure 3.15 ^{1}H-NMR spectrum of QA (300 MHz, D$_2$O)

3.4.2 Nicotinic acid mononucleotide biosynthesis

Once QA could be prepared enzymatically, we focused on introducing this compound into the NADP$^+$ biosynthetic pathway. This can be achieved by reacting phosphoribosyl pyrophosphate (PRPP) with QA in the QAPRT reaction to form nicotinic acid mononucleotide (NAMN). In turn, PRPP can be produced *in situ* from ribose-5-phosphate (R5P) using PRPP synthetase (Scheme 3.7), which was expressed as a his-tag recombinant protein from a vector kindly provided from Professor Jamie Williamson at The Scripps Research Institute (255).
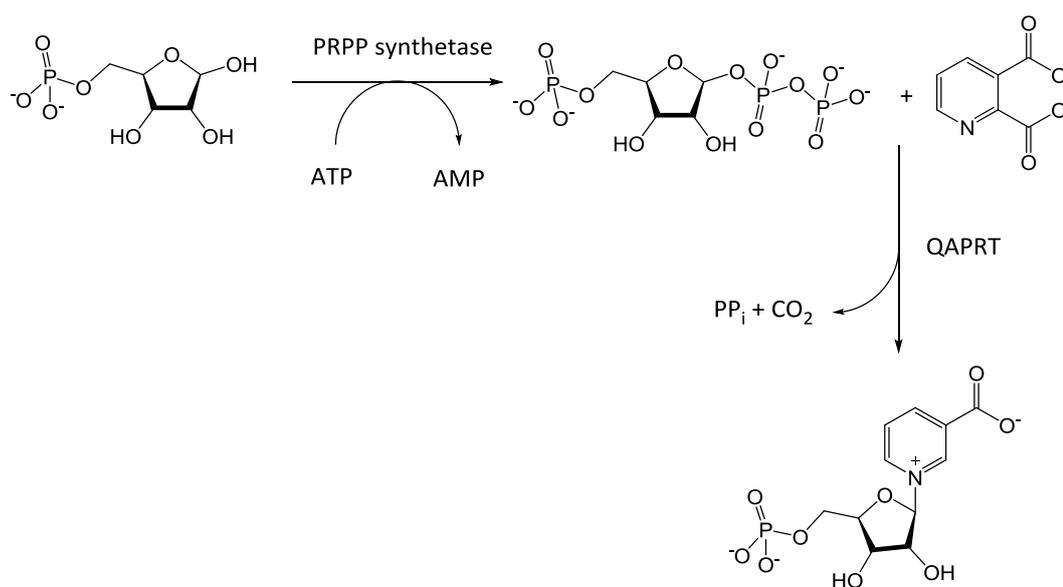
Scheme 3.7 *In situ* **production of PRPP and the subsequent QAPRT catalysed reaction.**

It has been reported that PRPP synthetase requires a high concentration of phosphate buffer to function, therefore the reaction was performed using K_iPO_4 (50 mM, 20 mM $MgCl_2$, 50 mM KCl, pH 7) (256, 257). The reaction was driven to completion by using a threefold excess of R5P and the addition of myokinase, PK and PEP to recycle the ATP. The reaction was performed at 37 °C for 3-4 hours before purification via anion-exchange chromatography (Figure 3.16). Reverse phase HPLC was also performed using a C18 column, but no separation could be achieved. On the other hand, even though NAMN does not bind to the anion-

exchange column, the other reagents, including QA, AMP, ADP and ATP, do and elute at a later time. Hence, anion-exchange chromatography was preferred to 'semi-purify' NAMN for subsequent enzymatic reactions.
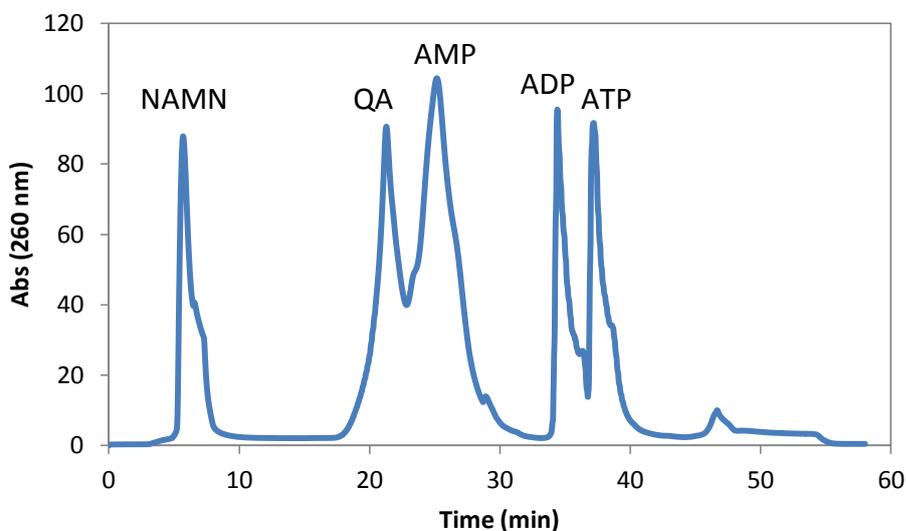
Figure 3.16 Chromatogram showing the purification of NAMN via SAX. The peaks are labelled with their corresponding product/reactant.

NAMN was characterised by ^{1}H-NMR spectroscopy (Figure 3.17). The ribose moiety is indicated by the following signals, including: the distinct doublet at 6 ppm that corresponds to the anomeric proton, a two-proton multiplet at 4.4 ppm that corresponds to 5'-Hs, as well as the one-proton multiplets at 4.3, 4.0 and 3.8 ppm that accounts for the 2'-, 3'- and 4'-Hs. The nicotinic acid ring is evident by the four proton signals found in the aromatic region (8-10 ppm). A combination of two doublets and one triplet are characteristic coupling pattern among the neighbouring protons at the C4, C5 and C6 positions. Most importantly, the singlet signal for the C2 proton at 9.1 ppm confirms the activity of QAPRT, since the starting material needs to be decarboxylated at the corresponding position in order to yield the product. Lastly, ESI-MS was also used to analyse the product peak from the QAPRT reaction. Since NAMN has a positively charged nicotinamide ring, the peak at 335.94 m/z corresponds to the (M+0) state of the compound (Figure 3.18).
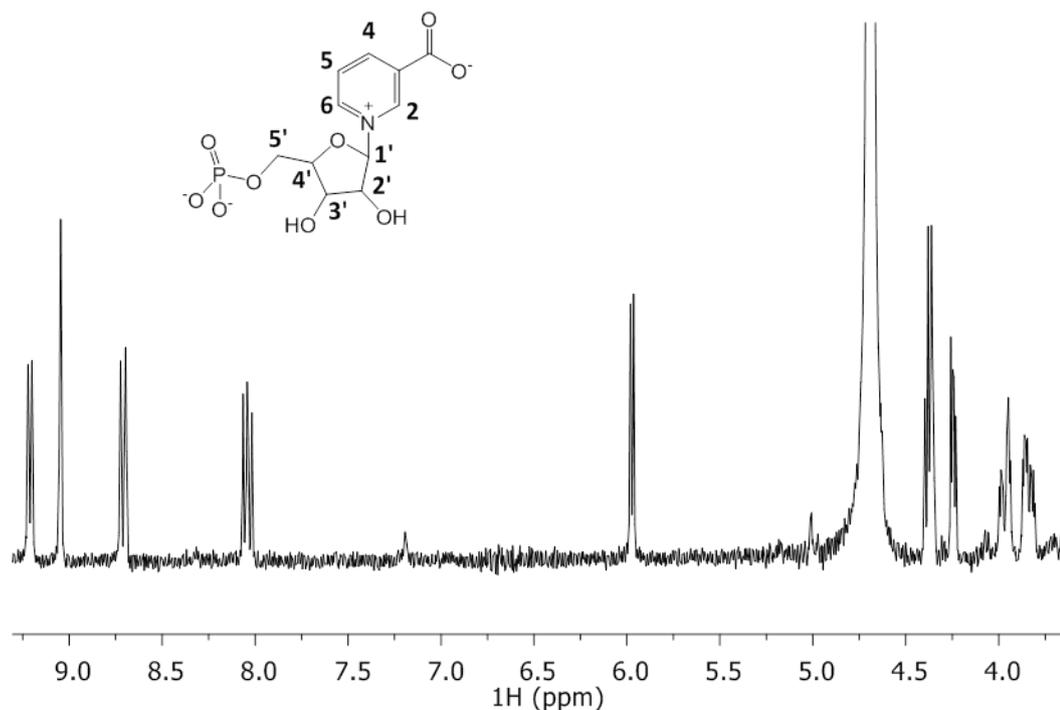
Figure 3.17 ¹H-NMR spectrum of NAMN (300 MHz, D₂O)
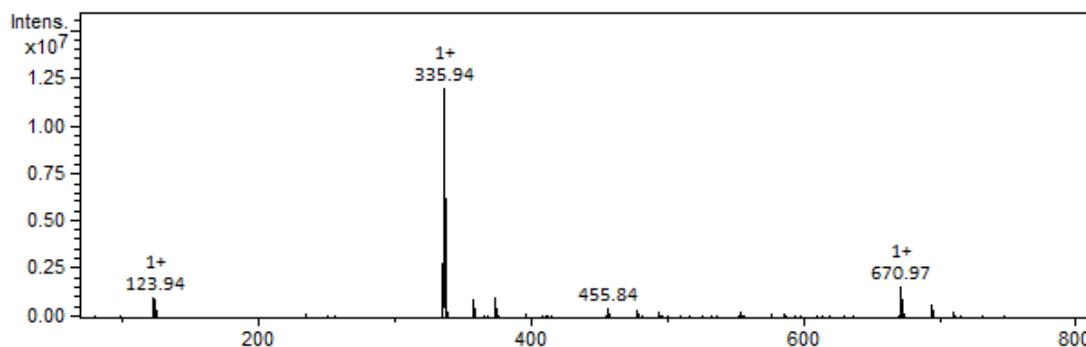
Figure 3.18 ESI-MS of NAMN

3.4.3   Nicotinic acid adenine dinucleotide biosynthesis

The adenosine moiety of NADP⁺ originates from ATP, which is introduced to the biosynthetic pathway under the action of NAMNT. This enzyme converts the intermediate NAMN to nicotinic acid adenine dinucleotide (NAAD), consuming one equivalent of ATP during the reaction.
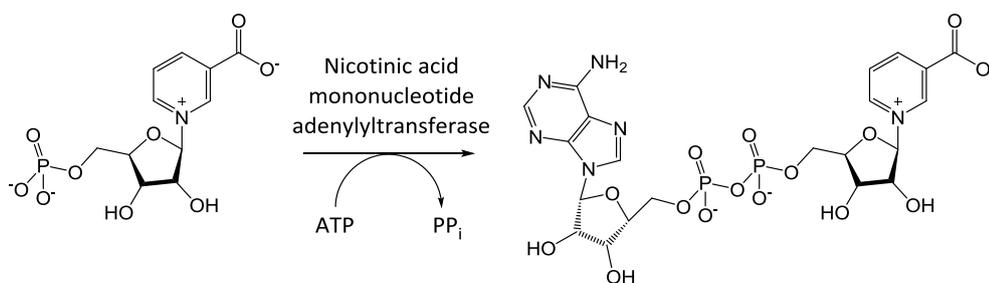
Scheme 3.8 NAMNT catalysed addition of adenosine to NAMN.

Similar to the QAPRT reaction, an excess of ATP was used to drive the reaction towards completion. However, this led to difficulty in purifying NAAD, because it was co-eluted with ATP in the C18-HPLC run. Even though a reasonably shallow gradient was used (0.5% increase of acetonitrile per minute), separations between these compounds still could not be resolved. To solve this problem, the mixture was first semi-purified by anion-exchange chromatography to which ATP binds more tightly than NAAD. The starting material NAMN and NAAD were eluted in the void volume, and they were consequently separated by HPLC (Figure 3.19). The reaction does not seem to reach completion even in the presence of excess ATP. Nevertheless, the starting materials were always recycled for additional adenosine addition reactions.
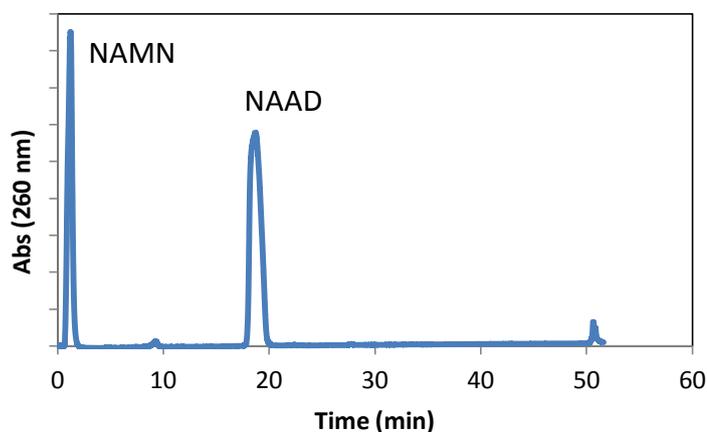
Figure 3.19 Chromatogram showing NAAD purification via HPLC

The ^{1}H-NMR spectrum of NAAD revealed signals that originate from the adenosine moiety (Figure 3.20). These include the additional anomeric proton signal at 6 ppm as well as the 'duplicated' multiplets at 4-4.5 ppm. Moreover, in the aromatic region, there are two new singlet peaks at 8 and 8.3 ppm, which originate from the two protons in the adenine base. In the ESI-MS analysis, the m/z peak of 665.02 matches well with the calculated MW of this compound (MW = 665.10, Figure 3.21).
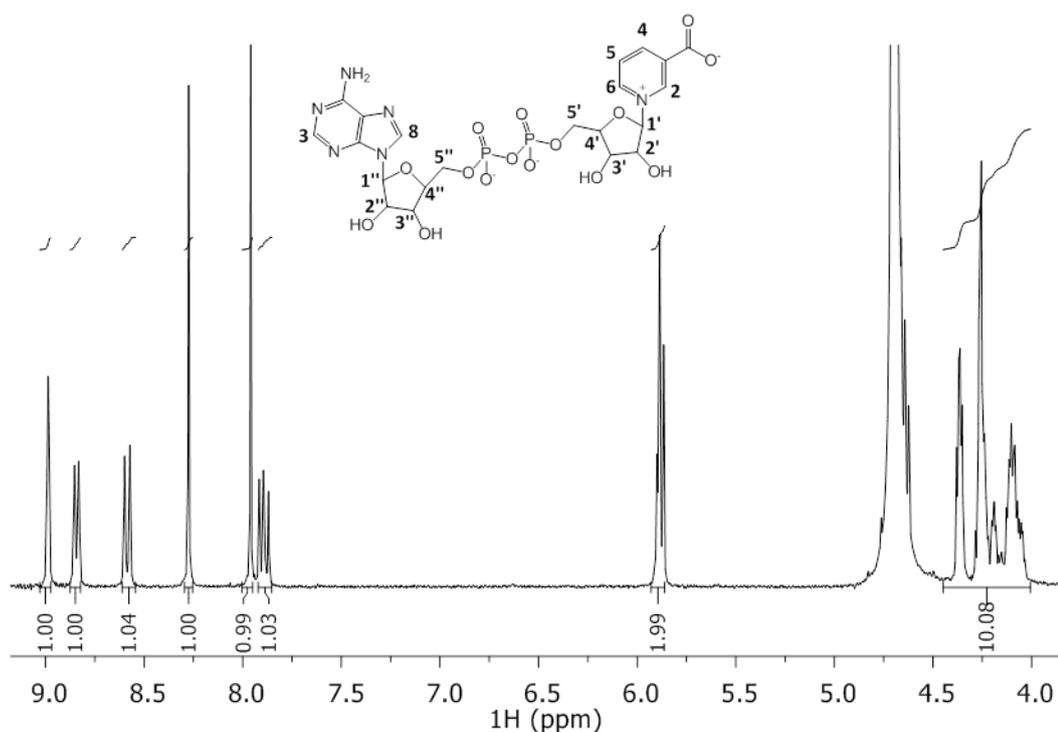
Figure 3.20 ^{1}H-NMR spectra of NAAD (300 MHz, D$_2$O)

An attempt to carry a one-pot conversion from QA to NAAD was made. However, there was no evidence of NAAD being synthesised. In fact, there was no report regarding this one-pot reaction. It is suspected that one or both of the participating enzymes QAPRT and NAMNT were inhibited by an intermediate.
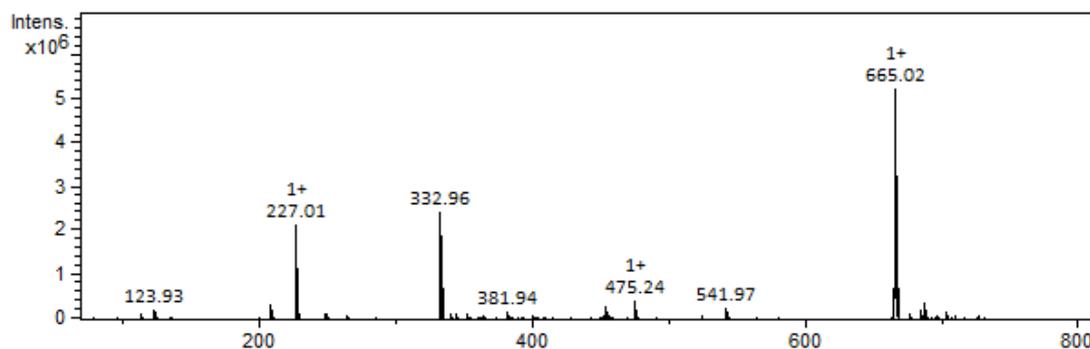
Figure 3.21 ESI-MS of NAAD (MW = 665.10)

3.4.4 NAD⁺ biosynthesis

The subsequent step in the biosynthetic pathway introduces an ammonia group into the nicotinic acid moiety of NAAD to yield NAD$^+$ (Scheme 3.9). Two classes of NAD synthetase exist in nature; both of which require ATP for catalysis, but one class is glutamine-dependent while the other is ammonia-dependent (213). The *E. coli* enzyme is preferred, as it utilises ammonia instead of glutamine as the source of nitrogen (258). Hence, it will facilitate ^{15}N-labelling of the amide group for future experimental use from the significantly cheaper and readily available ^{15}NH$_4$Cl. This *E. coli* enzyme was reported to be inhibited by the previous intermediates of the reaction pathway (248), therefore no attempt was made to carry out the one-pot reaction. In a potassium buffer solution (50 mM, 20 mM MgCl$_2$, 50 mM KCl, pH 7), NAD synthetase was added with 2 equivalents of ATP and ammonium chloride. As shown in the HPLC trace, the peak corresponding to NAAD decreased as the reaction progressed, indicating that the starting material was consumed (Figure 3.22).
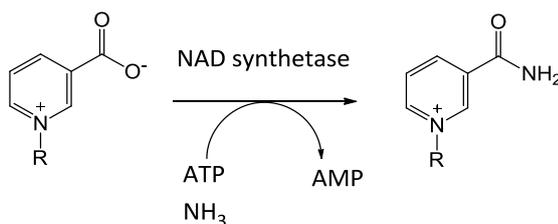
Scheme 3.9 Reaction catalysed by NAD synthetase
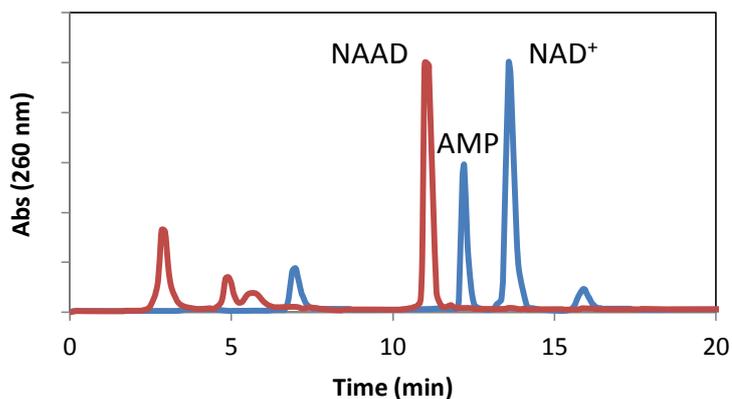
Figure 3.22 Chromatogram showing NAAD (red) and NAD⁺ (blue) purification via HPLC

The ^{1}H-NMR spectrum of NAD$^+$ highly resembles to that of NAAD. This is not surprising because of their structural similarity (Figure 3.23). However, all of the aromatic signals in NAD$^+$ were slightly shifted downfield by 0.25 ppm. This agrees with previous studies on NAD synthetase activity (259). The ESI-MS analysis showed a positive molecular ion peak at 664.12, confirming the introduction of an amide group (calculated MW is 664.12 for NAD$^+$ vs. 665.10 for NAAD, Figure 3.24)
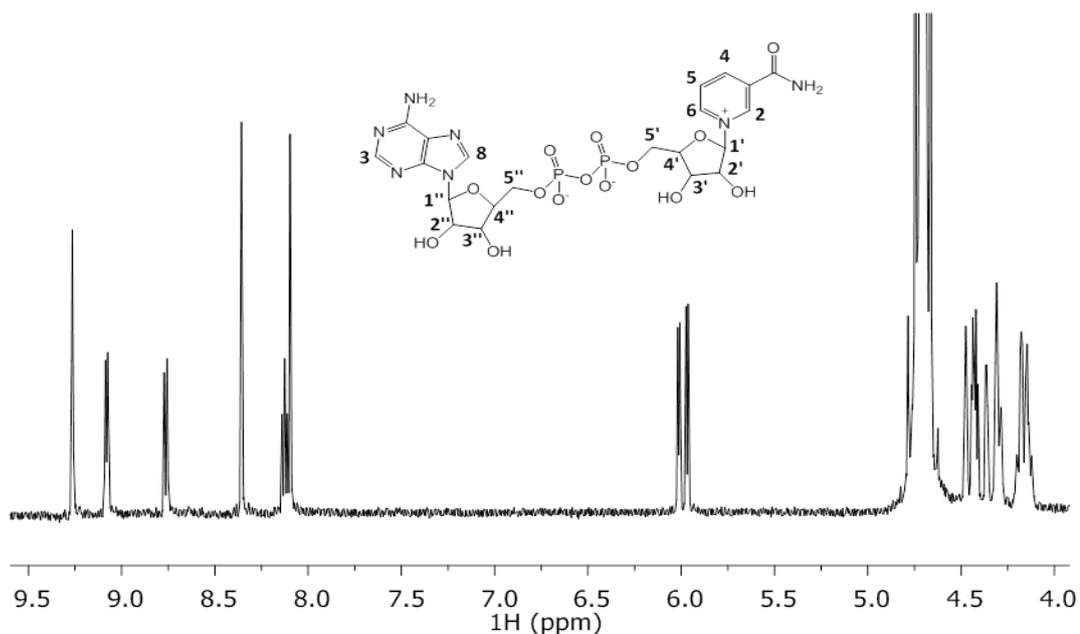
Figure 3.23 ^{1}H-NMR spectrum of NAD$^+$ (500 MHz, D$_2$O)
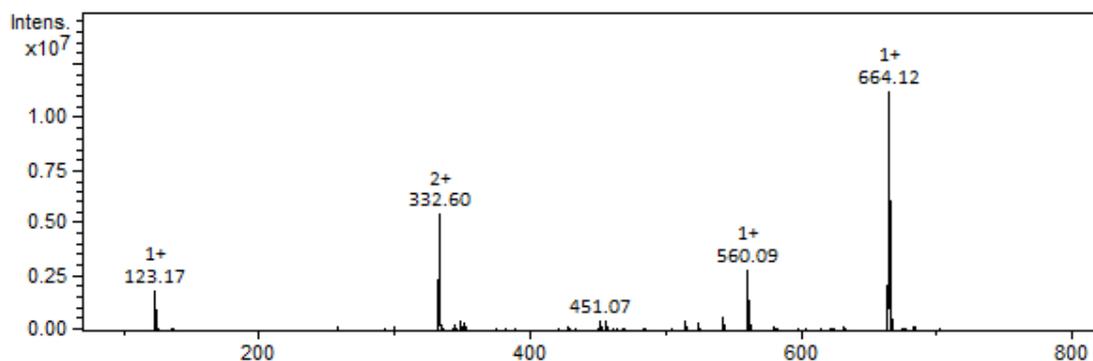
Figure 3.24 ESI-MS of NAD$^+$

Even though our research goal was to synthesise isotopically labelled NADPH, many enzymology models, such as lactate dehydrogenase, employ NADH as cofactor. It is therefore worthwhile to verify if NADH can be made using this biosynthetic pathway, and yeast alcohol dehydrogenase (yADH) was chosen to reduce NAD$^+$ to NADH using ethanol as the hydride source (260). The purified NAD$^+$ was reacted with yADH for 1 hour in 50 mM K$_i$PO$_4$, 50 mM KCl before subjecting it to anion-exchange chromatographic separation. Elution of NADH could be readily identified, as it is the only compound that absorbs strongly at 340 nm (Figure 3.25). Furthermore, the ESI-MS showed a (M-H)$^-$ peak at 664.31 m/z (calculated MW is 665.13, Figure 3.26).
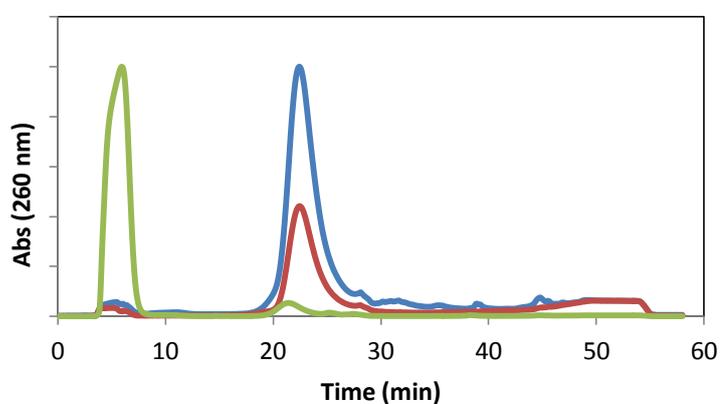
Figure 3.25 Chromatogram showing elution of NAD$^+$ (blue, 260 nm and red, 340 nm) and NADH (green, 260 nm) from the anion-exchange column.
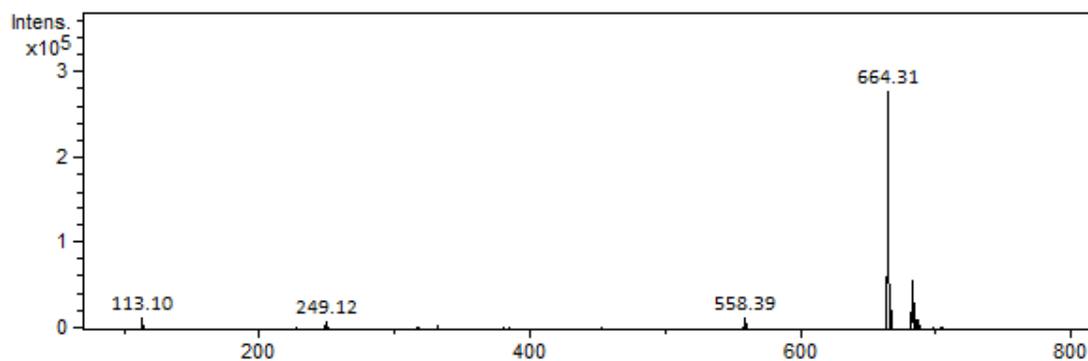
Figure 3.26 ESI-MS of NADH (calculated MW 665.13).

3.4.5 NADP⁺ and NADPH biosynthesis

With the use of NAD kinase, a phosphate group is added onto NAD^+ to yield $NADP^+$. This enzymatic reaction was performed in K_iPO_4 buffer (50 mM, 50 mM KCl, pH 7), supplemented with 20 mM of the cofactor $MgCl_2$ and an excess of ATP (3 equivalents). After 3 hours incubation at 37 $^{\circ}$C, the crude reaction was loaded onto anion-exchange chromatography column, and the product $NADP^+$ was eluted before ADP and ATP. Similarly, the product ^{1}H-NMR spectrum is reminiscent of the NAD^+ and NAAD spectra. The downfield shift of the 2''-H signal (5 ppm) is caused by the introduction of a strongly electron-withdrawing phosphate group.
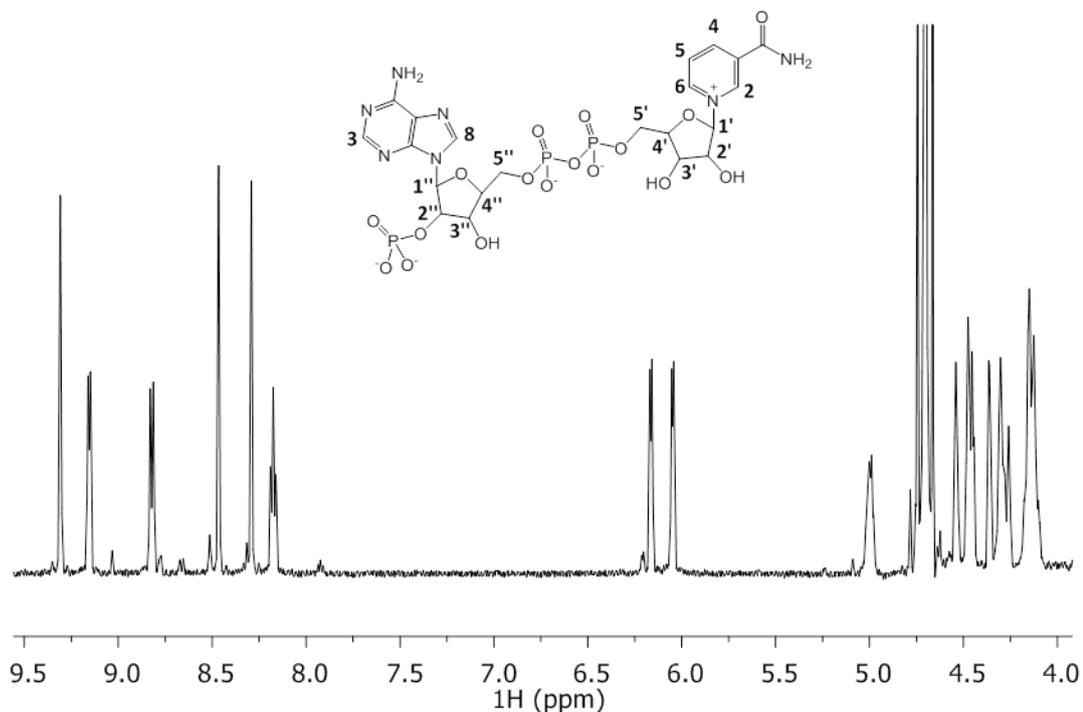
Figure 3.27 ¹H-NMR spectrum of NADP⁺ (500 MHz, D₂O).

Finally, we obtained our final product NADPH (or NADPD) via a hydride transfer reaction that is catalysed by a thermophilic ADH from *Thermoanaerobacter brockii* (TbADH). In this reaction, 2-propanol (or $^2H_8$-2-propanol) serves as the hydride source for reducing NADP⁺ (253). The reaction product was also purified via anion-exchange chromatography and identified by its characteristic absorbance at 340 nm. Because NADPH elutes at the same concentration of salt as ATP does, it is more practical to perform stepwise reactions rather than a one-pot biosynthesis from NAD⁺ to NADPH.

3.5 Labelled NADPH biosynthesis

Since we had successfully synthesised NADPH by feeding simple starting materials into the biosynthetic pathway, we then focused on generating isotopically labelled products. As the DHFR reaction centres at the nicotinamide group of NADPH, isotopic labels will be incorporated into this functionality by synthesising ¹⁵N and

[13]C-labelled quinolinic acid. In this thesis, commercially available [15]NH$_4$Cl and [1,6-[13]C$_2$]-glucose were used for heavy isotope labelling.

3.5.1   [4-^{13}C]-QA and [1-^{15}N]-QA biosynthesis

This first goal was to introduce a ^{13}C label into C4 of QA, at which hydride transfer happens in the DHFR reaction. This can be established by feeding [1-^{13}C]-dihydroxyacetone phosphate ([1-^{13}C]-DHAP) into the biosynthetic pathway of NADP[+]. In turn, this labelled starting material could be made using the glycolysis pathway. Due to the symmetrical production of DHAP in this pathway, both the C1 and C6 position of glucose need to be labelled (Figure 3.28), thus [1,6-[13]C$_2$]-glucose was converted to [4-^{13}C]-QA by the simple one-pot method, as described in Section 3.4.1.
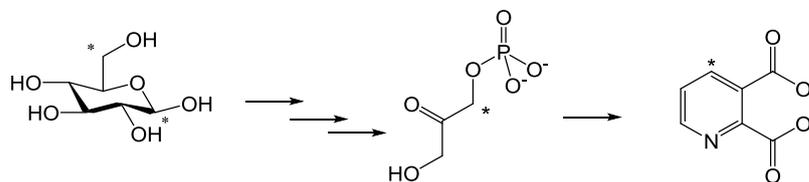
Figure 3.28   [1,6-[13]C$_2$]-glucose and the subsequent labelled [1-^{13}C]-DHAP and [4-^{13}C]-QA. [13]C isotope represented by the *.

Three reactions were performed using separate batches of freshly purified QS to convert a total of 100 mg labelled glucose to QA, and the product from each trial was pooled and purified via HPLC. In the ^{1}H-NMR spectrum of [4-^{13}C]-QA, the doublet corresponding to the C4-H further splits into a doublet of doublets, due to direct coupling to the adjacent ^{13}C atom (Figure 3.30). Also, the ^{13}C-NMR spectrum showed a single peak corresponding to the C4 position at 137 ppm that was verified by comparing it to a standard ^{13}C-spectrum of unlabelled QA (Figure 3.29). There were no other signals originating from QA, indicating that the label was incorporated specifically into the C4 position, although a small impurity was present as indicated by the peak at 185 ppm. The overall yield of [4-^{13}C]-QA synthesis was

1%, slightly lower than that of the unlabelled one (5% yield). It is possible that some of the commercial enzymes had become less active over time, thereby contributing to the lower yield. In the future, freshly prepared recombinant enzymes will be employed to maximise the yield.
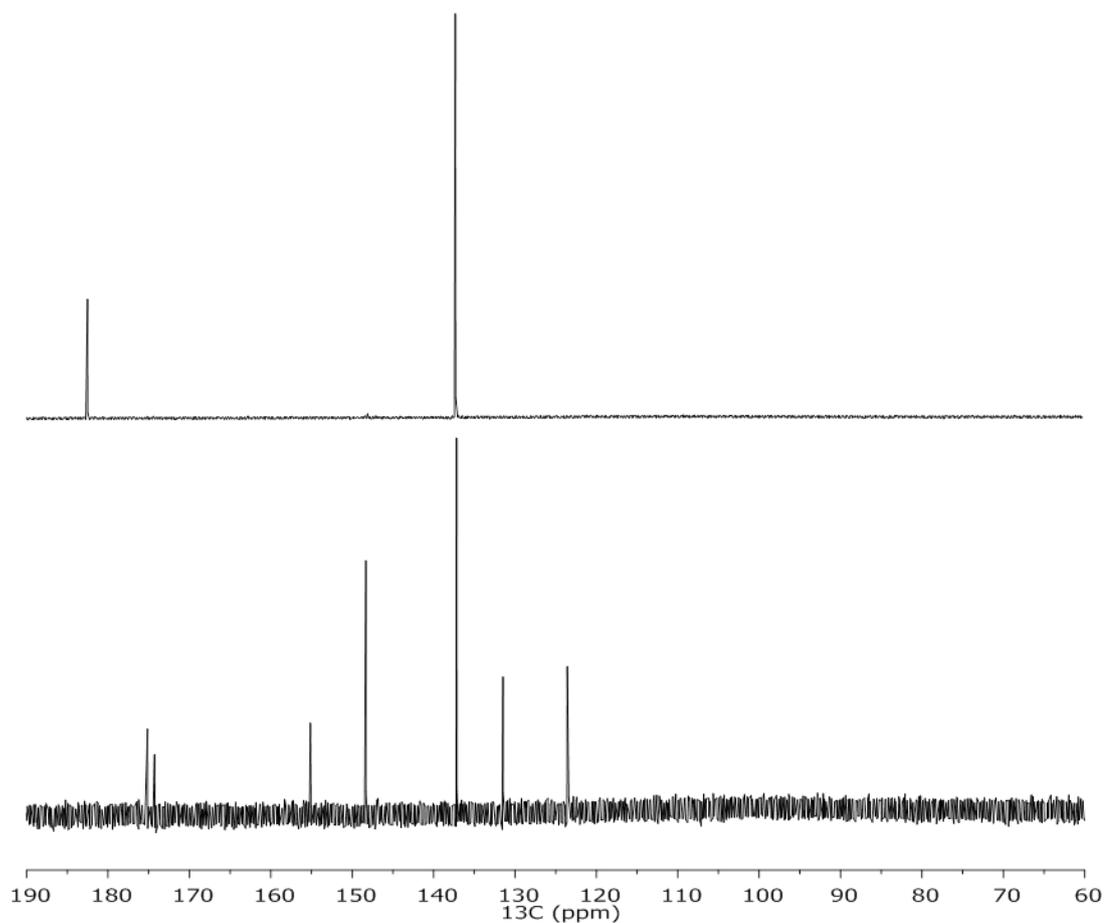
Figure 3.29 ^{13}C-NMR spectra of [4-^{13}C]-QA and unlabelled QA (150 MHz, D$_2$O)
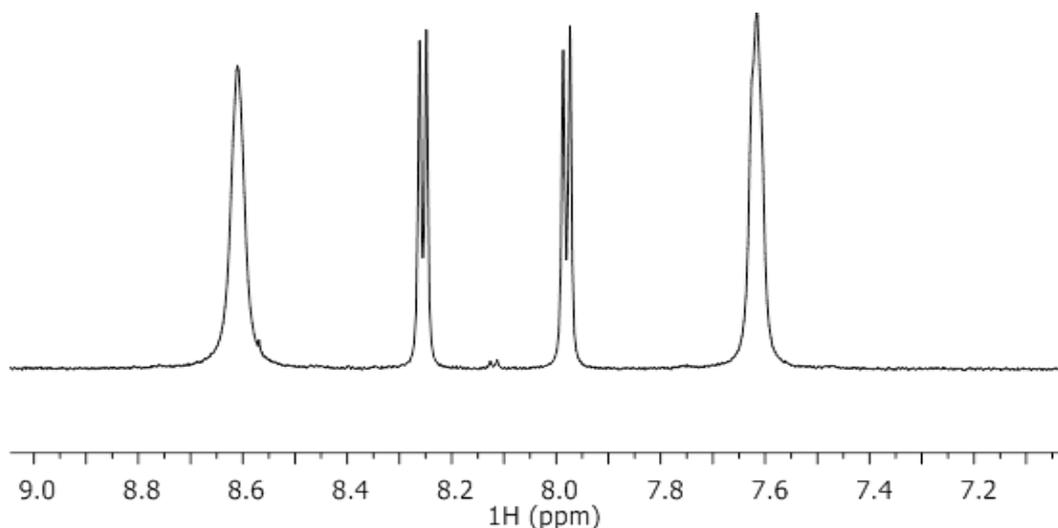
Figure 3.30 ^{1}H-NMR spectrum of [4-^{13}C]-QA (600 MHz, D$_2$O)

Another isotopologue of QA was made by incorporating a ^{15}N label into the aromatic ring. ^{15}N-iminoaspartate was prepared chemically by condensing oxaloacetate with ^{15}NH$_4$Cl, and it is reacted with DHAP in the QS reaction to generate [1-^{15}N]-QA. As the early glycolysis enzymes are not needed to synthesise this isotopologue, most of the starting materials were retained, and a final yield of 36% was achieved after HPLC purification. The ^{1}H-NMR spectrum of ^{15}N-QA showed three peaks in the aromatic region as expected. Furthermore, all of them showed additional splitting, due to the coupling with the ^{15}N atom (Figure 3.31). Likewise, there are two doublets (123 and 148 ppm) in the ^{13}C spectrum, because the C2 and C6 couple to the neighbouring ^{15}N nucleus (Figure 3.32). Lastly, a single sharp peak appears in the ^{15}N-NMR spectrum of the ^{15}N-QA (Figure 3.33) as anticipated. Despite the success in [1-^{15}N]-QA synthesis, [1-^{15}N]-NADPH has not been made yet, as this isotopic label is fairly distant from the reaction centre with regards to the DHFR reaction. Hence, kinetic studies were focused on [4-^{13}C]-NADPH (see Chapter 4). Nevertheless, these experimental results together prove that any position of the quinolinic acid can be labelled with heavy atom by using the *de novo* biosynthetic pathway.
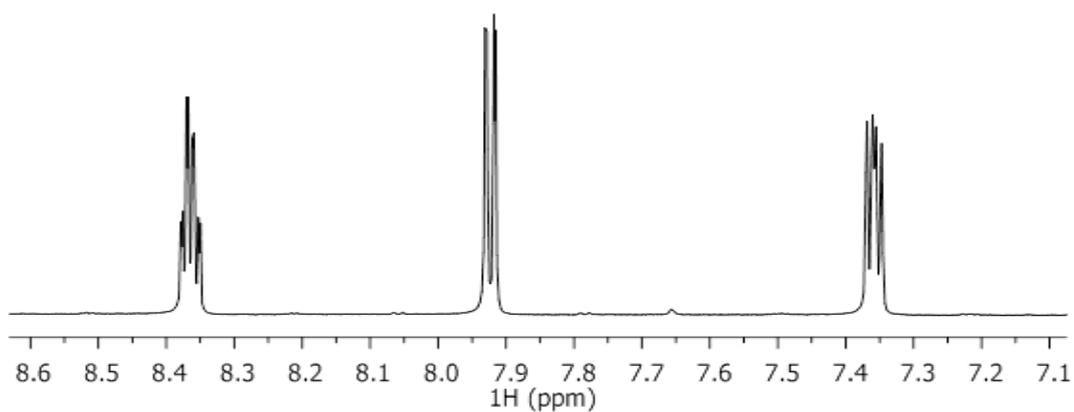
Figure 3.31 ^{1}H-NMR of ^{15}N-QA (600 MHz, D$_2$O)
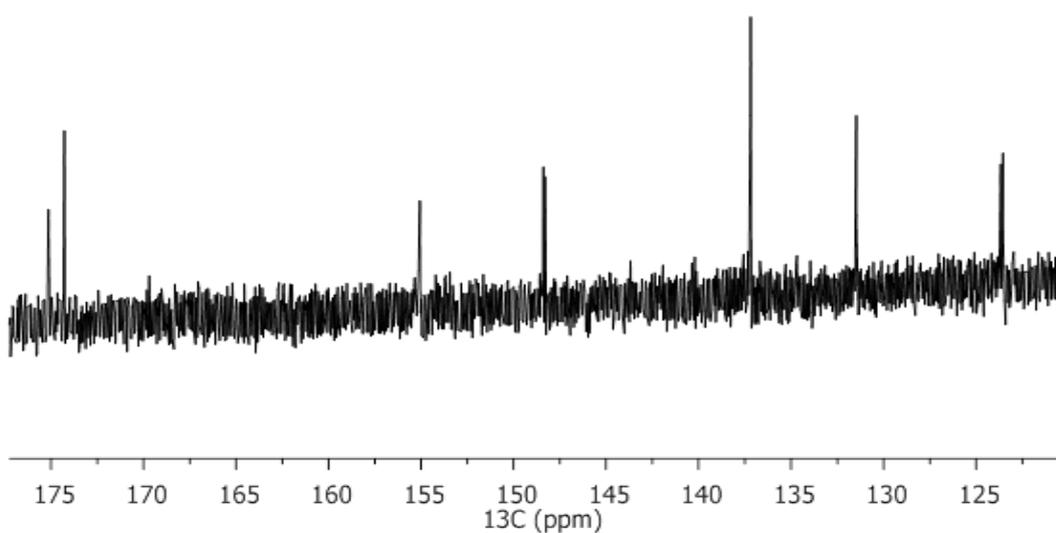
Figure 3.32 ^{13}C-NMR of ^{15}N-QA (150 MHz, D$_2$O)
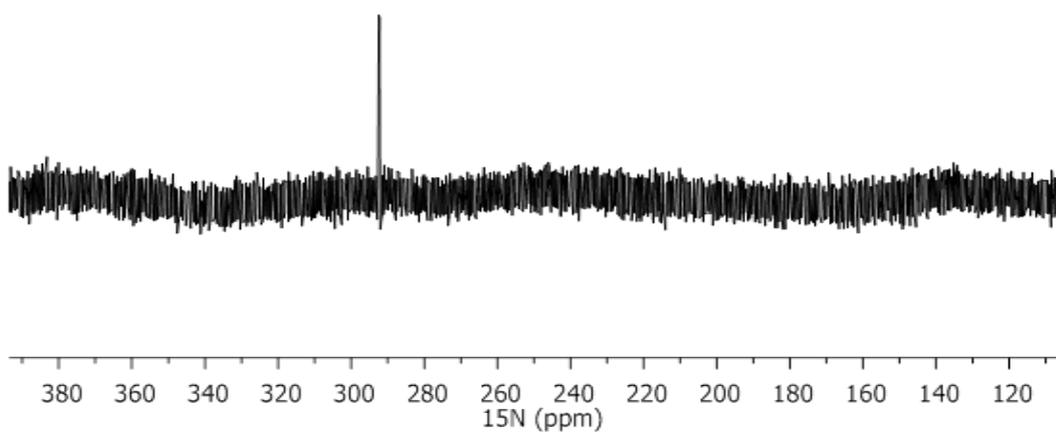
Figure 3.33 ^{15}N-NMR spectrum of ^{15}N-QA (60 MHz, D$_2$O)

3.5.2 [4-¹³C]-NAD⁺ biosynthesis

[4-¹³C]-QA was converted to [4-¹³C]-NAD⁺ using the enzymatic reactions that were tested with the unlabelled reactants earlier (section 3.4). A ribose group was added in the QAPRT reaction to yield [4-¹³C]-NAMN. After anion-exchange chromatography purification, the intermediate transformed into [4-¹³C]-NAAD in the NAMNT reaction. The ¹H-NMR spectrum of the HPLC-purified [4-¹³C]-NAAD showed evidence of direct coupling between the ¹³C and ¹H atoms at the 4-position, giving a measured *J* constant of 160 Hz, whilst other signals remain identical to those of the unlabelled one (Figure 3.34).
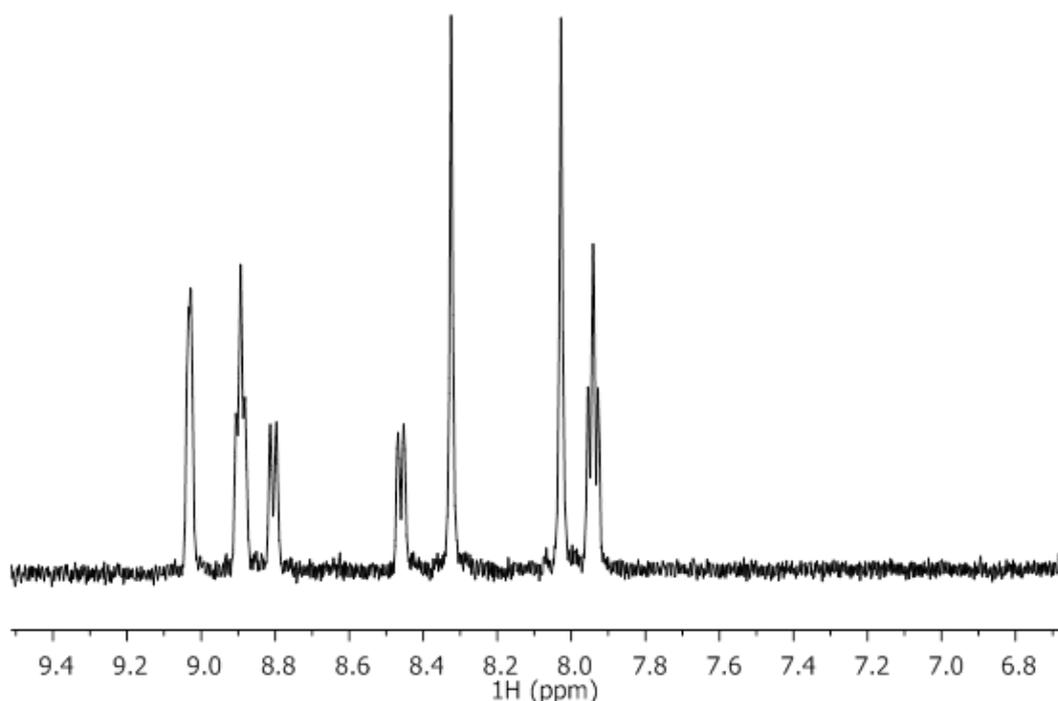
Figure 3.34 ¹H-NMR spectrum of [4-¹³C]-NAAD (500 MHz, D₂O).

Next, [4-¹³C]-NAD⁺ was made in the NAD synthetase reaction and was purified by HPLC. Similarly, the signature ¹³C-¹H coupling was observed in the ¹H-NMR spectrum (*J* = 185 Hz, Figure 3.35). ESI-MS analysis of [4-¹³C]-NAD⁺ showed an molecular ion peak at 665 m/z, showing an increase of one a.m.u. from that of the unlabelled compound (Figure 3.36). These data together strongly supports the conclusion that [4-¹³C]-NAD⁺ was successfully synthesised from [4-¹³C]-QA. To

further confirm its chemical structure, a small amount of the labelled product was incubated with yADH. This enzyme readily catalyses the hydride transfer and generate [4-^{13}C]-NADH, as indicated by absorbance reading at 340 nm. The yield of [4-^{13}C]-NADH was estimated by absorbance reading (ε_{340} = 6220 M^{-1} cm^{-1}), and approximately 41% of [4-^{13}C]-QA was converted to product. This is comparable to other reported yields for enzymatic NAD$^+$ biosynthesis that utilise the NAD$^+$ salvage pathway to produce unlabelled NAD$^+$ (261).
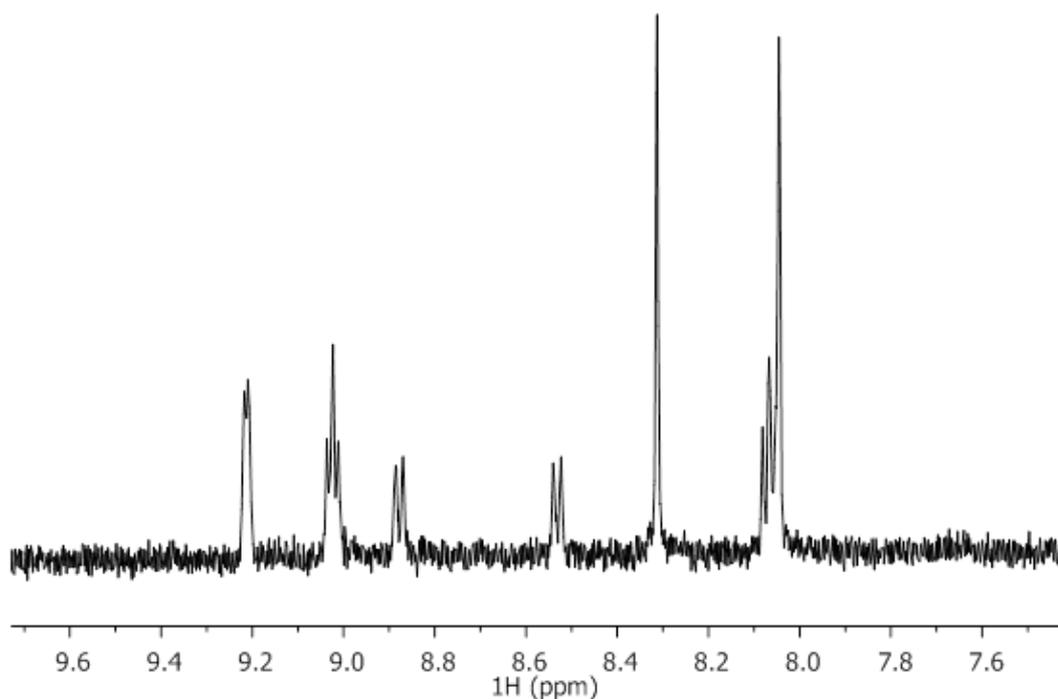
Figure 3.35 ^{1}H-NMR spectrum of [4-^{13}C]-NAD$^+$ (500 MHz, D$_2$O)
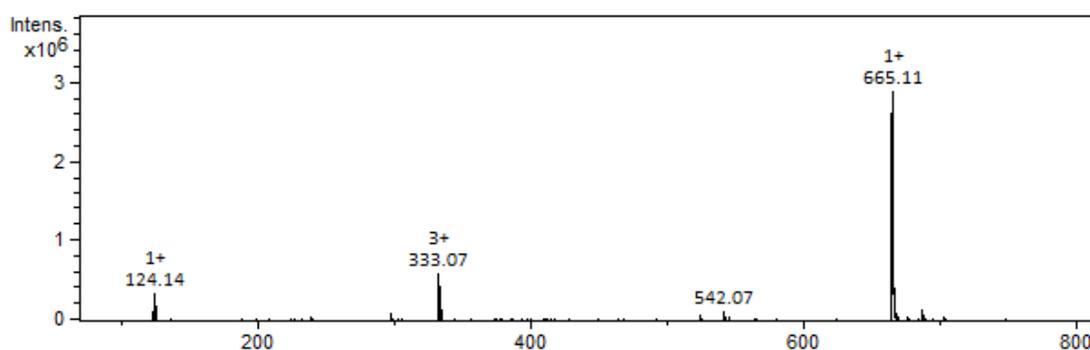
Figure 3.36 ESI-MS of [4-^{13}C]-NAD$^+$

3.5.3   [4-^{13}C]-NADPH and [4-(*R*)-^{2}H,4-^{13}C]-NADPH biosynthesis

[4-^{13}C]-NADPH synthesis was accomplished by first phosphorylating [4-^{13}C]-NAD$^+$ in the NAD kinase reaction and then transferring a hydride/deuteride in the following TbADH reaction. The purified [4-^{13}C]-NAD$^+$ was reacted with excess ATP and the NAD kinase before purification via anion-exchange chromatography. A considerable amount of labelled material was lost (60%) during the phosphorylation reaction. It was not clear the cause of this likely hydrolysis, as it was not observed when unlabelled substrate was used.

Nevertheless, enough product could be isolated for basic NMR and MS analysis. The ^{1}H-NMR spectra shows the characteristic ^{13}C-^{1}H coupling at the 4-position, as well as the downfield shift of 3''-H signal to 5 ppm as a result of phosphate addition (Figure 3.37). The ESI-MS shows a M-2 ion peak of 743 m/z, corresponding to the negatively charged molecular ion (Figure 3.38). As there is one a.m.u. increase from the molecular ion peak of the unlabelled NADP$^+$, it implies that a ^{13}C atom has been successfully incorporated into this compound. Unfortunately, the ^{13}C-NMR spectra only showed one singlet at 160 ppm, which derives from the contaminant ammonium carbonate, and the expected doublet for 4-^{13}C seems to be overshadowed even after multiple scans were collected due to the loss of material during the reaction.
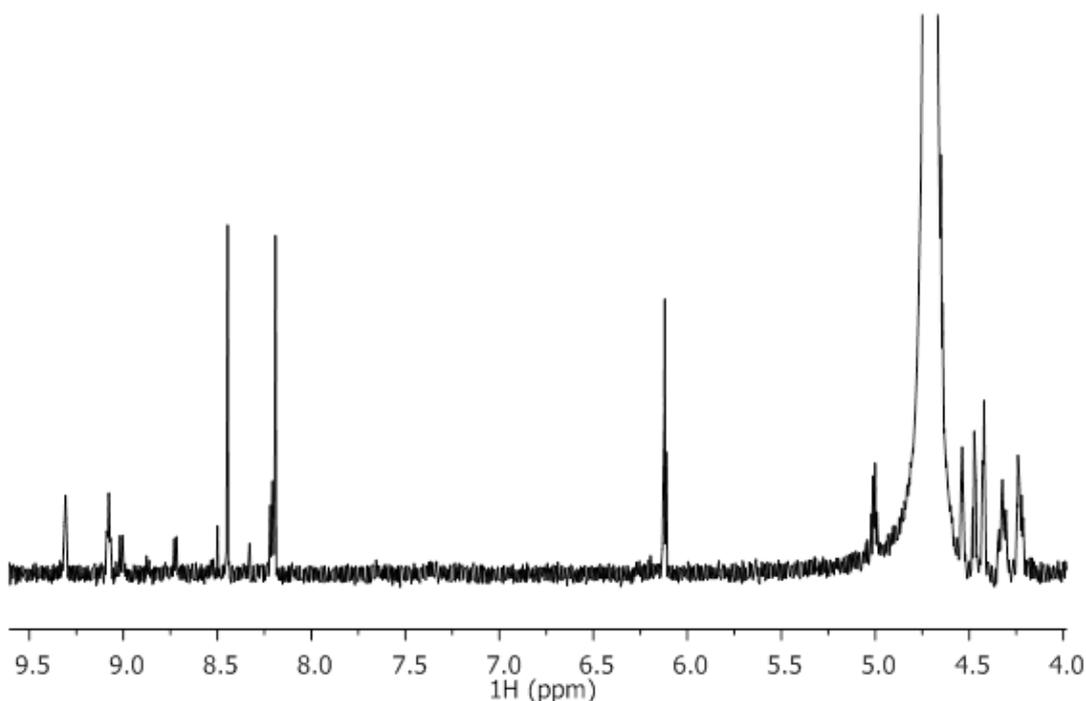
Figure 3.37 ^{1}H-NMR spectrum of [4-^{13}C]-NADP$^+$ (600 MHz, D$_2$O)
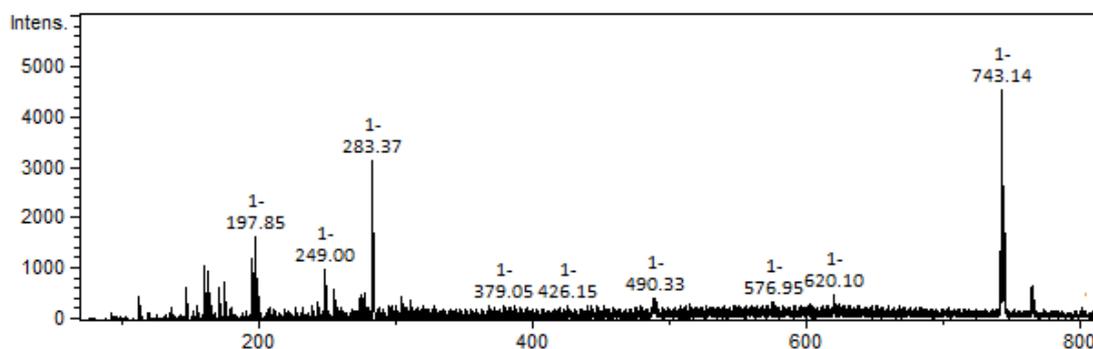
Figure 3.38 ESI-MS spectrum of [4-^{13}C]-NADP$^+$ (MW 745)

To generate useful compounds for DHFR studies, the labelled NADP$^+$ was reduced to [4-^{13}C]-NADPH and [4-(*R*)-^{2}H,4-^{13}C]-NADPH with TbADH serving as the catalyst. Isopropanol or its deuterated isotopologue (^{2}H$_8$-isopropanol) were used as substrate. The reaction was followed by absorbance reading at 340 nm and the products were purified by anion-exchange chromatography. Due to the low yield from the previous NAD kinase reaction, the concentration of the protiated product [4-^{13}C]-NADPH was too low to be characterised by NMR spectroscopy. On the other

hand, a reasonably clear ^{1}H-NMR spectrum of the deuterated product [4-(*R*)-^{2}H,4-^{13}C]-NADPH was collected, and it showed that signals of the oxidised nicotinamide ring has disappeared (Figure 3.39). Instead, evidence for deuteride transfer has become prominent. These include the new singlet at 7 ppm, which corresponds to the C2-H of the reduced nicotinamide ring, as well as the doublet of broad peaks at 2.67 ppm ([13]C-[1]H *J* = 60 Hz), which represents the (*S*)-hydrogen at the C4 position. The hydrogen signal at 2.76 ppm that is found in unlabelled NADPH could not be seen, indicating that the pro-(*R*) hydrogen at the C4 position has been replaced. The ESI-MS also confirmed the reduction of NADP$^+$ by TbADH, as the negative (M-H)$^{-1}$ peak at 745.15 is one a.m.u. higher than the unlabelled analogue (Figure 3.40).
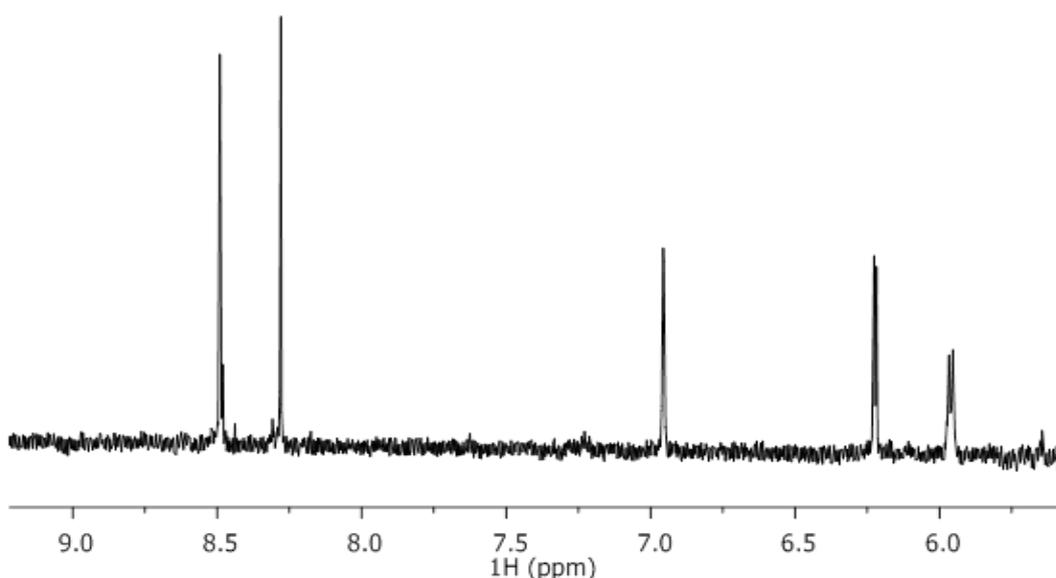
Figure 3.39 ^{1}H-NMR spectrum of [4-^{13}C]-NADPD (600 MHz, D$_2$O)
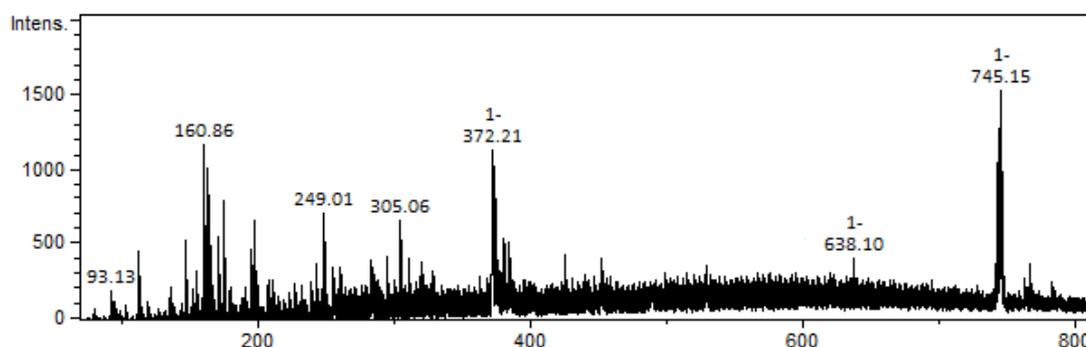
Figure 3.40 ESI-MS of [4-^{13}C]-NADPH (MW = 746.09)

3.5.4 [amino-^{15}N]-NAD$^+$

Compared to the ^{13}C-labelled NADPH, preparation of the [amino-^{15}N]-NAD$^+$ is relatively simple and only requires the use of ^{15}NH$_4$Cl in the NAD synthetase reaction. ^{1}H-NMR spectrum of the HPLC-purified product is identical to that of unlabelled NAD$^+$. However, the ^{15}N-NMR spectrum shows a single peak at 200 ppm, signifying the incorporation of ^{15}N (Figure 3.41, free ^{15}NH$_4$Cl in solution produces a peak at 20 ppm). The ESI-MS shows a molecular ion peak of 665.05, agreeing well with molecular weight of this inherently positive compound (MW 665.10, Figure 3.42). Although this ESI-MS data is extremely similar to that of the starting material NAAD (M+1 = 665 m/z), all of these analytical characterisations together (^{1}H-NMR, ^{15}N-NMR and HPLC analysis) indicate that the enzymatic reaction was successful. Lastly, since yADH was able to transform this compound to [amino-^{15}N]-NADH, it further confirms the chemical structure of [amino-^{15}N]-NAD$^+$.
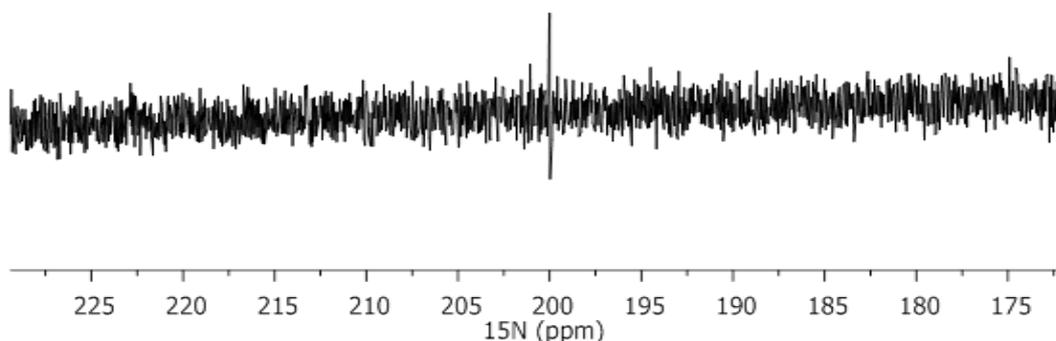
Figure 3.41 ^{15}N-NMR spectrum of [amino-^{15}N]-NAD$^+$ (50 MHz, D$_2$O)
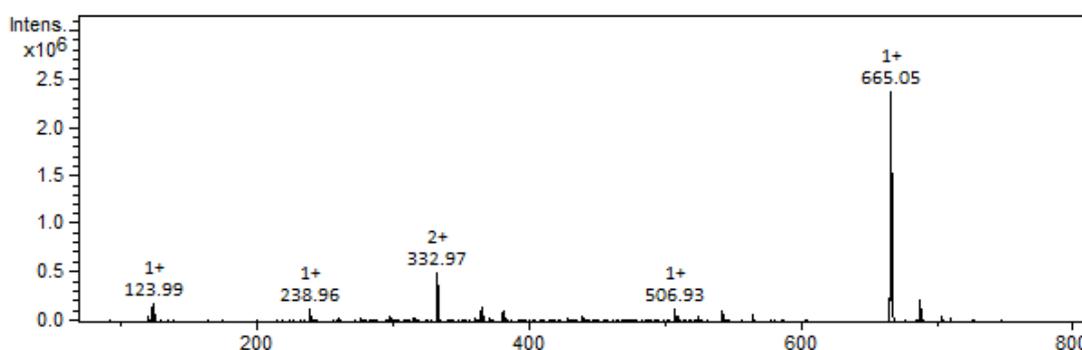
Figure 3.42 ESI-MS of [amino-^{15}N]-NAD$^+$

3.6 Conclusions

In summary, the *E. coli de novo* NADP$^+$ biosynthetic pathway has been reconstituted in the laboratory. The six *E. coli* enzymes involved in NADP$^+$ biosynthesis were successfully cloned and purified, and together they could synthesise the desired isotopically labelled cofactors. This allows a one-pot synthesis of selectively labelled quinolinic acid. This is exemplified by the production of ^{15}N-QA and [4-^{13}C]-QA. As demonstrated in this chapter, [4-^{13}C]- and [amino-^{15}N]-NAD$^+$ as well as [4-^{13}C]-NADP$^+$ were prepared accordingly. To the best of our knowledge, this is the first use of the *de novo* biosynthetic pathway to synthesise isotopically labelled NAD$^+$ and NADP$^+$. Commercially available ^{13}C-aspartate (instead of ^{13}C-glucose) would enable the labelling of the C2, C3 and the carbonyl group positions around the ring in the future using this method developed here. This finding is particularly exciting, as this work indicates that a large variety of NADPH isotopologues can be made, with heavy atom labelled at any desired location. In turn, these compounds have numerous applications in enzymology studies, most notably for KIE and NMR studies.

4   ANALYSIS OF THE DHFR REACTION COORDINATE

4.1 Preface

Kinetic isotope effect studies have been used to probe the nature of the hydrogen transfer in enzymatic reactions. In the DHFR reaction, the pro-*R* hydride at the C4 position of the nicotinamide ring is transferred to DHF, and primary hydrogen KIE measurements were performed by employing cofactor that carries a deuterium (or tritium) label at this position (87, 88). Homologues from different species, including the mesophile *E. coli* (EcDHFR) and psychrophilic *M. profunda* (MpDHFR) (113) as well as the thermophiles *G. stearothermophilus* (BsDHFR) (132) and *T. maritima* (TmDHFR) (120), have been tested. While the chemistry of hydride transfer has become well-characterised by this method, our understanding on enzyme catalysis can be enormously advanced by measuring KIEs with isotopologous cofactors that carry label(s) in the other positions of the nicotinamide ring. These alternative experiments include: α-secondary (2$^\text{o}$) hydrogen KIE measurements, which employ cofactor that carries a deuterium label at the pro-*S* position, and heavy atom KIE measurements, which use ^{13}C-labelled cofactor. Results derived from these measurements can provide insightful details regarding the impact of hybridisation change upon hydride transfer. Accordingly, in combination with the primary KIE data, the short-lived transition state can be precisely illustrated at an atomic scale. Perhaps most importantly, this fundamental knowledge will help in unravelling the role of protein dynamics in enzyme reactions, and specifically to the controversial existence of "promoting motions", an area of research that is thought to be essential to our understanding of the immense catalytic power of enzymes.

Eyring first proposed the idea of an activated complex that is formed during enzyme catalysis. Whilst this has long been accepted, analysis of this activated complex (*i.e.* the transition state) remains difficult. Transition state theories (TST), including the Bell tunnelling correction, were found to be inadequate to explain complex enzyme systems, particularly those involving quantum mechanical tunnelling. This has led to more complex models, including variational TST and Marcus-like models (see Introduction for a more detailed explanation). In the case of DHFR, if the enzyme has evolved to promote hydride transfer by modifying the barrier width, these non-statistical motions likely act via the neighbouring atoms, such as the C4 and the pro-

S hydride, in order to narrow the donor-acceptor distance (DAD). To verify the proposed existence of these "promoting motions", it is of great interest to analyse the role of these atoms in the transition state of hydride transfer.

The 2° KIE was previously determined for TmDHFR at pH 7 (155), whereas for EcDHFR this parameter was only determined at pH 9, where hydride transfer is mostly rate-limiting and the authors could measure this parameter by a competitive KIE method. However, it is desirable to measure this parameter under physiologically relevant conditions, as it permits characterisation of the transition state most resembling that in its natural environment. Even though the magnitude of 2° KIE is generally small (1.05-1.20), we perceived that this parameter can be measured by direct comparison using a pre-steady state stopped-flow setup, because this method can distinguish rates that are differed by only ~3% (153, 155). Similarly, ^{13}C heavy atom isotope effects can be measured by the same method, when this parameter is above 1.03-1.04. We believed that the ^{13}C KIE parameter would likely be higher than this value (at least at low temperature), as quantum mechanical tunnelling dominates in the DHFR reaction (87, 88, 240). In addition to the factors mentioned above, ^{13}C heavy atom KIE measurement has never been reported in the studies of DHFR, likely because of the inherent difficulties in preparing [4-^{13}C]-NADPH.

Here, we report the α-secondary hydrogen KIEs for EcDHFR, MpDHFR and BsDHFR measured under physiologically relevant conditions. Also, we report the first ^{13}C KIE measurement on the hydride transfer step catalysed by EcDHFR. These results shed light onto the mechanism of the DHFR reaction.

4.2 Results and Discussion

4.2.1 α-secondary hydrogen kinetic isotope effects

2° KIEs were measured between NADPH and (*S*)-NADPD at pH 7 for EcDHFR as well as the extremophiles MpDHFR and BsDHFR. The labelled cofactor was prepared based on the method developed by Ottolina and coworkers (262), which uses glucose dehydrogenase to introduce a deuterium label into the *si* face of the C4 in NADP$^+$ (Scheme 4.1).
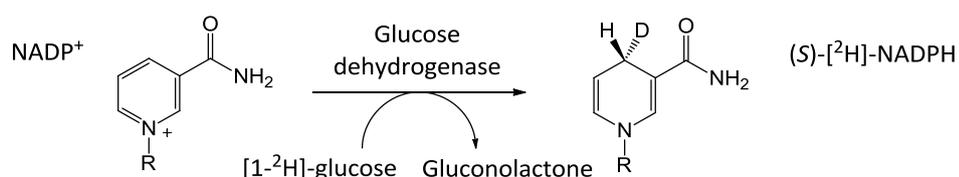
Scheme 4.1 Synthesis of (*S*)-[^{2}H]-NADPH from NADP$^+$ and [1-^{2}H]-glucose through the action of glucose dehydrogenase.

These parameters were measured using a stopped-flow setup, which monitored the decrease in FRET signal. Since the C4 position of NADPH changes from sp^3 to sp^2 hybridisation during catalysis, a normal secondary KIE over 1 was expected (see Introduction for more details). The average 2° KIE values were evaluated to be 1.044 ± 0.01 for MpDHFR, 1.058 ± 0.02 for EcDHFR and 1.048 ± 0.02 for BsDHFR (Table 4.1, Figure 4.1). At 10 °C, the 2° KIE for BsDHFR contains a relatively large error (1.092 ± 0.054), likely due to the poor signal-to-noise ratio in the FRET measurement (see Chapter 2. For future reference, we will likely perform the experiment again by measuring the decrease of absorbance at 340 nm.).

Table 4.1 Temperature dependence of α-secondary hydrogen KIEs for Mp, Ec, Bs and TmDHFR at pH 7.

T (°C)	MpDHFR	EcDHFR	BsDHFR	TmDHFR[A]
5	1.040 ± 0.010	1.066 ± 0.020	-	1.068 ± 0.003
10	1.040 ± 0.031	1.055 ± 0.028	1.092 ± 0.054	1.063 ± 0.010
15	1.055 ± 0.018	1.059 ± 0.033	1.044 ± 0.016	1.064 ± 0.007
20	1.036 ± 0.025	1.083 ± 0.031	1.044 ± 0.021	1.032 ± 0.009
25	1.037 ± 0.023	1.049 ± 0.012	1.037 ± 0.015	1.072 ± 0.007
30	1.059 ± 0.006	1.060 ± 0.043	1.037 ± 0.029	1.056 ± 0.006
35	-	1.033 ± 0.013	1.053 ± 0.027	1.043 ± 0.005
40	-	-	1.049 ± 0.027	1.054 ± 0.007
45	-	-	1.045 ± 0.021	1.040 ± 0.013
50	-	-	1.032 ± 0.009	1.049 ± 0.018
55	-	-	-	1.040 ± 0.040
60	-	-	-	1.032 ± 0.014
Average	1.044 ± 0.01	1.058 ± 0.02	1.048 ± 0.02	1.052 ± 0.02
[A] Values taken from ref (155)				
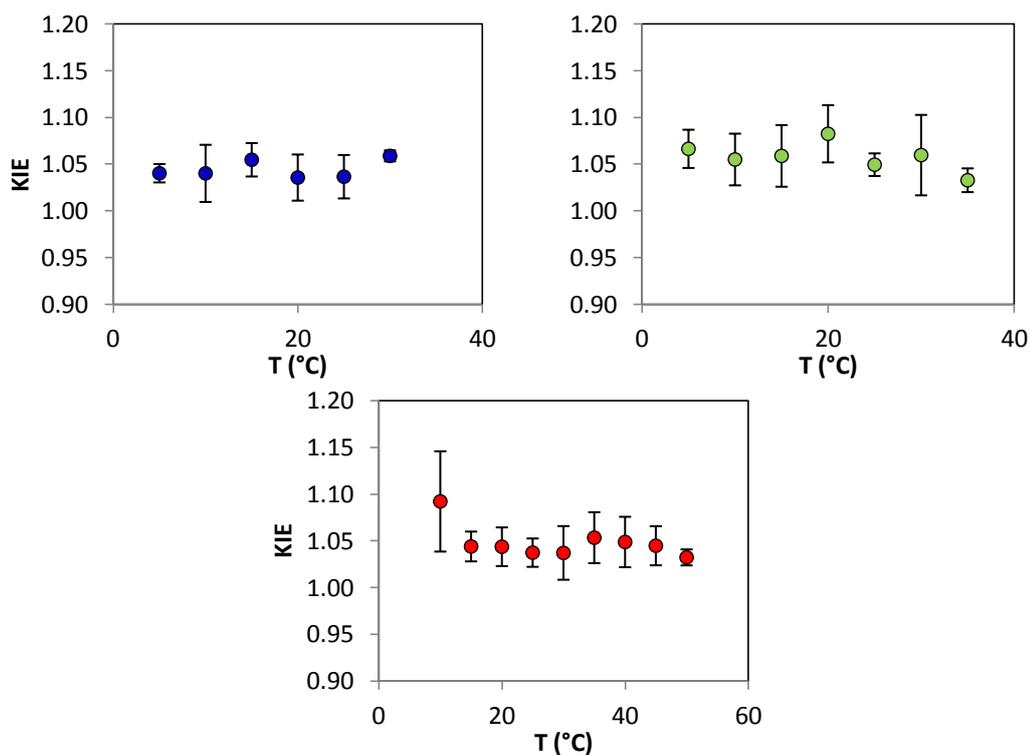
Figure 4.1 Temperature dependence of α-secondary KIEs at pH 7 for A) MpDHFR, B) EcDHFR and C) BsDHFR.

There seemed to be no apparent difference for the 2° KIE values between the ones measured in this investigation and those measured for TmDHFR (155) (Table 4.1). Kohen and coworkers previously studied EcDHFR at pH 9 and showed that the intrinsic 2° KIE is 1.052 ± 0.007. This correlates well with the value measured in this investigation (1.058 ± 0.02), although the experimental methods differ greatly (88). All of these parameters are temperature independent (Table 4.2), with no significant difference in activation energies between the isotopologous cofactors.

Table 4.2 Average α-2° KIE and ΔE$_a$ for Mp, Ec, Bs and Tm DHFR at pH 7.

	MpDHFR	EcDHFR	BsDHFR	TmDHFR[A]
Average α-2° KIE	1.044 ± 0.01	1.058 ± 0.02	1.048 ± 0.02	1.052 ± 0.02
ΔE$_a$ (kJ mol^{-1})	0.7 ± 1.1	0.53 ± 0.99	0.52 ± 0.59	0.4 ± 0.5
[A] Values taken from ref (155)				

In order to extrapolate the meaning of these findings, we have to compare these values with the equilibrium isotope effect (EIE). EIE is defined as the equilibrium distribution of isotopes between two stable states, *i.e.* the reactant and product states, whereas KIEs describes the difference between ground state and transition state (148). In other words, comparing the magnitude of KIE with that of EIE would reveal details about the hybridisation state of the transition state. Huskey and Schowen have hypothesised that, when a 2° KIE exceeds the equilibrium isotope effect (EIE), there is a coupling of the motion between the primary and secondary hydrogen, with a concomitant contribution from tunnelling (152). This theory was applied to the investigation of yeast ADH and showed exalted 2° KIEs above the experimental EIE (150, 151). In the case of the DHFR reaction the EIE was determined to be 1.13 (149). The lower magnitude of 2° KIE measured here implies that there is no coupling between the motion of the α-secondary hydrogen and the primary hydride that is transferred in all three DHFR enzymes.

Additionally, the relationship between primary and secondary KIE has been investigated by Scrutton and colleagues. In their studies, morphinone reductase

(MR) and pentaerythritol tetranitrate reductase (PETNR), a pair of related enzymes that use NADH as hydride source, were tested. Although both of these enzymes catalyse the hydride transfer to flavin mononucleotide (FMN) and they are structurally similar, the temperature dependence of primary KIE widely differed. In MR the hydride KIE drops noticeably with increasing temperature, while the corresponding KIE in PETNR remains constant at most temperatures (154). In contrast, the 2° KIEs for both of these enzymes are essentially identical and are not sensitive to temperature changes (153). Consequently, the authors hypothesised that the 2° KIE actually reports on the reaction ready configuration (RRC), which refers to the NADH-FMN arrangement that is subsequent to conformational sampling/preorganisation but immediately prior to tunnelling (153). The similarity in the magnitude of secondary KIE suggested that MR and PETNR adopt similar NADH-FMN arrangement, and a temperature independent 2° KIE was considered as proof that this arrangement does not alter significantly over the tested temperature range (153).

In the case of DHFR, the pattern of primary KIE changes greatly among homologues. For EcDHFR and MpDHFR the primary KIEs are only weakly temperature dependent (87, 113), but for BsDHFR there is a biphasic plot on 1° KIE (Chapter 2), which has a strong temperature dependence at 5-20 °C. Since the measured 2° KIEs are apparently temperature-independent and are essentially the same among all tested DHFR homologues, this parameter also likely reports on the configuration immediately before the tunnelling event (153). Even though these enzymes greatly differ in their flexibility, conformational preferences and likely their conformational sampling/preorganisation behaviours (113, 117, 236), the secondary KIE results suggest that substrate-cofactor configurations immediately before tunnelling are essentially the same among these enzymes (155).

4.2.2 $^{12}C/^{13}C$ Isotope effects on EcDHFR catalysis

This thesis also presents the first ^{13}C isotope effect studies on the DHFR reaction. [4-^{13}C]-NADPH and (*R*)-[4-^{2}H,^{13}C]-NADPH were prepared via the *E. coli* NADP$^+$

biosynthetic machinery (Chapter 3). In short, these cofactors were prepared by incorporating a ^{13}C label into the C4 position of the nicotinamide ring in NADP[+], which was then reduced stereoselectively by *Thermoanaerobacter brockii* alcohol dehydrogenase (Figure 4.2). Although carbon isotope effects have been used in a number of enzyme reactions, to our knowledge this technique has only been used on one previous to investigate enzymatic hydrogen tunnelling reaction (263). Additionally, this study only utilised carbon isotope effects at a single temperature, compared to the temperature dependence of the isotope effect studied here (263).
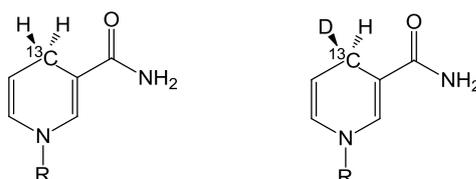
Figure 4.2 Structures of [4-^{13}C]-NADPH and (R)-[4-^{2}H,^{13}C]-NADPH showing the positions of the isotopic labels.

To determine the effect of ^{13}C labelling on the EcDHFR catalysed hydride transfer, the rates of the chemical step were measured between 5 and 35 °C at pH 7 (Figure 4.3). When NADPH and its ^{13}C-labelled isotopologue were employed, the ^{12}C/^{13}C isotope effect (k_{12C}/k_{13C}) is close to unity above 15 °C (~1.02) but increases sharply to 1.10 ± 0.03 at 5 °C (Table 4.3, Figure 4.3). Interestingly, the magnitude and temperature dependence of the carbon isotope effects reported in this investigation have never been observed in any other enzymatic system. Previous calculations on simple molecules, such as methane, propane and other gaseous alkanes, have indicated that a typical vibrational frequency difference between a ^{12}C-H and ^{13}C-H bond is approximately 20-60 cm[-1] (264, 265). Accordingly if there is a 20 cm[-1] vibrational frequency difference, the theoretical ^{12}C/^{13}C isotope effect calculated by the Arrhenius equation would be 1.04 at 25 °C and it would decrease by approximately 0.05-1% over the temperature range of ~30 °C (28). Owing to the structural complexity of NADPH, the ^{12}C-H and ^{13}C-H stretching vibrational

frequencies have not been determined at this time. Furthermore, the EcDHFR active site involves numerous interactions with the nicotinamide ring, most notably hydrophobic interaction with Ile14 and Tyr100, and a hydrogen bond network involving Ser49 and a water molecule (76). Nevertheless, for the purpose of discussion, we hypothesise that there might be a theoretical vibrational frequency difference of 40 cm^{-1} between NADPH and ^{13}C-NADPH. This would result a ^{12}C/^{13}C isotope effect of 1.1 at 5 °C, a similar magnitude to the observed value. However, this argument does not hold true for the observed temperature dependence. The theoretical ^{12}C/^{13}C KIE would only drop by 1% from 5 to 20 °C, but the experimental KIE dropped markedly by ~10%.

Whilst the stretching vibrational mode will contribute most to the observed isotope effect, a contribution from other vibrational modes should also be taken into consideration. In the DHFR reaction, NADPH is oxidised to NADP$^+$ therefore the C4 atom is rehybridised from sp^3 to sp^2. Since the out-of-plane bending modes differ significantly between sp^3 and sp^2 hybridised states, this might contribute to the observed KIE. As discussed above, the measured secondary hydrogen KIEs are temperature-independent, whereas the observed ^{13}C heavy atom isotope effect displays a relatively strong temperature dependence. Rehybridisation alone therefore could not be used to explain the sudden increase in magnitude at low temperature.

Table 4.3 Temperature dependence of ^{12}C/^{13}C isotope effects between 5 and 35 °C at pH 7.

T (°C)	NADPH/[4-^{13}C]-NADPH KIE	(*R*)-[4-^{2}H]-NADPH/(*R*)-[4-^{2}H,^{13}C]-NADPH KIE
5	1.10 ± 0.03	1.15 ± 0.02
7	-	1.13 ± 0.02
10	1.10 ± 0.03	1.12 ± 0.02
12	-	1.10 ± 0.02
15	1.05 ± 0.02	1.11 ± 0.03
20	1.01 ± 0.04	1.11 ± 0.03
25	1.02 ± 0.03	1.07 ± 0.04
30	1.03 ± 0.03	1.07 ± 0.04
35	1.03 ± 0.03	1.03 ± 0.03
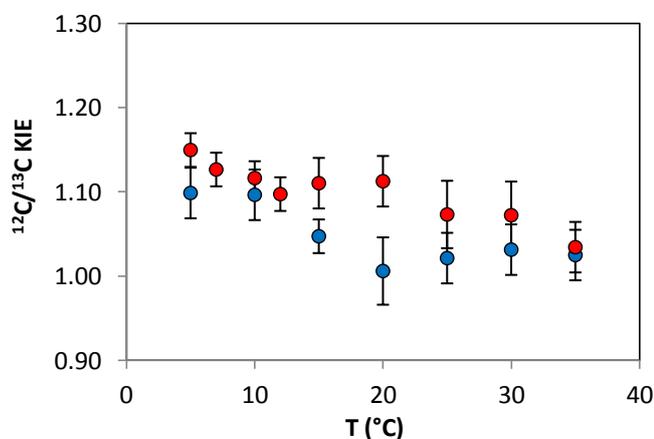
Figure 4.3 Temperature dependence of the ^{12}C/^{13}C KIE on hydride transfer (blue) and ^{12}C/^{13}C KIE on deuteride transfer (red) at pH 7.

The reacting atoms in the DHFR reaction were further investigated by measuring the heavy atom isotope effect with deuterated cofactors, NADPD and ^{13}C-NADPD. A similar trend of temperature dependence was observed, with ^{13}C/^{12}C isotope effect decreasing from 1.15 ± 0.02 at 5 °C to 1.03 ± 0.03 at 35 °C. Because the reduced mass difference is greater upon deuterium incorporation (Table 4.4), it is not unexpected that the isotope effects are generally higher than that of the protiated equivalent. Moreover, the increased mass difference could also lead to a greater

disparity in vibrational frequency. For a theoretical difference of 60 cm^{-1} between the labelled and unlabelled cofactors, the calculated KIE would be 1.16 at 5 °C. However, it would also lead to an isotope effect that greatly differs from the experimental one at 20 °C. This implies that a difference in vibrational frequencies could not be the sole factor affecting the observed ^{12}C/^{13}C isotope effects.

Table 4.4 The reduced mass of a C-H and C-D bond involving a ^{12}C and ^{13}C atom.

	^{12}C-H	^{13}C-H	^{12}C-D	^{13}C-D
Reduced mass	0.923	0.929	1.714	1.733
Difference between ^{12}C and ^{13}C	0.006		0.019	

Recently, Singleton and coworkers reported an exalted ^{12}C/^{13}C isotope effect of 1.16 at -100 °C on the ring opening of a cyclopropylcarbinyl radical species (Figure 4.4). Based on their VTST-MT calculations, the exalted isotope effect can be explained by a contribution from heavy atom tunnelling (266). Nevertheless, in the DHFR reaction, the transfer of the hydrogen dominates the chemical coordinate, thus the likelihood for a significant contribution from heavy atom tunnelling is extremely low. An alternative explanation to the observed isotope effect is consequently needed.
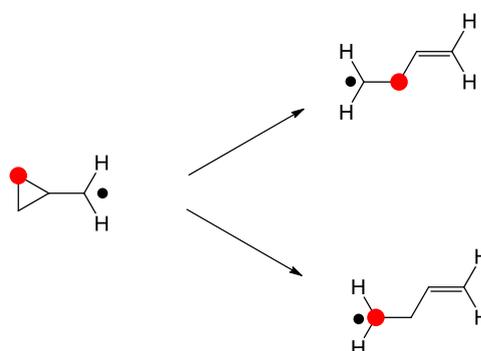
Figure 4.4 Ring-opening reaction of cyclopropylcarbinyl radical studied by Singleton and coworkers (266). The atoms subjected to heavy atom KIE measurements are noted by red circles.

Two explanations are available to explain the observed $^{12}C/^{13}C$ isotope effect. Firstly, the heavy atom incorporation may affect either the tunnelling probability or the frequency of recrossing events. As explained above, Marcus-like models invoke enzyme dynamics that include a distance sampling term. As this variable is mass- and temperature dependent and likely channels through the C4 position of NADPH, the exalted $^{12}C/^{13}C$ isotope effect at low temperature can be viewed as a result of changing the barrier width *i.e.* changing the tunnelling probability. Upon ^{13}C-labelling, the frequency of fs bond vibration of the C-H decreases and this might subsequently affect hydrogen transfer if the chemical coordinate is coupled to "promoting motions" (Figure 4.5). However, a previous investigation of EcDHFR by enzyme isotope substitution revealed that protein dynamics are not coupled to the barrier height and/or width. Instead, fs bond vibrations in EcDHFR coupled to the reaction coordinate through the frequency of dynamic recrossing rather than the shape of the reaction barrier (117). Thus, the observed increase in $^{12}C/^{13}C$ at low temperature may be due to changes in the recrossing coefficient between the light and heavy cofactor.
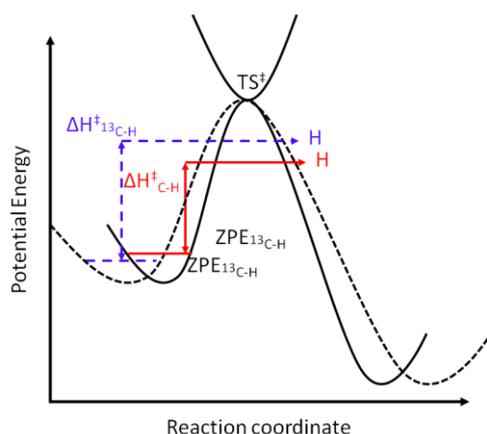
Figure 4.5 Example of increased barrier width upon direct coupling of the C4 atom to the primary hydrogen. The coupled and uncoupled reaction pathways are shown in blue and red respectively.

Alternatively, the large effects at low temperature may link to a microscopic mechanism that affects the observed pre-steady state rate. That is, the observed isotope effect could report on the step of hydride transfer as well as other step(s)

that are isotopically sensitive in the DHFR catalysis. This hypothetical step likely occurs after electrostatic preorganisation, which should not be isotopically sensitive as indicated by the Born-Oppenheimer approximation. The C-H bond vibration could affect step(s) immediately before the formation of the RRC. Given that the RRC is relatively static as revealed by the 2° KIE studies, the experimental ^{13}C isotope effect could be accounted by a change in reorganisation energy of the reactants. At high temperatures reorganising the substrate dipole is not rate determining, but when temperature is below 15 °C the energy commitment for reorganisation becomes partially rate limiting and produces the profound carbon isotope effects (Table 4.5). When a deuteride is inserted at the C4 position, the postulated reorganisation energy further increases and provides a greater difference in activation energy between ^{12}C-NADPD and ^{13}C-NADPD (ΔE_a 4.3 kJ mol^{-1}) (Table 4.5).

Table 4.5 Activation energy of NADPH and labelled isotopologues

	NADPH	[4-^{13}C]-NADPH	(*R*)-[4-^{2}H]-NADPH	(*R*)-[4-^{2}H,^{13}C]-NADPH
E$_a$ (kJ mol^{-1})	31.8 ± 0.8	33.7 ± 1.0	41.7 ± 0.5	46.0 ± 1.0
ΔE$_a$ (kJ mol^{-1})	1.8 ± 1.2		4.33 ± 1.1	

4.2.3   Effect of ^{13}C isotopic substitution on the 1° KIE of EcDHFR

Numerous Marcus-like models suggest that heavy-atom motions promote the rate of reaction by modulating the barrier width between the substrate and reactant. If promoting motions indeed enhance hydride tunnelling via the C4 atom, the magnitude of primary hydride KIE would consequently be affected upon ^{13}C labelling (Figure 4.6). This would theoretically allow the ^{12}C/^{13}C isotope effect observed above to be designated as a promoting motion or as a result of recrossing or reorganisational changes.
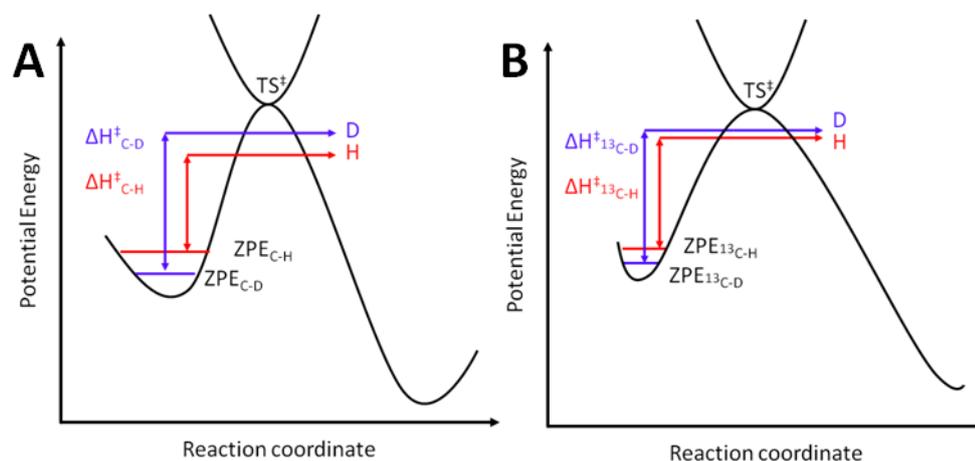
Figure 4.6 A) Reaction barrier of EcDHFR as implied from previous KIE studies, B) Theoretical reaction barrier upon decreasing hypothetical "promoting motions", causing the barrier width to increase.

To verify this hypothesis, the unlabelled ($k_{12C-NADPH}/k_{12C-NADPD}$) primary KIE was compared to the ^{13}C-labelled 1° KIE ($k_{13C-NADPH}/k_{13C-NADPD}$). The resulting ^{13}C-labelled primary hydrogen KIE was calculated to be 3.66 at 5 °C and decreased to 2.01 at 35 °C (Table 4.6). When these values are compared to the unlabelled (^{12}C) primary KIE, there seems to be no significant difference (Figure 4.8).
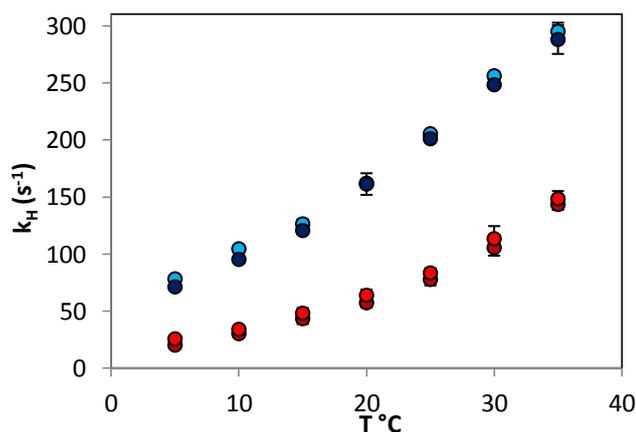
Figure 4.7 Rate of hydride transfer with NADPH (light blue), [4-^{13}C]-NADPH (dark blue), (*R*)-[4-^{2}H]-NADPH (light red) and (*R*)-[4-^{2}H,^{13}C]-NADPH (dark red) at pH 7.

Table 4.6 Temperature dependence of the primary hydrogen KIE of EcDHFR with ^{12}C NADPH and ^{13}C NADPH at pH 7.

T (°C)	^{13}C labelled primary hydrogen KIE	Primary hydrogen KIE
5	3.66 ± 0.39	3.05 ± 0.29
10	3.12 ± 0.12	3.07 ± 0.11
15	2.79 ± 0.32	2.64 ± 0.31
20	2.81 ± 0.19	2.54 ± 0.17
25	2.59 ± 0.16	2.46 ± 0.13
30	2.36 ± 0.16	2.27 ± 0.24
35	2.01 ± 0.03	1.99 ± 0.05
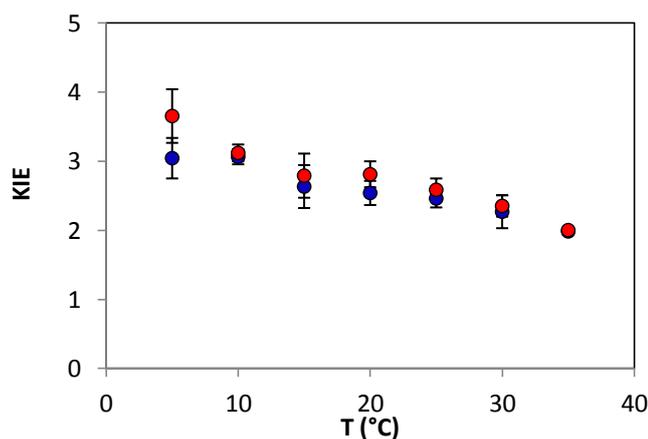
Figure 4.8 Temperature dependence of the primary hydrogen KIE (blue) and ^{13}C-labelled primary hydrogen KIE (red) at pH 7.

The lack of change of the ^{13}C-labeled primary KIE suggested that promoting motions do not act via the C4 atom in NADPH. A drop in primary KIE would be expected as a consequence of ^{13}C labelling because the barrier width for hydride transfer increases. Furthermore, the proposed recrossing or reorganisation effects (see above) should not cause any obvious changes in the primary KIE, as a similar magnitude of isotope effect is observed for both ^{13}C-NADPH and ^{13}C-NADPD thus cancelling out any effect. Overall, based on our speculations, the ^{12}C/^{13}C KIE measurements should report on the mass- and temperature-dependent reorganisation process, which occurred prior to tunnelling, or an alteration in the reaction pathway that results in more recrossing events upon isotopic labelling of

the C4 position. Nonetheless, current limitations in the experimental methods must be taken into account.

The primary hydrogen KIE results suggested that ^{13}C labelling is not causing any profound effect on the barrier width that affects hydride transfer. However, the difference imposed by ^{13}C labelling is minimal when compared to that by deuterium substitution, thus reactivity difference caused by ^{13}C labelling might be masked by the primary hydrogen KIE. Competitive KIE measurements might be needed to verify if there is any subtle difference on the primary KIE that might have been overseen in this investigation. However, these are incredibly difficult to measure at pH 7. Similarly, EA-VTST calculations could determine if any of the components of the transmission coefficients are being altered. This would locate which components of the reaction coordinate that are isotopically sensitive.

4.2.4 Future studies

Our investigations clearly showed that the pro-*S* hydride as well as the C4 atom participate in the transition state of the DHFR reaction. For future studies, it is of great interest to analyse the role of other atoms in the nicotinamide ring. For example, investigating the isotope effects of the C3 and C5 atoms would highlight if any coupling motions would act via these atoms as a method to enhance the rate of hydride transfer. Furthermore, nitrogen isotope effect studies on the N1 position and the amide moiety could reveal effect of charge change and the function of hydrogen bonding in the reaction, respectively. Lastly, a competitive method that can be performed at pH 7 will need to be developed in order to fully characterise the transition state of DHFR under physiologically relevant conditions. An alternative DHFR homologue in which hydride transfer is more rate-limiting might be needed. This will provide more insights, as the intrinsic KIEs and the forward commitment factors can be extrapolated.

4.3   Conclusion

In this investigation, the α-secondary hydrogen isotope effects (2° KIEs) at pH 7 were measured for EcDHFR, MpDHFR and BsDHFR. The measured isotope effects are essentially the same among one another, including TmDHFR which was previously investigated. Furthermore, the observed 2° KIEs were significantly lower than the calculated EIE, ruling out the potential coupling between the primary and secondary hydrogen atoms at the C4 position of NADPH. This also coincides with the results for TmDHFR, and supports the argument that α-secondary hydrogen isotope effect reports on the RRC immediately prior to hydride transfer.

This thesis also documented the first ^{13}C isotope effect measurement on the hydride transfer catalysed by EcDHFR. These studies clearly showed that the C-4 position participate in the transition state of DHFR, but the temperature dependence of the ^{12}C/^{13}C isotope effects (on both hydride and deuteride transfer) cannot be explained by semi-classical or simple tunnelling models. While there is a strong carbon isotope effect at low temperature, ^{13}C labelling does not seem to change the magnitude of primary hydride KIE. This contradicts the hypothetical promoting motions postulated in the modified Marcus theories (25, 39, 41). Instead, these results indicate that carbon isotope effect might arise as a consequence of reorganisation effects that occur immediately before hydride transfer. However, limitations on the pre-steady state rate measurements might override the subtle difference caused by ^{13}C labelling, especially on the primary KIE calculations. Precise isotope effect measurements by a competitive method and/or additional computational studies will therefore be needed in future studies.

In summary, the ^{13}C isotope effect results could not be explained by current Marcus-like theories. In comparison to previous experimental results, work presented in this chapter highlights the importance of the reaction ready configuration in DHFR catalysis and shows that many different factors contribute to the formation of a reaction coordinate conducive to hydride tunnelling.

5 RECONSTITUTING THE DIHYDROFOLATE BIOSYNTHETIC PATHWAY

5.1 Preface

A generic method to regiospecifically label dihydrofolate (DHF) with heavy isotopes would be enormously useful in enzymology studies. However, similar to the situation of isotopically labelled NADPH (Chapter 3), synthesising DHF isotopologues is difficult and not cost-effective. DHF and derivatives are relatively complex molecules that serve as one-carbon carriers in key metabolic pathways, but enzymes tend to employ only the pterin ring moiety for chemical reactions. As a matter of fact, these reactions typically centre around positions at N5, C6 and/or N10. One example would be the DHFR reaction, which catalyses a NADPH-dependent hydride transfer to the C6 position and protonation at the N5 position of DHF. The anti-cancer drug target thymidylate synthase is another example, in which these atoms in the pterin ring are responsible for the chemistry of the enzyme catalysis. With the use of N^5,N^{10}-methylene-THF, this enzyme performs a reductive methylation of deoxyuridine monophosphate to generate deoxythymidine monophosphate in DNA synthesis. Accordingly, to facilitate the studies of these DHF-related enzymology studies, it is of great interest to synthesise DHF with isotopes labelled at these positions. Current synthetic methods involve adding aminobenzoic acid and glutamate moieties to an oxidised pterin ring thus yielding labelled folate (267, 268). However, not only were these methods low yielding, the regioselectivity of isotopic label incorporation is undesirable due to the symmetrical nature of the reagents involved. Furthermore an additional step of reduction is required in order to obtain the desired product for this investigation. Hence, a new, simple and high-yielding method to synthesise DHF isotopologues is urgently needed.
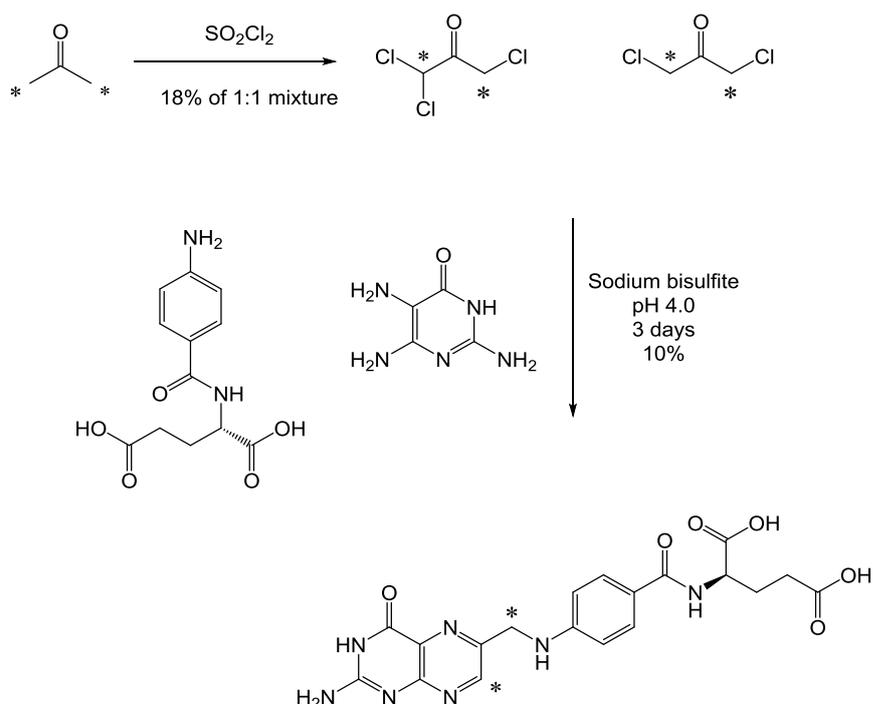
Scheme 5.1 Synthesis of [7,9-^{13}C$_2$]-folate from [1,3-^{13}C$_2$]-acetone, taken from ref (267)

In nature DHF is made via a ring-opening pathway of GTP, which in turn is synthesised from glucose. All the enzymes have been biochemically characterised with many of them involving interesting chemistry, such as dihydropteroate synthetase (DHPS), a well-established drug target (269). With a wealth of information available (227, 228, 232-234), it seems feasible to use the DHF biosynthetic pathway for heavy isotope incorporation (Scheme 5.2). Indeed, the research group of Bacher previously used this pathway to synthesise the isotopically labelled intermediate tetrahydropteroate. Interestingly, these enzymes have not been employed to label DHF with heavy isotopes (220), even though only simple, relatively inexpensive starting materials, such as isotopically labelled glucose, are required. Hence, it would be ideal to re-construct the DHF biosynthetic pathway in our laboratory.
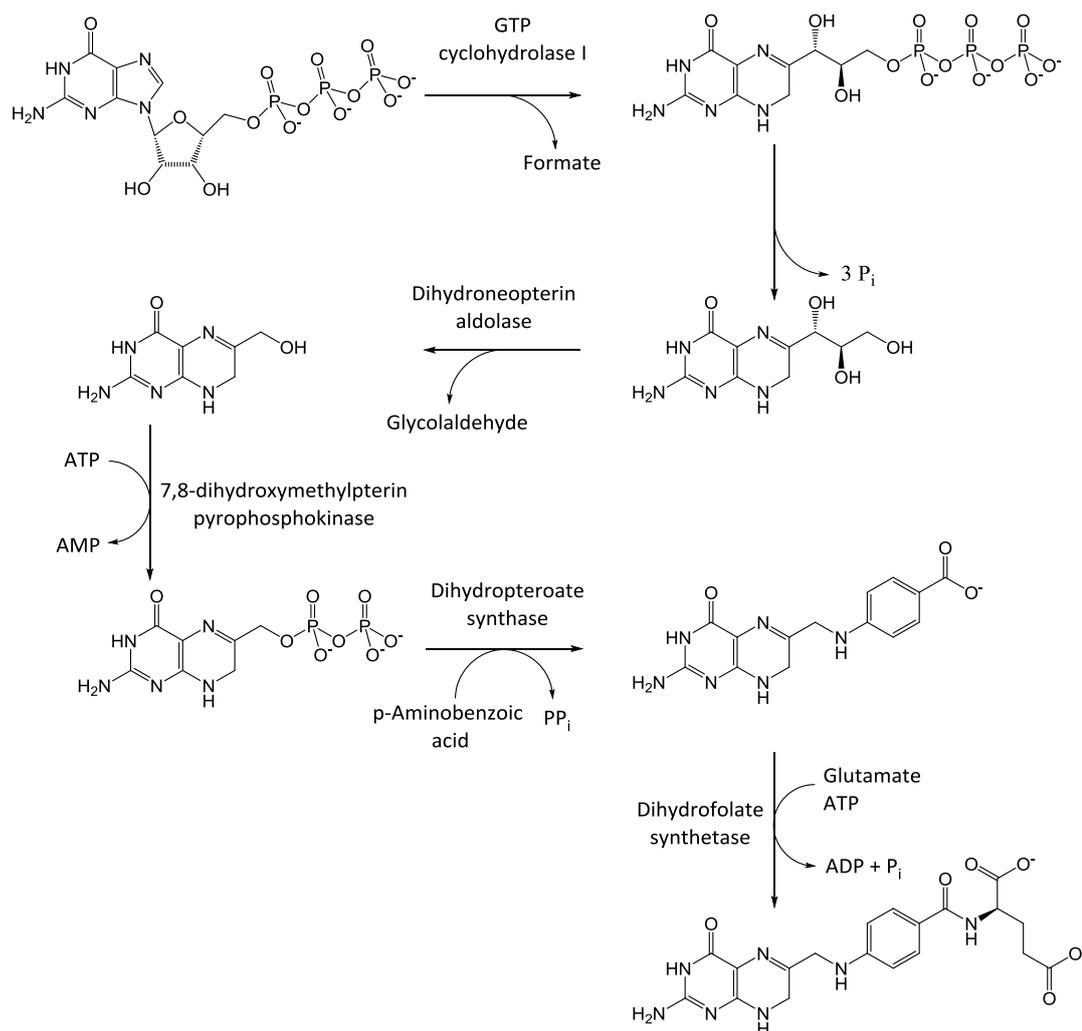
Scheme 5.2 Dihydrofolate biosynthesis pathway from GTP in *E. coli*.

This chapter outlines our efforts towards preparing the enzymes needed to synthesise DHF. As described above, DHF synthesis is extended from the biosynthetic pathway of GTP. Fortunately, at the time of this investigation most of the enzymes in the GTP pathway, except PRPP synthetase and xanthine-guanine phosphoribosyltransferase, were commercially available. On the other hand, enzymes that participate in DHF biosynthesis cannot be sourced from any biochemical company. Consequently, the genes for these missing enzymes were prepared by a variety of methods. Furthermore, the resulting enzymes were isolated and initial optimisations were performed.
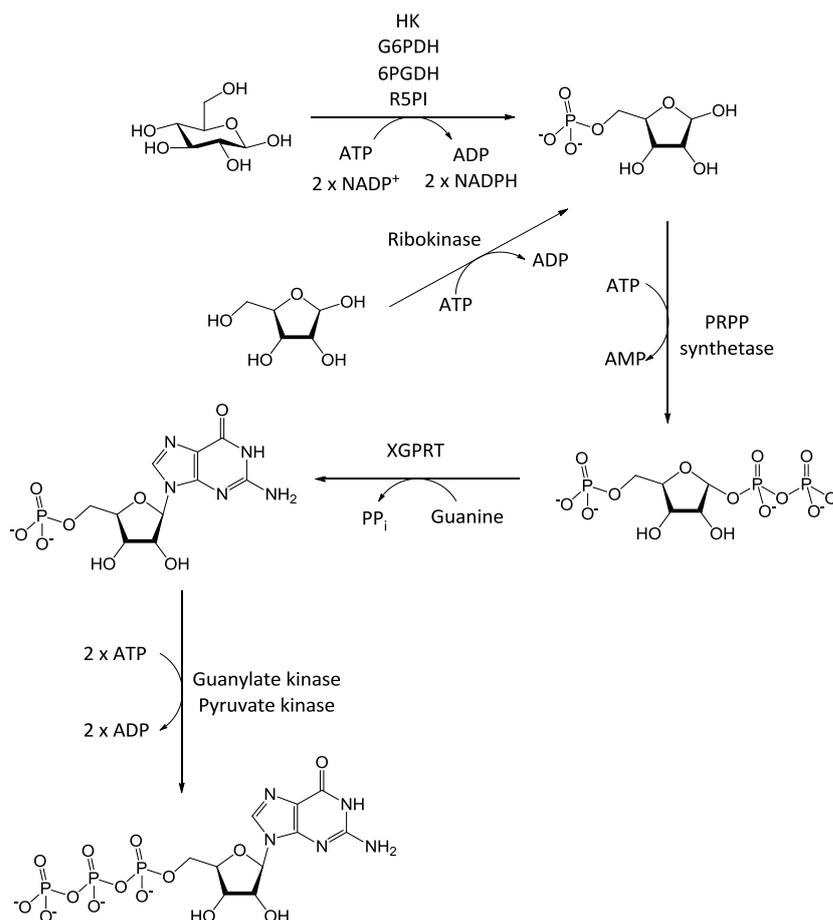
Scheme 5.3 Conversion of glucose to GTP via pentose-phosphate pathway, taken from (223). An alternate method from ribose is also shown. HK = hexokinase, G6PDH = glucose-6-phosphate dehydrogenase, 6PGDH = 6-phosphogluconate dehydrogenase, R5PI = ribose-5-phosphate isomerase and XGPRT = xanthine-guanine phosphoribosyltransferase.

5.2 Production and purification of DHF biosynthesis enzymes

5.2.1 PCR amplification and construction of expression vectors

Gene sequences encoding DHF biosynthesis enzymes were sourced from *E. coli* K-12 using the GenBank® database. *folB, C, E, K,* and *gpt* were amplified from *E. coli* strain XL1-Blue chromosomal DNA and cloned into expression vectors in identical fashion to the NADPH biosynthesis genes (Chapter 3). A temperature gradient was used to determine the optimal annealing temperature for each set of primers (Table 5.1), and agarose gel electrophoresis was used to visualise each gene. The *folC* gene was amplified successfully via PCR, however, one or both of the primers also bound to numerous locations on the gene resulting in multiple PCR products

(Figure 5.1). PCR amplification of *folP*, encoding dihydropteroate synthetase, was unsuccessful despite numerous conditions attempted. As a result this gene and the *folC* gene were chemically synthesised and cloned into pET-21a expression vector (Epoch Life Sciences).

The purified expression vector constructs were sequenced (by Eurofins) to confirm that no mutations had occurred during recombinant DNA manipulation. The gene sequence of *folB* did not match any sequence in the GenBank® database, likely due to the fact that the specific *E. coli* XL1-Blue *folB* gene sequence has not been uploaded at this time. To remove the possibility of inactive enzyme being produced, three site-directed mutations were performed to align the sequence with a known *E. coli* sequence in the database. Similarly, the ligation of *folE* into pET21a had resulted in the gene being out of frame, so a single point mutation was performed immediately before the start codon of the gene.

Table 5.1 Conditions used for PCR amplification of the genes involved DHF biosynthesis.

Step	Target gene			
	folB	*folC, folK*	*folE*	*gpt*
Denaturation*	10 s, 98 °C			
Annealing	30 s, 42 °C	30 s, 43 °C	30 s, 67 °C	60 s, 69 °C
Extension	45 s, 72 °C			
* The denaturation, annealing and elongation step was cycled 35 times				
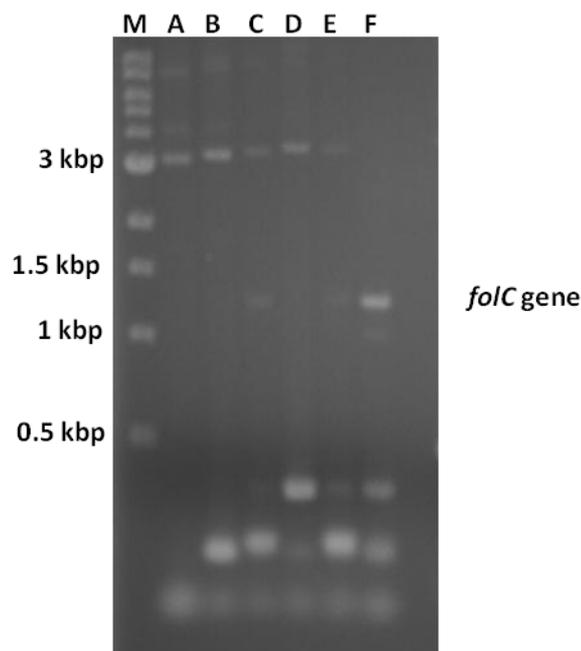
Figure 5.1 Agarose gel visualising the multiple products obtained from PCR amplification of *folC*. M = marker, A - F = PCR amplifications of *folC*.

5.2.2 Production and purification of enzymes

E. coli BL21 Star[TM] (DE3) cells were choose as host for expressing the genes for DHF biosynthetic enzymes, which were cloned into the pET-21a, -21d and -28a vectors. Since these genes were regulated by a lac operator/T7 promoter system, the corresponding enzymes were translated upon the addition of IPTG. pET-28a encodes a N-terminal his tag, thus the expressed protein can be purified by Ni^{2+} affinity chromatography. Accordingly, the genes for FolB, FolK and Gpt were expressed at 37 °C, following IPTG induction when the OD$_{600}$ of the host cell culture reached 0.6. Test expressions were carried out to determine the optimal induction time for each protein. In all cases, a profound band of protein was observed after 4 hrs of induction. Dihydroneopterin aldolase (DHNA), 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK) and xanthine-guanine phosphoribosyltransferase (XGPRT) were purified by Ni^{2+}-affinity chromatography (Figure 5.2). All these proteins were isolated with reasonable purity as indicated by SDS-PAGE analysis.
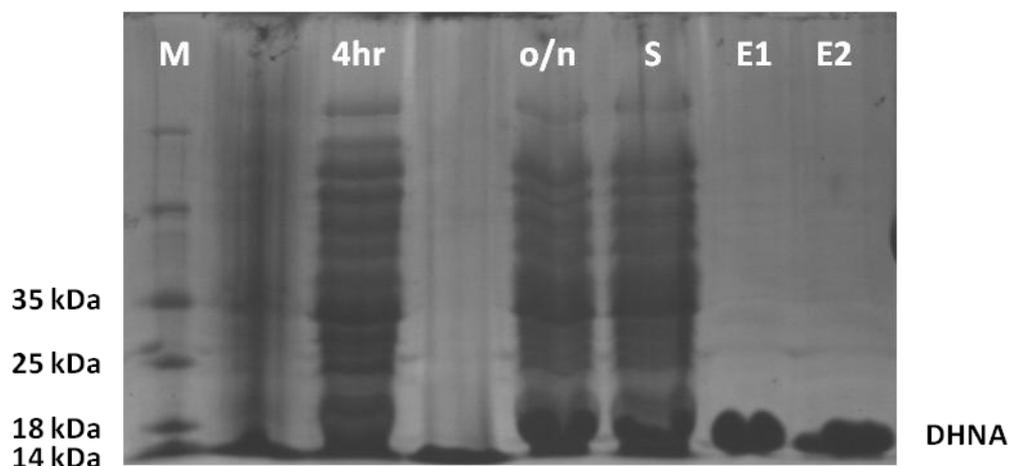
Figure 5.2 SDS-PAGE analysis of *folB* expression and DHNA purification. M = marker, 4 hr = 4 hours after induction, o/n = overnight after induction, S = supernatant, E1 and E2 = elution fractions.

The pET-21a and pET-21d vectors do not encode a N-terminal his-tag, thus anion-exchange chromatography was used to purify GTP cyclohydrolase I (GTPCI), dihydroneopteroate synthase (DHPS) and dihydrofolate synthetase (DHFS). *E. coli* BL21 Star™(DE3) cells were grown at 37 °C until an OD_{600} of 0.6 was reached. Subsequently, IPTG (0.5 mM) was added to induce gene expression. GTPCI showed a clear protein band in SDS-PAGE analysis after 4 hours of incubations at 37 °C. However DHPS and DHFS required expression overnight to produce significant amounts of protein. As each of these proteins have a theoretical pI below 7, anion-exchange chromatography was performed using 50 mM Tris, pH 8 buffers, 5 mM βME. All three enzymes bind to positively charged quaternary ammonium resins, and they were eluted at various concentrations of salt buffer (50 mM Tris, 1 M NaCl, pH 8), at 35% for GTPCI, at 40% for DHPS and 35% for DHFS. The SDS-PAGE analyses of these samples showed the correct monomeric size for each protein. All six proteins gave bands on the SDS-PAGE corresponding to the correct monomer size determined from the amino acid sequences.
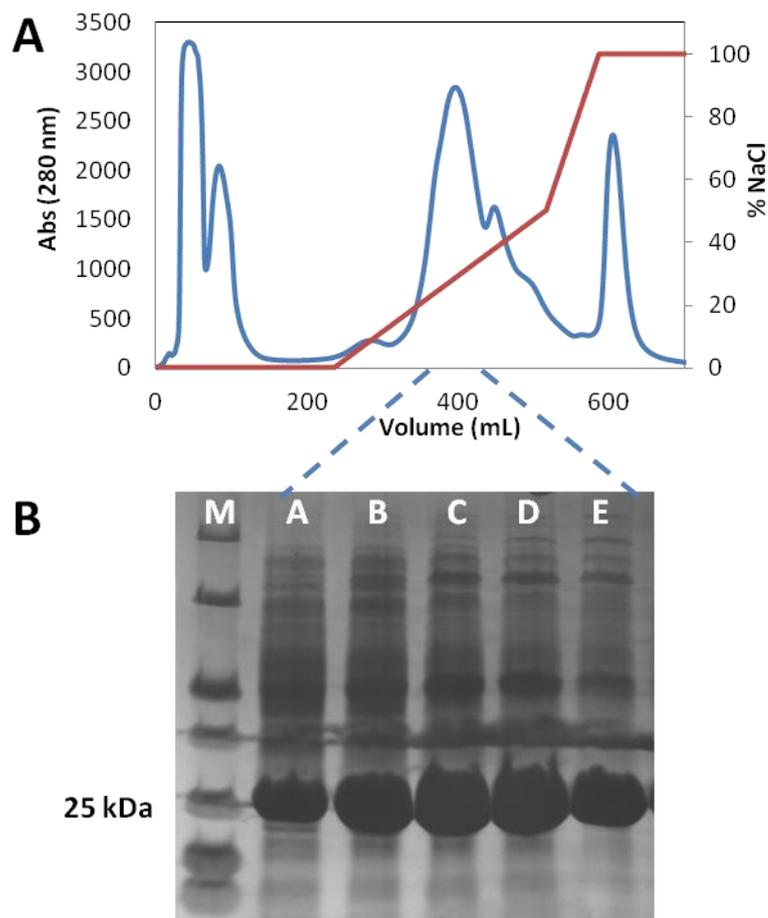
Figure 5.3 A) Chromatogram showing the purification of GTPCI by anion-exchange chromatography. UV absorption of eluted protein is shown in blue and concentration of NaCl in red. B) SDS-PAGE of fractions corresponding to elution peak from the anion-exchange column. M = MW marker, A-F = elution fractions corresponding to the region between the dashed lines.

5.3 Characterisation of the DHF biosynthesis enzymes

Size-exclusion chromatography was used to investigate the native quaternary structure of the purified enzymes. CD spectroscopy, in turn, was employed to determine if each protein contains the appropriate secondary structures. These results together indicated that all of the DHF biosynthesis enzymes were purified in their folded states.

5.3.1 Size-exclusion chromatography

Size-exclusion chromatography was used to determine the native state of each enzyme in the DHF biosynthetic pathway. In the Superdex200 size-exclusion chromatographic analysis, the retention volume for each recombinant protein was compared with standard proteins that have known size and quaternary structures (227, 228, 232-235). As previous studies reported, DHPS and DHFS are stable as dimer and monomer, respectively (228, 232). However, the remaining four proteins exhibit quaternary properties that are different from reported observations (227, 233-235). GTP cyclohydrolase I was found to be a trimer (MW approximately 75000), whereas this enzyme was previously shown to be decameric (MW 250000). DHNA and XGPRT reportedly form an octamer (MW 110000) and a tetramer (MW 64000), but presented as a tetramer (approximately 60 kDa) and a dimer (MW approximately 40) in this investigation. Although HPPK has previously been reported as a monomer with mass 25000 in the native state (234), both the monomer and dimer (MW 50000) were purified in this investigation, as there are two resolved peaks in the chromatogram (C, Figure 5.4) with the dimer form being predominant.

For XGPRT and DHNA, the differences in their quaternary states are likely due to the attachment of a N-terminal his-tag, which affects how the monomer interacts with one another. Even though the his-tag can be cleaved with the thrombin cleavage site that is inserted in front the recombinant enzyme, these were shown to be active despite the different oligomer states (see below). On the other hand, recombinant GTPCI does not contain a his-tag at either the N- or C-terminal. It is therefore not clear why the decamer species is not observed during this investigation. Similar size-exclusion chromatograms were obtained in the presence and absence of reducing agent, thus ruling out a role for disulfide bond formation as the cause.
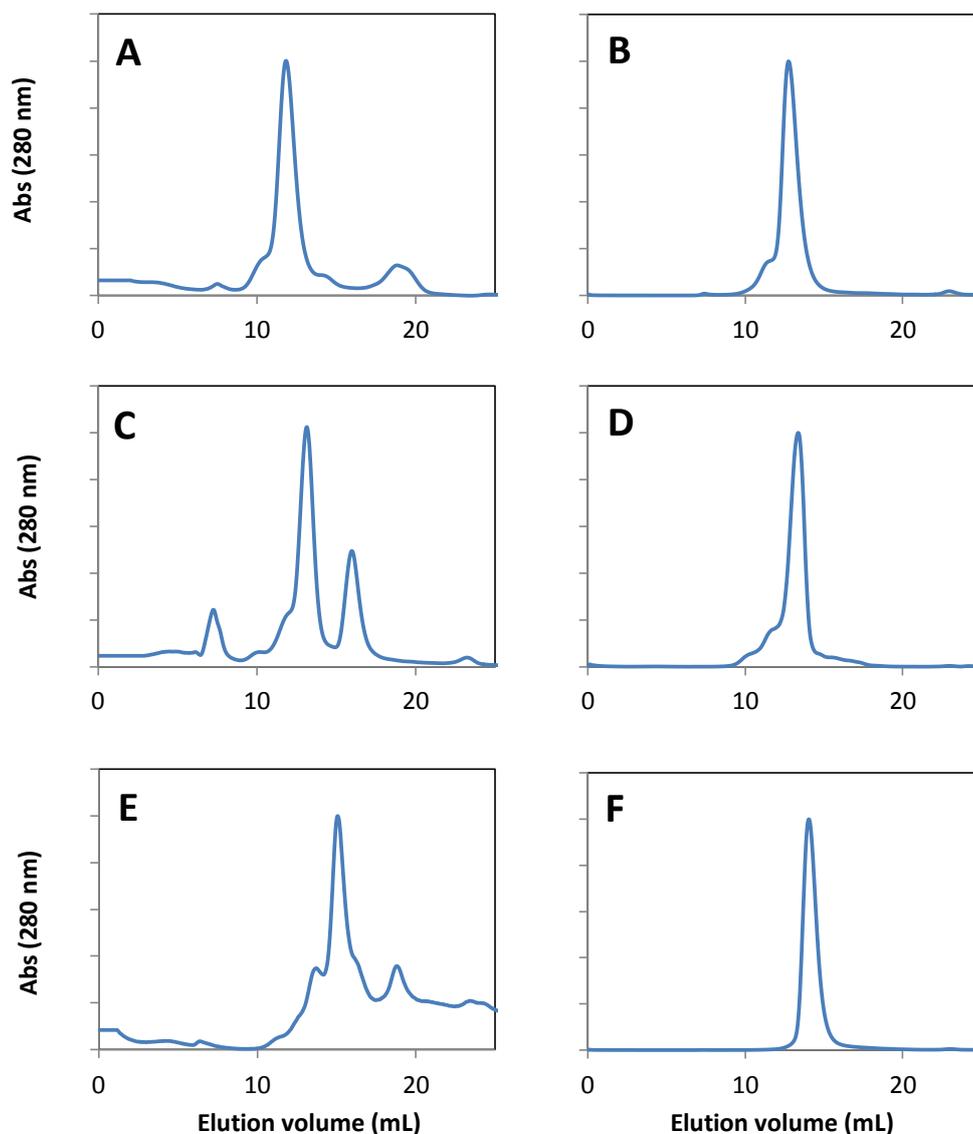
Figure 5.4 Chromatograms from Superdex200 size-exclusion column. A = GTPCI, B = DHNA, C = HPPK, D = DHPS, E = DHFS and F = XGPRT

5.3.2 CD spectroscopy

The secondary structural organisation of the recombinant proteins were analysed by CD spectroscopy. To obtain the mean residue ellipticity, CD signals of these samples were measured between 190 and 400 nm in phosphate buffer (10 mM at pH 7 and 20 °C). The minima at 222 and 206 nm as well as the maximum at 195 nm are indicative of α-helices and β-sheets, suggesting that these proteins are properly

folded. The CD spectrum of unfolded protein has a weak signal in the far UV region (190-250 nm), but this was clearly not the case in these enzymes.

Furthermore, the enzymes that exhibited different quaternary structures, including GTPCI, DHNA, and XGPRT, presented CD signals of folded proteins (A, B and F, Figure 5.5). This provided further additional evidence that protein folding is not strongly affected by the difference in the oligomeric state. The spectrum of HPPK showed that both the monomer and dimer species are folded (C, Figure 5.5). However, it is unclear at this moment whether the dimer species is active as this enzyme has not been optimised fully for DHF biosynthesis *in vitro* up to this point (see below).
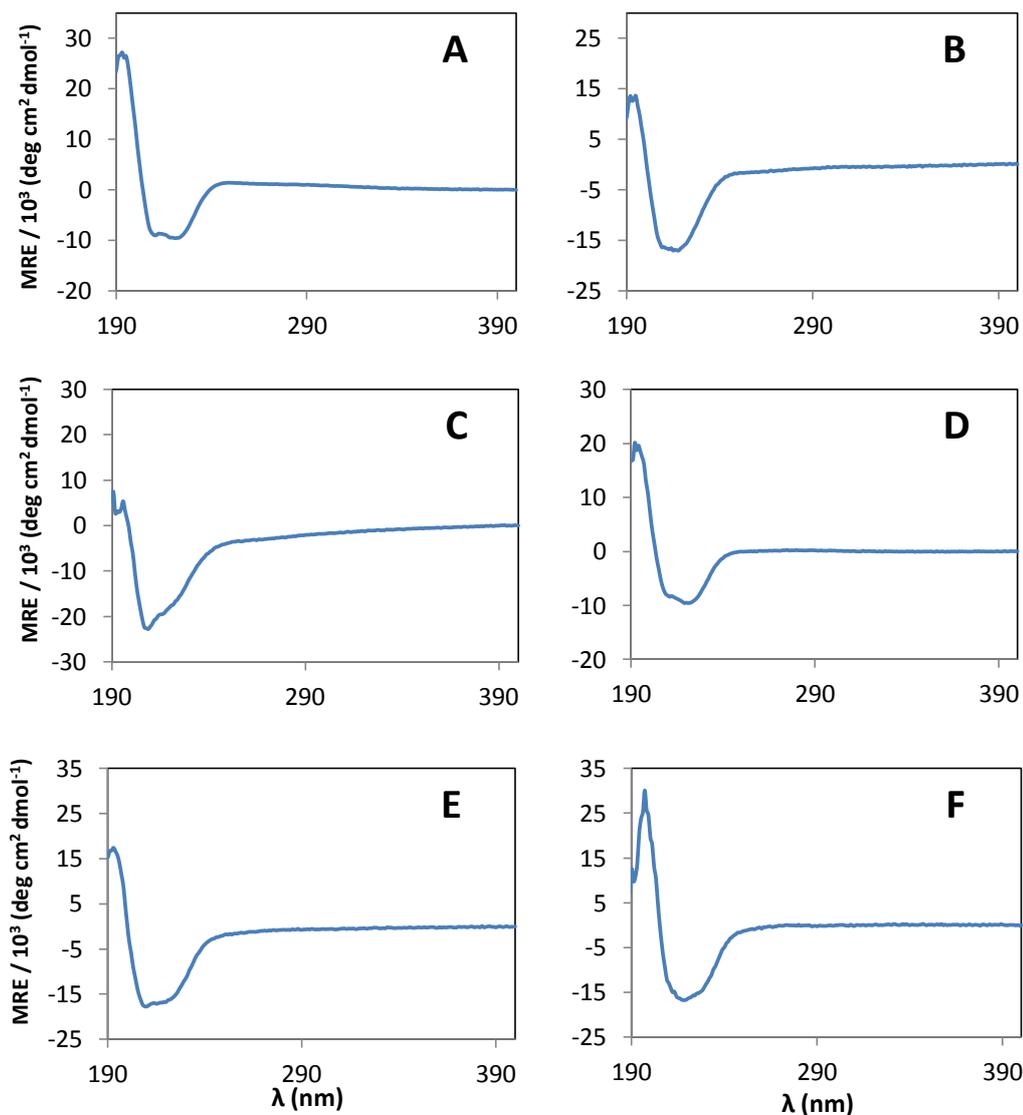
Figure 5.5 CD spectra of the DHF biosynthesis enzymes purified in this investigation. A) GTPCI, B) DHNA, C) HPPK, D) DHPS, E) DHFS and F) XGPRT.

5.4   Development of the DHF biosynthesis pathway

Attempts were made to optimise the activities of the recombinant enzymes, before assembling them to synthesise DHF. Analytical techniques, including HPLC, ion-exchange chromatography, NMR and ESI-MS, were used to characterise each enzyme. Reactions involving xanthine-guanine phosphoribosyltransferase, GTP cyclohydrolase I and dihydroneopterin aldolase have already been optimised. However, the final three steps of the pathway have yet to be optimised on a large

scale. Nevertheless, preliminary data showed promising results, and these enzymes can likely be used for generating a 'DHF factory'.

5.4.1 Xanthine-guanine phosphoribosyltransferase (XGPRT)

PRPP synthetase and XGPRT are the two enzymes in the GTP biosynthetic pathway that are not commercially available. Fortunately, we received the gene of PRPP synthetase cloned into a lac operon/T7 expression vector as a kind gift of Professor Jamie Williamson from The Scripps Research Institute. This enzyme was isolated using the procedure developed by his group (255-257). XGPRT in turn was cloned from the chromosome of *E. coli* K12 and purified as described above.
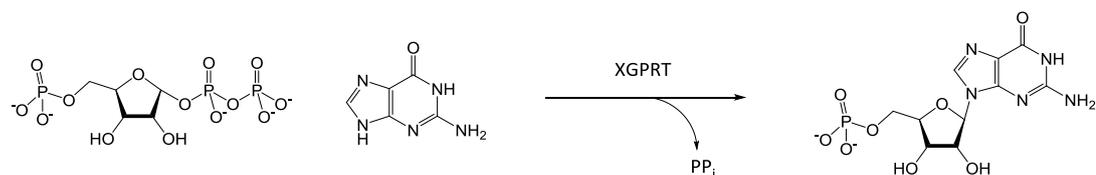
Scheme 5.4 Reaction catalysed by XGPRT.

To synthesise GTP, the starting materials ribose-5-phosphate and ATP were incubated with these enzymes in K_iPO_4 buffer (50 mM, pH 8), supplemented with 50 mM KCl, 20 mM $MgCl_2$ for 3 hours at 37 °C. The system was also included an ATP recycling machinery to prevent product inhibition of PRPP synthetase (256, 257).

The low solubility of guanine at neutral pH presented a major problem in this investigation. While it dissolves readily in 1 M NaOH, guanine precipitated immediately when introduced to the reaction buffer at pH 8, which was optimal for the activity of PRPP synthetase. To circumvent the problem, the nitrogenous base was maintained at a concentration below the precipitation limit, by adding small aliquots in 10 min intervals. After incubation, excess precipitated guanine base was removed by centrifugation. The crude was then subjected to anion-exchange chromatography, and a sharp single peak at 260 nm corresponding to GMP was observed in the trace (Figure 5.6).
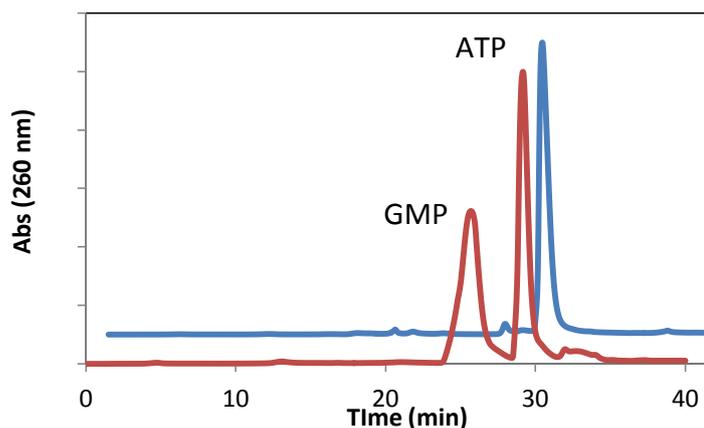
Figure 5.6 Chromatogram showing elution of the XGPRT reaction (red) and the blank reaction without XGPRT (blue).

Purified GMP was analysed by ^{1}H-NMR spectroscopy (Figure 5.7). In the aromatic region, a singlet corresponding to the 8-H on the purine ring was observed at 8.2 ppm. The characteristic anomeric 1'-H is clearly shown as a doublet at 5.9 ppm. The multiplets in the 3.9-4.5 ppm region originate from the ribose signals at C-2',-3' and -4'. However, the signal to noise ratio is low due to the presence of salt-hydrate complex. On close inspection the signals at 5.1 and 5.3 ppm are in fact two triplets, and thus is an impurity as it does not correspond to any proton in the structure of GMP. Additionally, the ESI-MS of the XGPRT reaction crude showed a peak at 363.97 m/z, which matches well with the (M+1) state of GMP (calculated MW = 363.22).
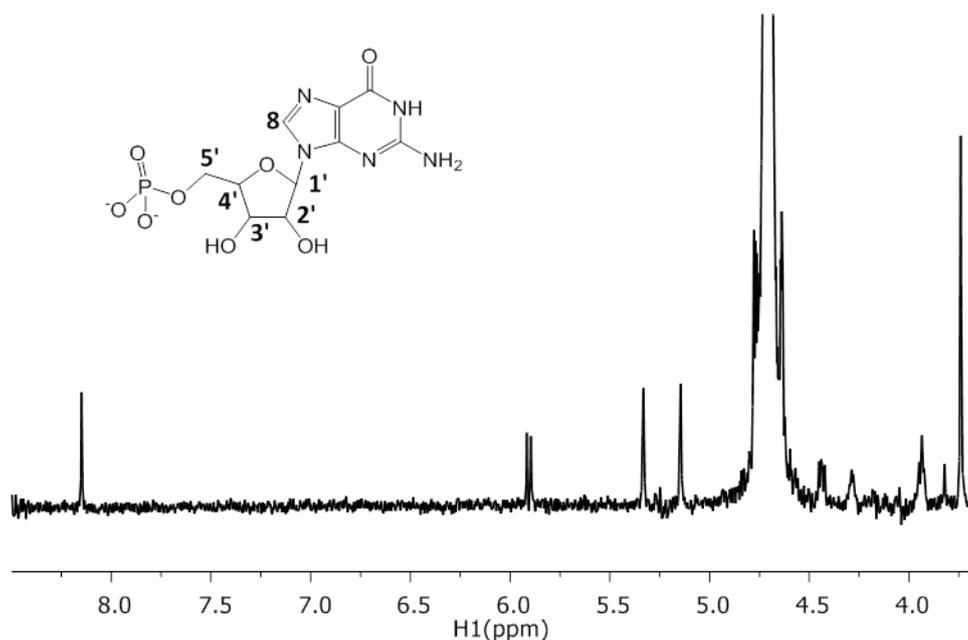
Figure 5.7 ^{1}H-NMR spectrum of the product purified from XGPRT reaction. (300 MHz, D$_2$O)
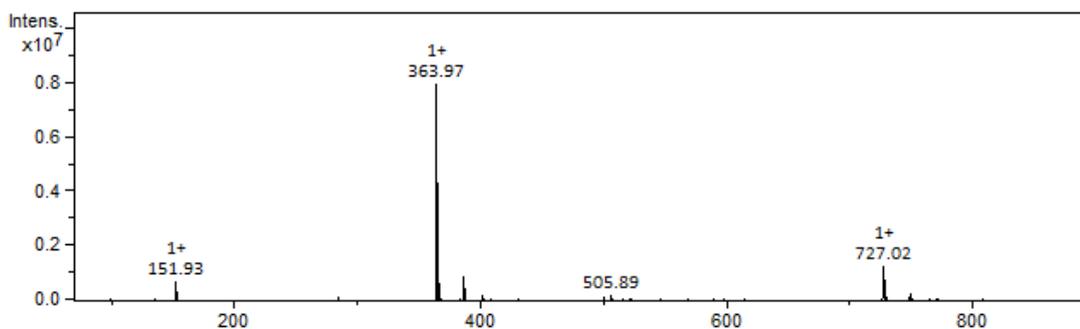
Figure 5.8 ESI-MS of GMP (MW: 363.22)

5.4.2 GTP cyclohydrolase I (GTPCI)

The first committed step of DHF biosynthesis transforms GTP to dihydroneopterin triphosphate under the action of GTP cyclohydrolase I (GTPCI) (Scheme 5.5). This reaction was reported to be rate-limiting in the pathway, and previous studies evaded this problem by using large amounts of enzyme (>30 mg freeze-dried enzyme/mL) (220). Additionally, GTPCI requires Zn^{2+} for reaction (270), but the enzyme's affinity to this cofactor is unknown. To optimise the conditions, the GTPCI reaction was carried out at various pH values (7-9) with or without the addition of

ZnCl$_2$. The enzyme was found to be active under all the tested conditions, but the highest yield obtained using a phosphate buffer (50 mM) at pH 7.

The addition of ZnCl$_2$ (1 mM) did not cause an observable effect on the reaction yield. This could be that the cationic cofactor was tightly bound to the enzyme. Alternatively, the cysteine residues responsible for cofactor binding were oxidised during the purification. Indeed, the addition of β-mercaptoethanol in the purification and reaction buffers increased the yield, as demonstrated by Antonio Angelastro who currently continues the development of this project. The production of dihydroneopterin triphosphate was traced by absorbance reading at 330 nm (ε = 6300 M^{-1} cm^{-1}, Figure 5.9), and a final yield of 58% was obtained.
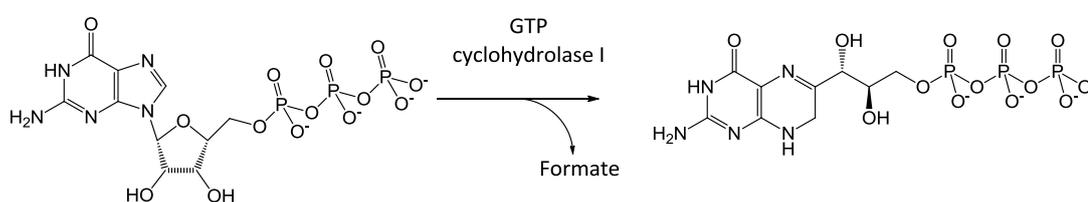
Scheme 5.5 Reaction catalysed by GTPCI.
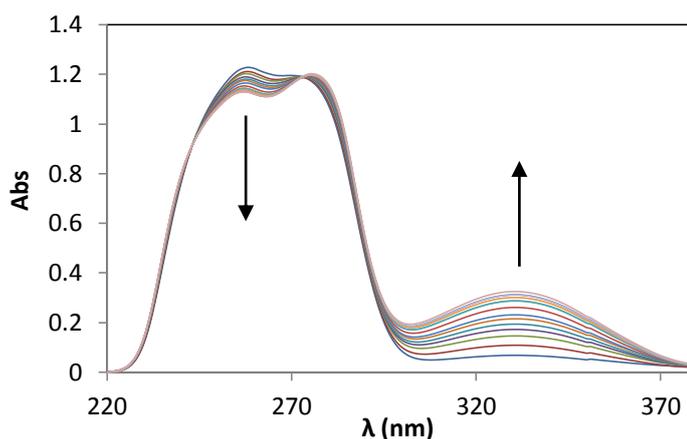
Figure 5.9 UV-Vis spectrum of the conversion of GTP to dihydroneopterin triphosphate. Measurements were taken every 5 minutes.

Although relatively good yields of dihydroneopterin triphosphate were achieved, we experienced a number of difficulties during purifications. HPLC did not seem to be applicable, when trying to separate the product from the starting material GTP,

both of which contain a triphosphate moiety. Anion-exchange chromatography offered marginally improved separation, however it was impossible to completely separate the starting material from the product. Furthermore, the product is prone to oxidation at C-7 and N-8 positions over time, and neopterin triphosphate is often observed during purification. This was indicated by the two maxima at 330 nm in the UV-Vis; a clear (M-H) peak 492.33 m/z corresponding to the oxidised product (MW = 493.17) in the negative ESI-MS study. As well as this, the (M-H) state of GTP (522 m/z, MW = 523.18) was present, highlighting the difficulty of purification (Figure 5.11). At this stage it is not certain when the product oxidation occurred. The presence of a reducing agent may decrease the rate of oxidation and thus is a possible avenue of future investigation.
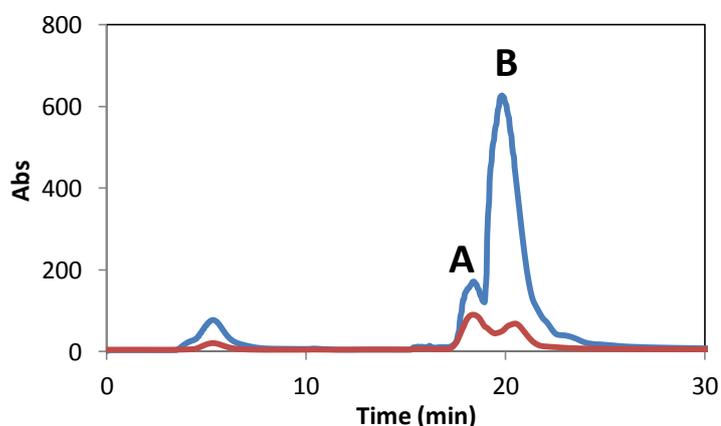
Figure 5.10 Chromatogram of the purification of the GTP cyclohydrolase I reaction at 260 nm (blue) and 330 nm (red). A and B correspond to the ESI-MS below (Figure 5.11).
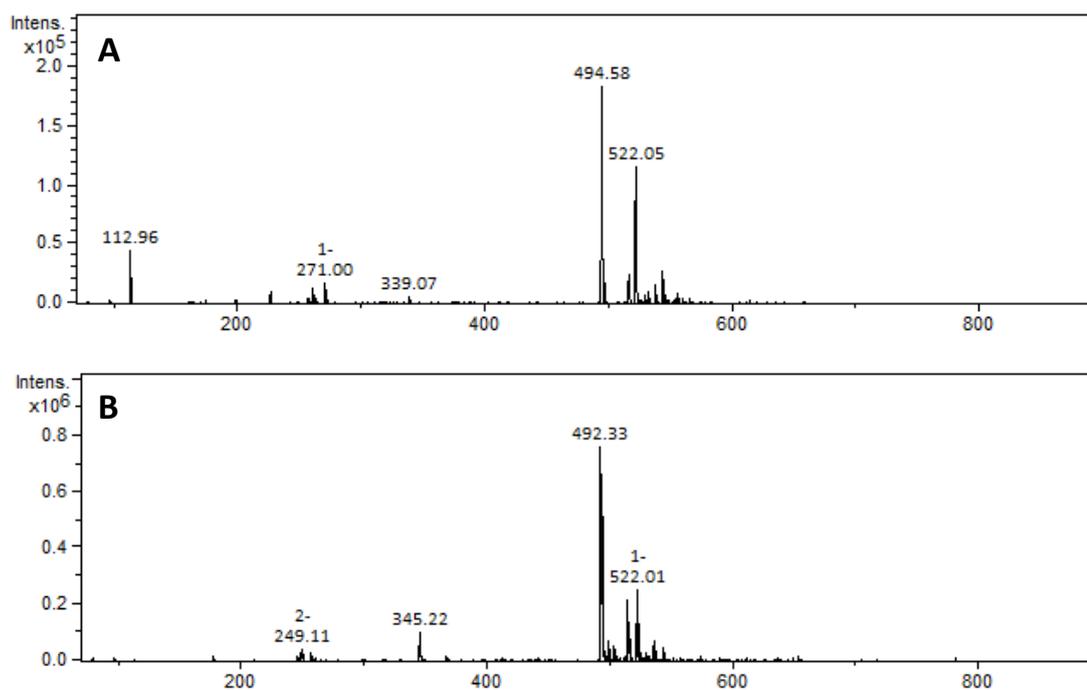
Figure 5.11 ESI-MS of dihydroneopterin triphosphate (A, MW = 495.17) and the oxidised form, neopterin triphosphate (B, 493.15). In both cases unreacted GTP (MW = 523.18) is also present.

5.4.3 Dihydroneopterin aldolase

Following the GTP rearrangement reaction, the phosphate groups from the intermediate are hydrolysed to dihydroneopterin. It remains unclear which enzyme is responsible for this reaction in nature, however bovine alkaline phosphatase (BAP), a commercially available alkaline-resilient enzyme, was used as the catalyst in laboratory. The intermediate dihydroneopterin underwent a retro-aldol cleavage reaction to produce 6-hydroxymethyl-dihydroneopterin, under the action of dihydroneopterin aldolase (DHNA) (Scheme 5.6).
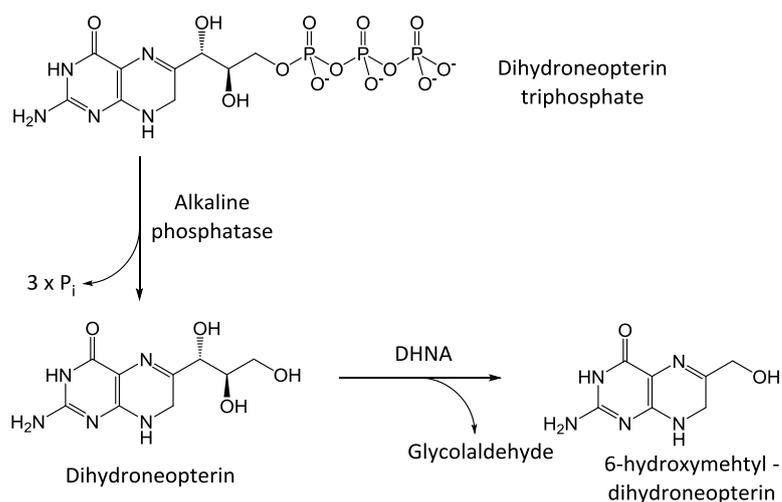
Scheme 5.6 Reaction catalysed by DHNA.

Because of the difficulties in purifying dihydroneopterin triphosphate, it was decided to perform the phosphatase and aldolase reactions simply using the GTPCI reaction crude. GTPCI was removed by filtering the reaction through MW 10000 spin-column, and the phosphatase and aldolase were added. As BAP requires a divalent cation as a cofactor, 1 mM of $ZnCl_2$ was added in the alkaline buffer (50 mM Tris, pH 9) and the reaction was incubated at 37 °C for 4 hours. A C-18 column (semi-prep, 0.22 μM particle size) was used in HPLC purification of the reaction crude (Figure 5.12). Likewise, a number of problems presented during this process. Two product peaks corresponding to neopterin and 6-hydroxymethyl neopterin were seen in the HPLC trace (A and B, respectively). Unfortunately, oxidation of these compounds was observed. The ESI mass spectrum also showed the peaks for neopterin (251.90 m/z) and 6-hydroxymethyl-neopterin (193.90 m/z), which matches well with the calculated MW of these compounds (253.08 and 195 respectively). At this time it is not clear if DHNA is able to react with neopterin or if the tetramer form of DHNA purified in this investigation is affecting the reaction in any way.
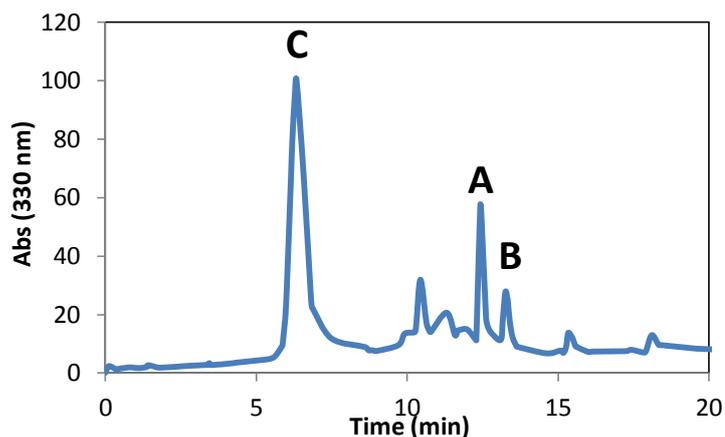
Figure 5.12 Chromatogram showing the purification of DHNA reaction purified via HPLC. A) Neopterin, B) 6-hydroxymethyl-neopterin and C) solvent front.
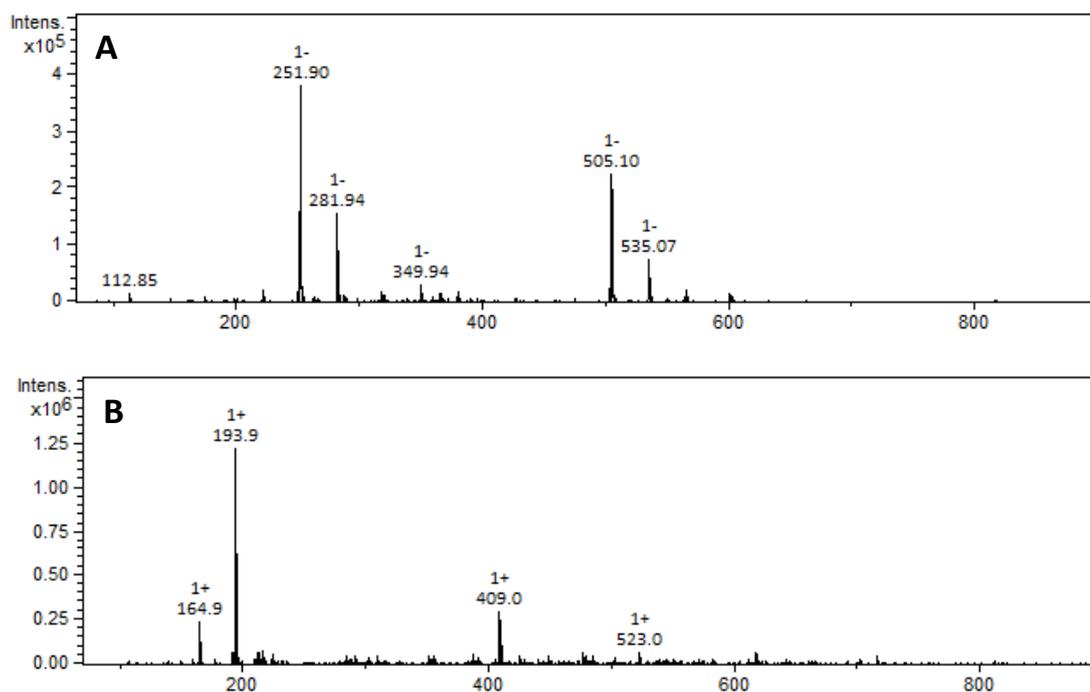
Figure 5.13 ESI-MS of neopterin (A, MW = 253.08) and 6-hydroxymethyl neopterin (B, MW = 193.06). In both cases the dihydroneopterin ring has been oxidised to neopterin.

5.4.4 Further optimisation of the DHF pathway

Continuation of the work described here, by Antonio Angelastro, has led to the production of unlabelled DHF in a 2 pot synthesis from GTP. This indicates that the recombinant HPPK, DHPS and DHFS purified in this investigation are indeed active.

162

5.5 Conclusion

The DHF biosynthesis enzymes have been cloned and purified successfully from *E. coli* chromosomal DNA. The CD spectroscopic analysis indicated that the purified enzymes exhibited properly folded secondary structures. Interestingly, GTPCI, DHNA and HPPK were found to adopt quaternary structures different from previously observed oligomerisation states. The latter two enzymes are affected by the N-terminal histidine fusion tag used during Ni^{2+}-chromatographic purifications. It is still unclear why GTPCI forms a trimer rather than a decamer, as previously reported. Although purification of the intermediate dihydroneopterin triphosphate was unsuccessful, the reaction crude could be used for the subsequent phosphatase and retro aldol cleavage reactions. Nevertheless, it is unclear if the enzymes are operating at a maximum rate at this stage.

For the GTP biosynthetic pathway, XGPRT was cloned and purified, thus a *de novo* biosynthesis of this nucleotide can be prepared in the laboratory. Even though XGPRT was reported to adopt a tetrameric complex, our studies indicated a dimeric enzyme. Nevertheless, this enzyme was shown to be active for catalysis.

Overall the DHF biosynthesis enzymes have successfully been cloned and purified. The DHF biosynthesis pathway has begun to be developed in the laboratory, although further optimisation of these enzymes is required. Antonio Angelastro, the Ph.D. candidate who continues this project, could indeed synthesise DHF from this pathway and showed major improvement on the reaction yields following further optimisations.

6   GENERAL CONCLUSIONS

This thesis aimed to elucidate the role of femtosecond (fs), short-range bond vibrations in enzymatic reactions. Two distinct approaches were employed. These include the kinetic investigation of a "heavy", isotopically labelled DHFR as well as heavy isotope effect measurements of an NADPH isotopologue labelled at the reaction centre.

6.1   The role of fs bond vibrations in the enzyme on catalysis

Primary KIE measurement of BsDHFR has led to the discovery of an interesting biphasic Arrhenius plot. Similar trends have been observed in other thermophilic enzymes, such as BsADH (21) and TmDHFR (24), and the strong temperature dependence at low temperature can be viewed as a consequence of poor conformational sampling. Alternatively, the biphasic behaviour could be explained by the existence of multiple conformers, in which equilibrium favourable for hydride transfer is temperature dependent. Future NMR studies could help identify the dynamic nature of BsDHFR and the possible existence of different conformers at low temperature.

BsDHFR was further characterised by enzyme isotope substitution. The mesophile EcDHFR, its catalytically compromised mutant and the hyperthermophile TmDHFR have also been investigated by the same approach (117, 118, 236). Steady-state kinetic analysis produced a relatively large enzyme KIE (~2.6), and this can be reasoned by the highly flexible nature of BsDHFR. On the basis of the temperature independence, a physical step in the BsDHFR catalysis is likely rate limiting, but it does not require a large conformational change (117, 236).

At physiologically relevant temperatures, the enzyme KIE on hydride transfer is insignificant but increases substantially at low temperature. Based on previous *in silico* studies on EcDHFR (117), we postulated that there are greater dynamic effects at low temperature. Preliminary EA-VTST calculations performed by our collaborators confirmed this. We also suggested that the intrinsic flexibility of BsDHFR contributes to the large enzyme KIE. At low temperature there is a greater entropic burden in the process of electrostatic preorganisation, thus leading to an

increase in reorganisation energy needed for hydride transfer. While this requires additional computational study, the lack of enzyme KIE^H at high temperatures indicates the dynamic coupling is minimal at physiologically relevant conditions. This agrees well with the hypothesis, which stated that efficient enzymes tend to minimise the coupling of protein motions to the chemical coordinate, as it can cause an anti-catalytic effect. Overall, there seems to be no evolutionary advantage towards direct coupling of these ultrafast motions to enzyme catalysis in all DHFRs.

6.2 The role of fs bond vibrations in the substrate in hydride transfer reactions

α-Secondary hydrogen (2°) KIEs report on the effect of rehybridisation upon hydride transfer (28), and have been suggested to be insensitive to preorganisation or reorganisation effects that occur prior to the reaction. Instead, they report solely on the reaction ready configuration (RRC) adopted immediately prior to hydride transfer (153, 155). 2° KIEs were measured for EcDHFR, MpDHFR and BsDHFR with the intention of determining a general trend for all DHFR enzymes. Even though these enzymes displayed greatly different temperature dependence on the primary hydride KIE, all of them have essentially the same magnitude of 2° KIE and are temperature independent. These findings imply that there are no significant differences in the RRC across the experimental temperature range, and suggest that the RRC of the substrates is very similar in all members of this enzyme family (155).

NADPH labelled with a ^{13}C at the 4-position was used in heavy atom isotope effect measurement. Between 15 and 35 °C, the ^{12}C/^{13}C KIE on hydride transfer is approximately 1.03. Interestingly, it increases sharply to 1.10 at 5 °C. To the best of our knowledge, this is the largest reported carbon isotope effect on hydride transfer. There is a greater ^{12}C/^{13}C isotope effect observed on deuteride transfer (1.15 at 5 °C), likely due to the greater reduced mass increase. This magnitude and temperature dependence cannot be explained using semi-classical theory, as simple ZPE differences between the stretching or bending vibrational frequencies of isotopologues cannot explain the sharp increase at low temperature (28).

[13]C-labelling may affect either the tunnelling probability or the frequency of recrossing events. When applying a Marcus-like model, the increased magnitude of isotope effect at low temperature could be due to weakening of the "promoting" motion that modulates the donor acceptor distance. This would result in a wider reaction barrier and an increased isotope effect. Alternatively, it could result from an increase in the dynamic recrossing coefficient as a consequence of fs vibrational coupling. Lastly, the ^{13}C labelling could affect the reorganisation of the RRC, which is theoretically isotopically sensitive. The observed ^{12}C/^{13}C on hydride transfer implies that the proposed reorganisation is at least partially rate determining at low temperature. At high temperature hydride transfer is rate limiting, thus "normal" carbon isotope effects are measured.

These possibilities were probed further by measuring the carbon isotope effect on the primary KIE. [13]C labelling has no apparent effect on the primary hydrogen KIE, showing values that are essentially the same as the ones obtained from ^{12}C labelled cofactors. This indicated that any hypothetical "promoting" motion is unlikely exerting its effect via the C4 atom. Future experimental and computational studies will focus on any potential difference that could be masked by the limitations of stopped-flow spectroscopic measurement.

Overall, fs bond vibrations in the substrate surrounding the reaction centre do couple to enzymatic reactions. However, at present they do not seem to directly couple to the barrier width and/or height. Hence, they cannot be classed as "promoting motions".

6.3   Reconstitution of the NADP[+] biosynthesis pathway

Most bacterial species synthesise NADP[+] from aspartate and dihydroxyacetone phosphate, in a biosynthetic pathway involving six enzymes (194). Previous studies identified the *nad* gene cluster in *E. coli*, and each protein was characterised biochemically (200, 201, 208, 211, 214, 216, 217). Accordingly, this pathway can be an effective tool towards the production of various isotopically labelled NADP[+] analogues, particularly those carrying labels in the nicotinamide moiety.

The genes encoding the enzymes in the biosynthesis of NADP$^+$ were cloned from *E. coli* K12 substrains and ligated into expression vectors. Each enzyme was purified using Ni^{2+} affinity chromatography, and subsequent size-exclusion chromatography and CD spectroscopy analysis confirmed successful production of native protein. With the addition of commercially available glycolysis enzymes, [1,6-^{13}C$_2$]-glucose was converted to [3-^{13}C]-dihydroxyacetone phosphate and then to [4-^{13}C]-NADP$^+$. Additionally, these enzymes were used to produce [amino-^{15}N]-NAD$^+$ and ^{15}N-quinolinic acid for our future studies. Our success in synthesising isotopically labelled NAD(P)H helps further advance the studies of DHFR, and it also opens the possibility to investigate other interesting enzymes that employ this cofactor.

6.4 Future approaches

This investigation has determined the role that short-range, fs bond vibrations play in the hydride transfer step catalysed by DHFR via the use of isotopically labelled enzyme and starting material, separately. While both studies cause an observable isotope effect, it would be of great interest to how a "heavy" enzyme reacts with labelled starting materials. This would enable any coupling between specific enzyme dynamics and those from the substrate to be observed. This current investigation cannot confirm the exact source for the dynamic coupling between enzyme and transition state. Hopefully, this proposed experiment will provide additional insights.

While the NADP$^+$ biosynthetic pathway is successfully reconstructed in the laboratory, we are currently focused on fully developing an efficient chemo-enzymatic synthetic pathway for DHF. Isotope labelling on position, such as C6/N5 of DHF would facilitate the characterisation of the TS structure. This in turn could lead to the development of TS analogues which could be potent inhibitors and drug candidates (271). This is particularly relevant in the case of DHFR, as it is valuable anti-bacterial and anti-malarial drug target and only few new candidates are currently available.

Lastly, we would also like to further advance our techniques in characterising the transition state of DHFR. In order to measure the heavy atom KIE in a highly precise manner required for full TS analysis (271), competitive heavy atom KIE measurements will be carried out. Alternative DHFR homologues in which hydride transfer is more rate-limiting at pH 7 will be used to minimise the problem of kinetic complexity. Other NAD(P)H-dependent enzymes, such as alcohol dehydrogenase, will also be investigated in the future.

6.5 Summary

Overall, this investigation has highlighted the role of fs bond vibrations both from the enzyme and the cofactor in DHFR catalysis. Results suggest that protein motions are crucial for the cofactor-substrate arrangement immediately prior to hydride transfer. However, there is no direct evidence that there is a short-range vibrational motion that "promotes" hydride transfer. Together with our previous works on DHFR, we concluded that dynamic effects greatly vary among this family of enzymes. While investigators often consider EcDHFR as a "paradigm" in enzymology studies, the dynamic characteristics of this enzyme are unlikely a universal phenomenon.

7   MATERIALS & METHODS

7.1 Bacterial strains and their preparation

7.1.1 Bacterial strains

Numerous cloning and expression strains of *Escherichia coli* were used in this investigation (Table 7.1).

Table 7.1 *E. coli strains* used during this investigation

E. coli strain	Use
XL1-Blue	PCR amplification of all genes
JM101	Expression of *nadK* from pNCO-nadK
BL21(DE3)	Large scale production of *E. coli* and *M. profunda* DHFR
BL21-CodonPlus(DE3)-RP	Expression of *nadA*, *B*, *C*, *folB*, *K* from pET28a expression vectors
BL21-Star™(DE3)	Expression of *nadD*, *E*, *folE*, *P*, *C*, *prsA* and *gpt* from pET expression vectors
Stratagene competent cells[A]	Production of cloning vectors from StrataClone Blunt vector kit
XL-10 Gold Kan ultra-competent cells[B]	Production of cloning vectors from pPCR-Script Amp SK (+) vector kit
A: purchased from Agilent as part of the Strataclone Blunt PCR cloning kit B: purchased from Agilent as part of the PCR-Script Amp cloning kit	

7.1.2 Competent cells

7.1.2.1 *Calcium chloride buffer I*

Calcium chloride (1.11 g) was dissolved in 80 mL of deionised water. The total volume was then taken to 100 mL. The solution was sterilised via autoclave at 121 °C for 20 minutes.

7.1.2.2 *Calcium chloride buffer II*

Calcium chloride (1.11 g) and glycerol (15% v/v) were added to 80 mL of deionised water. The total volume was then taken to 100 mL. The solution was sterilised via autoclave at 121 °C for 20 minutes.

7.1.2.3 Competent cell preparation

50 µL of the desired cell strain (Table 7.1) were incubated in 100 mL non-selective LB media overnight at 37 °C whilst shaking. In the case of XL1-Blue tetracycline was added to the media (50 mg/L). 1 mL of the overnight culture was used to inoculate 100 mL of non-selective LB media (50 mg/L tetracycline in the case of XL1-Blue) until an optical density of 0.6 at 600 nm was obtained. The cells were placed on ice for 15 minutes then harvested via centrifugation at 3400 g and the supernatant discarded. The cells were re-suspended in 40 mL calcium chloride buffer I (Section 7.1.2.1), incubated on ice for 15 minutes before the centrifugation step was repeated. The cells were re-suspended in 4 mL calcium chloride buffer II (Section 7.1.2.2), divided into 50 µL aliquots and frozen immediately in liquid nitrogen.

7.1.3 Super-competent cells

7.1.3.1 Rubidium chloride buffer I

Potassium acetate (0.294 g). rubidium chloride (1.21 g), calcium chloride (1.11 g), manganese chloride (0.692 g) and glycerol (15% v/v) were dissolved in 80 mL deionised water. The pH was adjusted to 5.8 using dilute acetic acid and the volume taken to 100 mL with deionised water. A 0.2 µm syringe filter was used to sterilise the solution and then stored at 4 °C.

7.1.3.2 Rubidium chloride buffer II

3-(N-morpholino)propanesulfonic acid (0.209 g), calcium chloride (0.832 g), rubidium chloride (0.121 g) and glycerol (15% v/v) were dissolved in 80 mL deionised water. The pH was adjusted to 6.5 using sodium hydroxide and the volume taken to 100 mL with deionised water. A 0.2 µm syringe filter was used to sterilise the solution and then stored at 4 °C.

7.1.3.3 Super-competent cell preparation

Super competent cells were prepared as with competent cells (Section 7.1.2.3) but with rubidium chloride buffer I and II (Section 7.1.3.1) in replace of calcium chloride buffer I and II (section 7.1.3.2) respectively.

7.2 Growth media and sterile solutions

7.2.1 Luria-Burtani (LB) medium

Tryptone (30 g), yeast extract (15 g) and NaCl (30 g) were dissolved in 2.5 L of deionised water and the final volume taken to 3 L. The solution was sterilised in an autoclave at 121 °C and stored at 4 °C. Immediately before use the required antibiotic was added under sterile conditions.

7.2.2 LB agar plates

Tryptone (10.0 g), yeast extract (5.0 g), sodium chloride (10.0 g) and agar (15.0 g) were dissolved in 0.9 L of deionised water and the total volume taken to 1 L. The solution was sterilised in an autoclave at 121 °C for 20 minutes. When cool enough to hold by hand the required antibiotic was added and the LB agar poured into plates under sterile conditions. The solution was allowed to cool and the plates stored at 4 °C.

7.2.3 Isotopically labelled media

7.2.3.1 Trace elements

EDTA (5 mg) was dissolved in 800 mL deionised H_2O and the pH adjusted to 7.5. $FeCl_3.6H_2O$ (0.83 g), $ZnCl_2$ (84 mg), $CuCl_2.2H_2O$ (13 mg), $CoCl_2.6H_2O$ (10 mg), H_3BO_3 (10 mg) and $MnCl_2.4H_2O$ (1.6 mg) were added and the volume adjusted to 1 L with deionised H_2O. The solution was sterilised in an autoclave at 121 °C, 15 lb.(sq. in)$^{-1}$ for 20 minutes and stored at room temperature.

7.2.3.2 Minimal (M9) media

Na_2HPO_4 (6.76 g), KH_2PO_4 (3 g), NaCl (0.5 g), 1 M $MgSO_4$ (2 mL), 1 M $CaCl_2$ (100 μL) and trace elements (600 μL) were added to 900 mL [$^2H_2$, 99.9%] H_2O. [^{15}N, 98%] NH_4Cl (1 g) and [99%, $^{13}C_6$; 1,2,3,4,5,6,6-$^2H_7$, 97%] glucose (4 g) were added and the volume taken up to 1 L with [$^2H_2$, 99.9%] H_2O. The media was passed through a sterile 0.2 μm filter and stored at 4 °C. Immediately before use the required antibiotic was added under sterile conditions.

7.2.4   Sterile solutions

7.2.4.1   Ampicillin solution

Ampicillin (200 mg) was dissolved in 4 mL deionised H_2O to give a concentration of 50 mg/mL. The solution was filtered with a sterile 0.2 μm syringe filter, aliquoted and stored at -20 °C.

7.2.4.2   Kanamycin solution

Kanamycin (200 mg) was dissolved in 4 mL deionised H_2O to give a concentration of 50 mg/mL. The solution was filtered with a sterile 0.2 μm syringe filter, aliquoted and stored at -20 °C.

7.2.4.3   Tetracycline

Tetracycline (200 mg) was dissolved in 4 mL deionised H_2O to give a concentration of 50 mg/mL. The solution was filtered with a sterile 0.2 μm syringe filter, aliquoted and stored at -20 °C.

7.2.4.4   Isopropyl-β-D-1-thiogalactopyranoside solution

Isopropyl-β-D-1-thiogalactopyranoside (IPTG) (360 mg) was dissolved in 3 mL deionised H_2O to give a concentration of 120 mg/mL. The solution was filtered with a sterile 0.2 μm syringe filter, aliquoted and stored at -20 °C.

7.2.4.5   50% v/v glycerol solution

100 mL glycerol was added to 100 mL deionised H_2O to give a 50% v/v mixture of glycerol. The solution was sterilised in an autoclave at 121 °C, 15 lb. (sq. in)$^{-1}$ for 20 minutes and stored at 4 °C.

7.3   Cloning

7.3.1   Polymerase chain reaction amplification of target genes

Polymerase chain reaction (PCR) was used to amplify all the target genes using *E. coli* XL1-Blue chromosomal DNA as the template (Section 7.3.1.1). The primers (Table 7.4) were designed from sequences in the Genbank library (Table 7.2 and

Table 7.3) with restriction sites to ligate the target genes into the pNCO expression vector (*Bam*HI and *Hin*dIII sites) and pET-21a, -21d and -28a expression vectors (*Nde*I/*Bam*HI, *Nco*I/*Hind*III and *Bam*HI/*Hind*III sites). Phusion® high fidelity polymerase was used and the components (Table 7.5) and conditions of the PCR reaction optimised for this specific polymerase and each set of primers (Table 7.6 and Table 7.7). The products were visualised via agarose gel electrophoresis (Section 7.3.3).

Table 7.2 NADP⁺ biosynthesis genes and enzymes

ENZYME	GENE	NCBI GENE ID
Quinolinate synthetase	*nadA*	12934034
L-aspartate oxidase	*nadB*	947049
Quinolinate phosphoribosyltransferase	*nadC*	12934550
Nicotinic acid mononucleotide adenylyltransferase	*nadD*	12932665
NAD synthetase	*nadE*	12933235
NAD kinase	*nadK*	947092

Table 7.3 DHF biosynthesis genes and enzymes

ENZYME	GENE	NCBI GENE ID
GTP cyclohydrolase I	*folE*	949040
Dihydroneopterin aldolase	*folB*	947544
Hydroxymethyl-dihydroneopterin pyrophosphokinase	*folK*	948792
Dihydropteroate synthetase	*folP*	947691
Dihydrofolate synthetase	*folC*	945451
Xanthine-guanine phosphoribosyltransferase	*gpt*	944817

Table 7.4 Target genes and primers used for PCR amplification

Target Gene	Primers
nadA	Fwd 5'-**CATATG**AGCGTAATGTTTGATCCAGACAC-3'
	Bwd 5'-**AAGCTT**ATTATCCACGTAGTGTAGCCGC-3'
nadB	Fwd 5'-**CATATG**AATACTCTCCCTGAACATTCATGTGA-3'
	Bwd 5'-**AAGCTT**ATTATCTGTTTATGTAATGATTGCCGGG-3'
nadC	Fwd 5'-**CATATG**CCGCCTCGCCGCTATAA-3'
	Bwd 5'-**AAGCTT**ATTAGCGAAAACGCATTGAAAGGTC-3'
nadD	Fwd 5'-GTCGAGAAT**GGATCC**ATGAAATCTTTCACG-3'
	Bwd 5'-CTGCGATGTG**AAGCTT**TCCGGGTAC-3'
nadE	Fwd 5'-CTTTTCTGTCTGGAGG**GGATCC**ATGACATTGCAAC-3'
	Bwd 5'-GGCGTGAACAAAGT**AAGCTT**TTTCACACAAACC-3'
nadK	Fwd 5'-GGACCTC**GGATCC**ATGAATAATCATTTCAAG-3'
	Bwd 5'-GAGGCTGG**AAGCTT**ATTAGAATAATTTTTTTGACC-3'
folB	Fwd 5'-**CATATG**GATATTGTATTTATAGAGCAACTTTC-3'
	Bwd 5'-**GGATCC**TTATTAATTATTTTCTTTCAGATTATTGCC-3'
folC	Fwd 5'-GCGGAA**CATATG**ATTATCAAACGC-3'
	Bwd 5'-CCACTAACCGA**GGATCC**TTATTACTTGCCAC-3'
folE	Fwd 5'-**CCATGG**ATGCCATCACTCAGTAAAGAAGC-3'
	Bwd 5'-**AAGCTT**TTATCAGTTGTGATGACGCACAGC-3'
folK	Fwd 5'-GGGTAC**CATATG**ACAGTGG-3'
	Bwd 5'-CGCAAT**GGATCC**TTATTACCATTTGTT-3'
gpt	Fwd 5'-**CATATG**AGCGAAAAATACATCGTCACCT-3'
	Bwd 5'-**AAGCTT**ATTAGCGACCGGAGATTGG-3'
glyA	Fwd 5'-**CATATG**TTAAAGCGTGAAATGAACATTGCCGATTAT-3'
	Bwd 5'-**AAGCTT**ATTATGCGTAAACCGGGTAACGTGC-3'
Restriction sites are shown in bold.	

7.3.1.1 E. coli *XL1-Blue chromosomal DNA*

E. coli XL1-Blue cells (~ 40 μL) were defrosted slowly on ice, streaked onto a tetracycline selective LB agar plate and incubated overnight at 37 °C. A single colony was then dipped into each PCR reaction mixture (Table 7.5) to provide a source of *E. coli* XL1-Blue chromosomal DNA.

Table 7.5 Components of the PCR reactions

Component	Volume (µL)
dNTPs (40 mM total, 10 mM each NTP)	1
Forward primer (25 µM)	1
Backward primer (25 µM)	1
E. coli XL1-Blue chromosomal DNA	One colony*
Phusion polymerase GC buffer	10
$MgCl_2$	2 µL
Sterile dH_2O	Make up to 50 µL
Phusion® high fidelity polymerase (1 unit/µL)	0.5
* One colony per reaction was scraped from a plate and dipped into the PCR mixture	

Table 7.6 Conditions used for PCR amplification of NADPH biosynthesis genes

Step	Target gene			
	nadA	*nadB, nadC*	*nadD, nadE*	*nadK*
Initial denaturation	180 s, 98 °C			
Denaturation*	10 s, 98 °C			
Annealing	60 s, 69 °C	60 s, 67 °C	30s, 44.5 °C	30 s, 65 °C
Extension	45 s, 72 °C	90s, 72 °C	45 s, 72 °C	30 s, 72 °C
Final Extension	600 s, 72 °C			
* The denaturation, annealing and elongation step was cycled 35 times, except for *nadD* which was cycled 40 times.				

Table 7.7 Conditions used for PCR amplification of dihydrofolate biosynthesis genes

Step	Target gene			
	folB	*folC, folK*	*folE*	*gpt*
Initial denaturation	180 s, 98 °C			
Denaturation*	10 s, 98 °C			
Annealing	30 s, 42 °C	30 s, 43 °C	30 s, 67 °C	60 s, 69 °C
Extension	45 s, 72 °C			
Final Extension	600 s, 72 °C			
* The denaturation, annealing and elongation step was cycled 35 times				

7.3.2 Site-directed Mutagenesis

Site-directed mutagenesis was used to insert or alter a number of DNA bases for two genes. Specific primers (Table 7.8) were designed to insert one base into the *folE* gene to shift the reading frame, and to introduce three different amino acids into the *folB* gene. 39 μL sterile dH$_2$O (sterilised in an autoclave at 121 °C, 15 lb.(sq. in)$^{-1}$ for 20 minutes), 5 μL Pfu-polymerase buffer, 1 μL dNTPs, 2 μL forward and backward primers (Table 7.8), 1 μL template DNA and 1 μL Pfu-polymerase were added to a PCR tube and underwent site-directed mutagenesis (Table 7.9). After the site-directed mutagenesis was complete the reaction incubated at 37 °C with 1 μL *Dpn*I for 1 hour. The reaction was cooled on ice before transformation into XL1-Blue *E. coli* cells.

Table 7.8 Primers for site-directed mutagenesis

Gene Mutation	Primers
pE21d-folE (frame shift)	Fwd 5'-GGAGATATACCATGATGCCATCACTCAG-3'
	Bwd 5'-CTGAGTGATGGCATCATGGTATATCTCC-3'
pET28a-folB K38N	Fwd 5'-GAAATGGCGTGGGATAACCGTAAAGCGGCGAAAAG-3'
	Bwd 5'-CTTTTCGCCGCTTTACGGTTATCCCACGCCATTTC-3'
pET28a-folB T70A	Fwd 5'-GGGGCGCGTTTTGCGCTGGTGGAAC-3'
	Bwd 5'-GTTCCACCAGCGCAAAACGCGCCCC-3'
pET28a-folB A108V	Fwd 5'-CGCGGGCGGCGAATGTTGGCGTAATCATTG-3'
	Bwd 5'-CAATGATTACGCCAACATTCGCCGCCCGCG-3'

Table 7.9 Conditions used for site-directed mutagenesis

Step	Temperature (°C)	Time (min)
Initial denaturation	95	3
Denaturation*	95	1
Annealing	55	2
Elongation	72	10
Final Elongation	72	10
* The denaturation, annealing and elongation step was cycled 15 times		

7.3.3   DNA visualisation

7.3.3.1   TAE buffer

Tris base (242 g), glacial acetic acid (57.1 mL) and 100 mL EDTA (0.5 M, pH 8.0) were dissolved in 900 mL deionised water. The total volume was taken to 1 L with deionised water. Immediately prior to use the solution was diluted in a 1:49 ratio with deionised water.

7.3.3.2   DNA loading buffer

Glycerol (3 mL) and 0.5% bromophenol blue solution (2 µL) were added to 7 mL deionised water.

7.3.3.3 *Agarose gel electrophoresis*

Agarose gels were prepared using dilute TAE buffer (Section 7.3.3.1) in a 1% w/v ratio. 5 µL DNA loading buffer (Section 7.3.3.2) was added to the DNA sample which was then loaded onto the gel and run at a constant voltage of 100 V for 1 hour before being placed in a dilute ethidium bromide solution for 10 minutes whilst shaking. The DNA bands were visualised using a Syngene GeneFlash UV light box (Syngene, Cambridge, UK) and compared to a molecular weight ladder run on the same gel.

7.3.4 Cloning of PCR products into holding vectors

The NADPH and DHF synthesis genes were cloned into pPCR-Script (*nadD* and *nadE*) and Strataclone holding vectors (all others) directly following the PCR amplification.

7.3.4.1 *Cloning into pPCR-Script vector*

pPCR-Script Amp SK(+) cloning vector(1 µL),PCR-Script 10× reaction buffer (1 µL), 10 mM rATP (0.5 µL), PCR product (2 µL), *Srf*I restriction enzyme (5 U/µL, 1 µL), T4 DNA ligase (4U/µL, 1 µL) and distilled water (3.5 µL) were incubated for 1 hour at room temperature. The reaction was then heat shocked for 10 minutes at 65 °C and the reaction frozen for later use or immediately used to transform ultra-competent cells (Section 7.3.8).

7.3.4.2 *Cloning into StrataClone PCR cloning vector*

StrataClone Blunt Cloning Buffer (3 µL), PCR product (2 µL) and StrataClone Blunt Vector Mix amp/kan (1 µL) were incubated at room temperature for 5 minutes and the reaction frozen for later use or immediately used to transform ultra-competent cells (Section 7.3.8).

7.3.5 DNA purification

7.3.5.1 *Miniprep of DNA*

A single colony from an agar plate, or a spike from a glycerol stock, harbouring the required transformed cells was used to inoculate 10 mL of LB medium (containing the appropriate selective antibiotic) and incubated overnight at 37 °C whilst

shaking. Cells were harvested by centrifugation at 3400g for 5 minutes. The QIAprep spin miniprep kit (QIAGEN, Crawley, UK) was used with EconoSpin All-in-1 mini spin columns (Epoch Biolabs, Inc, TX, USA) to purify the plasmid following the manufacturer's protocol.

7.3.5.2 *Agarose gel DNA purification*

The required samples underwent agarose gel electrophoresis (Section 7.3.3.3) and the DNA visualised by staining in dilute ethidium bromide solution. DNA bands were visualised using a Syngene GeneFlash UV light box (Syngene, Cambridge, UK) and the appropriate DNA bands cut from the agarose gel using a razor blade. DNA was extracted from the agarose gel using a QIAquick gel extraction kit (QIAGEN, Crawley, UK) with EconoSpin All-in-1 mini spin columns (Epoch Biolabs, Inc, TX, USA) following the manufacturer's protocol.

7.3.6 Restriction digest of DNA

Manufacturer's guidelines regarding temperature, buffers and denaturation were followed in all restriction enzyme digests (Table 7.10). Restriction digests involving *Nde*I were performed sequentially with the required restriction enzymes for 4 hours each at 37 °C where possible. *Nde*I was heat inactivated at 60 °C for 20 minutes before addition of the second enzyme. In cases where neither enzyme could be denatured the first restriction enzyme was incubated for 4 hours at 37 °C followed by the addition of the second restriction enzyme for a further 4 hours at 37 °C. Restriction enzymes were removed via agarose gel purification (Section 7.3.3).

Table 7.10 Enzymes used for restriction digestion of DNA

Gene (vector)	Restriction enzymes
nadA, nadB, nadC, gpt (pET28a)	*Nde*I and *Hind*III
nadD, nadE (pET28a)	*BamH*I and *Hind*III
nadK (pNCO)	*BamH*I and *Hind*III
folB, folK (pET28a)	*Nde*I and *BamH*I
folC, folP (pET21a)	*Nde*I and *Hind*III
folE (pET21d)	*Nco*I and *Hind*III

7.3.7 Ligation of DNA into expression vectors

Following restriction digest and agarose gel purification, the required gene and expression plasmid were ligated to form new expression vector constructs. Plasmid and gene in a 1:5 ratio was incubated with 2 μL T4 DNA ligase buffer and sterile dH$_2$O (sterilised in an autoclave at 121 °C, 15 lb.(sq. in)$^{-1}$ for 20 minutes) in a volume of 20 μL. 1 μL T4 DNA ligase was added and the ligation reaction incubated for 2 hours at 25 °C. The reaction was frozen for later use or immediately used to transform ultra-competent cells (Section 7.3.8).

7.3.8 Transformation of competent cells

The appropriate competent cells and plasmid DNA were allowed to thaw slowly on ice. 2 μL plasmid DNA solution (5-10 μL in the case of ligation reactions) was transferred to the competent cells and mixed gently. The cell/DNA mixture was incubated on ice for at least 20 minutes, heat shocked at 42 °C for 45 seconds and returned to ice for at least 2 minutes. Non-selective LB medium (1 mL) was added to the cell/DNA mixture and incubated at 37 °C for 1 hour whilst shaking for plasmids with ampicillin resistance and 1 hour 15 minutes at 37 °C whilst shaking for plasmids with kanamycin resistance. Cells were harvested by centrifugation for 1 minute at 3300 g (Eppendorf centrifuge 5415R) and the supernatant solution discarded. The cell pellet was re-suspended in 100 μL of LB medium, plated on agar plates containing the appropriate antibiotic and incubated at 37 °C overnight.

7.4 Gene expression

7.4.1 Unlabelled enzymes

E. coli cells containing a plasmid harbouring the target gene were incubated overnight in LB medium (100 mL) at 37 °C overnight whilst shaking and used to inoculate LB medium (5 mL of culture per 500 mL of LB medium). The cultures were incubated at 37 °C until an optical density of 0.6 AU at 600 nm was reached. Isopropyl-β-D-1-thiogalactopyranoside (IPTG) (60 mg) was added to the LB medium and incubated for the required time and temperature (Table 7.11 and Table 7.12). The induced cells were harvested via centrifugation at 3400 g for 20 minutes.

Table 7.11 Conditions used for expression of NADPH and DHF genes

Gene	*E. coli* expression strain	Induction temperature (°C)	Time (hours)
nadA *	BL21-CodonPlus(DE3)-RP**	25	Overnight (O/N)
nadB	BL21-CodonPlus(DE3)-RP	25	O/N
nadC	BL21-CodonPlus(DE3)-RP	25	O/N
nadD	BL21-Star^{TM}(DE3)	37	4
nadE	BL21-Star^{TM}(DE3)	25	O/N
nadK	JM101	37	4
folE	BL21-Star^{TM}(DE3)	25	O/N
folB	BL21-CodonPlus(DE3)-RP	37	4
folK	BL21-CodonPlus(DE3)-RP	37	4
folP	BL21-Star^{TM}(DE3)	25	O/N
folC	BL21-Star^{TM}(DE3)	25	O/N
prsA ***	BL21-Star^{TM}(DE3)	37	4
gpt	BL21-Star^{TM}(DE3)	25	O/N

* FeCl$_3$ was added to the media at induction to give a final concentration of 50 μM

** Cells contained the pB1282 plasmid, a kind gift from Prof Squire Booker, Pennsylvania State University. Arabinose was added at OD$_{600}$=0.3 to induce the expression of the *isc* operon encoded.

*** The vector containing PRPP synthetase was a kind gift from Prof Jamie Williamson, The Scripps Research Institute

Table 7.12 Conditions used for expression of DHFR genes

DHFR genes	*E.coli* expression cells	Induction temperature (°C)	Time (hours)
M. profunda DHFR	BL21 (DE3)	16	O/N
E. coli DHFR	BL21 (DE3)	37	4
B. stearothermophilus DHFR	BL21-Star^{TM}(DE3)	25	O/N

7.4.2   Isotopically labelled BsDHFR

E. coli BL21(DE3) Star cells harbouring the recombinant BsDHFR construct from an overnight culture in LB medium containing 100 μg/mL ampicillin were washed three times with M9 medium and then grown in 1 L M9 medium until the OD$_{600}$ reached

0.6. IPTG (0.5 mM) was added and the culture grown to an OD_{600} of 2.0. The cells were harvested by centrifugation for 30 min at 6000g, and the pellet was resuspended in 50 mM Tris (pH 9.0) containing 5 mM β-mercaptoethanol and sonicated for 10 min on ice. The cell lysate was heated to 55 °C for 15 min and precipitated proteins were separated by centrifugation for 30 min at 16000 g. The supernatant solution containing BsDHFR was fractionated by HiPrep 16/10 DEAE anion-exchange chromatography (GE Healthcare) in 50 mM Tris pH 9.0 with a gradient of 0-1 M NaCl. The pooled DHFR fractions were concentrated and subjected to size-exclusion chromatography using a Superdex75 column (320 mL column volume (CV), GE Healthcare) in 50 mM potassium phosphate (5 mM β-mercaptoethanol, pH 7). Labelled (heavy) BsDHFR was produced as described for the unlabelled enzyme in M9 medium using [$^2H_2$, 99.9%] H_2O supplemented with 1 g/L [^{15}N, 98%] NH_4Cl and 2 g/L [99%, $^{13}C_6$; 1,2,3,4,5,6,6-$^2H_7$, 97%] glucose (heavy medium). The purity of the enzyme was assessed by SDS-PAGE.

7.5   Enzyme purification

7.5.1   Phosphate stock solutions

7.5.1.1   1 M KH₂PO₄

KH_2PO_4 (68.1 g) was added to 450 mL deionised water. The volume was taken to 500 mL and the solution was stored at 4 °C.

7.5.1.2   1 M K₂HPO₄

K_2HPO_4 (87.1 g) was added to 450 mL deionised water. The volume was taken to 500 mL and the solution was stored at 4 °C.

7.5.2   Purification buffers

7.5.2.1   Purification buffer A

Tris base (6.1 g), NaCl (17.5 g), and imidazole (1.4 g) were added to 950 mL of deionised water. The pH was adjusted to 8.0 using 6 M HCl and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM Tris base, 300 mM NaCl and 20 mM imidazole. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored

at room temperature.

7.5.2.2   *Purification buffer B*

Tris base (6.1 g), NaCl (17.5 g), and imidazole (17.0 g) were added to 950 mL of deionised water. The pH was adjusted to 8.0 using 6 M HCl and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM Tris base, 300 mM NaCl and 250 mM imidazole. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at room temperature.

7.5.2.3   *Purification buffer C*

HEPES (11.9 g), NaCl (17.5 g), imidazole (1.4 g) and βME (2 mL) were added to 950 mL of deionised water. The pH was adjusted to 8.0 using 6 M HCl and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM HEPES, 300 mM NaCl, 20 mM imidazole and 0.2% v/v βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and used immediately.

7.5.2.4   *Purification buffer D*

HEPES (11.9 g), NaCl (17.5 g), imidazole (17.0 g) and βME (2 mL) were added to 950 mL of deionised water. The pH was adjusted to 8.0 using 6 M HCl and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM HEPES, 300 mM NaCl, 250 mM imidazole and 0.2% v/v βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and used immediately.

7.5.2.5   *Purification buffer E*

HEPES (11.9 g), NaCl (1.4 g), imidazole (17.0 g) and βME (700 μL) were added to 950 mL of deionised water. The pH was adjusted to 8.0 using 6 M HCl and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM HEPES, 300 mM NaCl, 20 mM imidazole and 10 mM βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein,

Germany) before bubbling with N$_2$ for 30 minutes and placed in a glove box with the cap removed for 24 hr.

7.5.2.6   *Purification buffer F*

HEPES (11.9 g), NaCl (17.5 g), imidazole (17.0 g) and βME (700 μL) were added to 950 mL of deionised water. The pH was adjusted to 8.0 using 6 M HCl and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM HEPES, 300 mM NaCl, 250 mM imidazole and 10 mM βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) before bubbling with N$_2$ for 30 minutes and placed in a glove box with the cap removed for 24 hr.

7.5.2.7   *Purification buffer G*

Tris base (6.1 g) and βME (350 μL) were added to 950 mL of deionised water. The pH was adjusted to 8.0 using 6 M HCl and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM Tris base and 10 mM βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at room temperature.

7.5.2.8   *Purification buffer H*

Tris base (6.1 g), NaCl (58.4 g) and βME (350 μL) were added to 950 mL of deionised water. The pH was adjusted to 8.0 using 6 M HCl and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM Tris base, 1 M NaCl and 5 mM βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at room temperature.

7.5.2.9   *Purification buffer I*

1 M K$_2$HPO$_4$ (30.5 mL, section 7.5.1.2), 1 M KH$_2$PO$_4$ (19.5 mL, section 7.5.1.1) and βME (350 μL) were added to 900 mL deionised water. The pH was adjusted to 7.0 using 1 M HCl and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM K$_i$PO$_4$ and 5 mM βME. The solution was filtered and

degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at room temperature.

7.5.2.10 Purification buffer J

1 M K_2HPO_4 (30.5 mL, section 7.5.1.2), 1 M KH_2PO_4 (19.5 mL, section 7.5.1.1), NaCl (58.4 g) and βME (350 μL) were added to 900 mL deionised water. The pH was adjusted to 7.0 and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM K_iPO_4, 1 M NaCl and 5 mM βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at room temperature.

7.5.2.11 Purification buffer K

Tris base (6.1 g) and βME (350 μL) were added to 950 mL of deionised water. The pH was adjusted to 9.0 using 6 M HCl and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM Tris base and 5 mM βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at room temperature.

7.5.2.12 Purification buffer L

Tris base (6.1 g), NaCl (58.4 g) and βME (350 μL) were added to 950 mL of deionised water. The pH was adjusted to 9.0 using 6 M HCl and the total volume taken to 1 L with deionised water to give a final concentration of 50 mM Tris base, 1 M NaCl and 5 mM βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at room temperature.

7.5.3   Purification Methods

Each protein was purified using the method shown (Table 7.13)

Table 7.13 Purification methods used for each protein

Purification Methods	Protein
Method 1	L-Aspartate oxidase, quinolinate phosphoribosyltransferase, nicotinic acid mononucleotide adenylyltransferase, NAD synthetase, NAD kinase, xanthine-guanine phosphoribosyltransferase, dihydroneopterin aldolase, hydroxymethylpterin pyrophosphokinase
Method 2	Quinolinate synthetase (aerobic purification)
Method 3	Quinolinate synthetase (anaerobic purification)
Method 4	GTP cyclohydrolase I, dihydropteroate synthetase, dihydrofolate synthase
Method 5	*E. coli* DHFR, *M. profunda* DHFR
Method 6	*G. stearothermophilus* DHFR

7.5.3.1 Method 1

Induced cells harvested after expression (Section 7.4) were resuspended in purification buffer A (Section 7.5.2.1) and lysed by sonication (5 second sonication period followed by 10 second rest period for a total of 15 minutes). The solution was clarified by centrifugation at 4 °C, 24300 g for 30 minutes. The supernatant was then applied to a Ni-NTA column pre-equilibrated with 5 CV binding buffer. The flow through was collected and the column was washed with 10 CV binding buffer before elution with 5 CV purification buffer B (Section 7.5.2.2). All fractions were analysed by SDS-PAGE and fractions containing the target protein were dialysed against 5 mM K_iPO_4, 5 mM βME.

7.5.3.2 Method 2

Induced cells harvested after expression (Section 7.4) were resuspended in purification buffer C (Section 7.5.2.3) and argon bubbled through the solution for 15 minutes before the cells were lysed by sonication (5 second sonication period followed by 5 second rest period for a total of 6 minutes). The solution was bubbled with argon for 3 minutes then clarified by centrifugation at 4 °C, 24300 g for 30 minutes. The supernatant was then applied to a Ni^{2+}-NTA column pre-equilibrated

with 5 CV binding buffer with a nitrogen line in the column. The flow through was collected and the column was washed with 10 CV binding buffer before elution with 5 CV purification buffer D (Section 7.5.2.4) under nitrogen. The elution was collected in 5 mL fractions and the most 3 most concentrated (by colour) were immediately passed through a Sephadex-G25 column, pre-equilibrated in purification buffer C, whilst under nitrogen.

7.5.3.3 Method 3

Induced cells harvested after expression (Section 7.4) were frozen in liquid nitrogen and transferred to a glove box. Cells were then defrosted and resuspended in purification buffer E (Section 7.5.2.5) lysed by sonication (5 second sonication period followed by 10 second rest period for a total of 15 minutes). The solution was sealed in air-tight centrifuge tubes before being transferred out of the glove box and then clarified by centrifugation at 4 °C, 24300 g for 30 minutes. The sealed tubes were transferred back into the glove box and the supernatant was then applied to a Ni^{2+}-NTA column pre-equilibrated with 5 CV binding buffer. The flow through was collected and the column was washed with 10 CV binding buffer before elution with 5 CV purification buffer F (Section 7.5.2.6). The elution was collected in 5 mL fractions and the 3 most concentrated (by colour) were immediately passed through a CentriPure P25 desalting column, pre-equilibrated in purification buffer E.

7.5.3.4 Method 4

Induced cells harvested after expression (Section 7.4) were resuspended in purification buffer G (Section 7.5.2.7) and lysed by sonication (5 second sonication period followed by 10 second rest period for a total of 15 minutes). The solution was clarified by centrifugation at 4 °C, 24300 g for 30 minutes and the supernatant loaded onto a Q-sepharose anion exchange column (~60 mL CV). Flow through was collected and the column washed with purification buffer G for 4 CV until baseline absorbance was reached. An NaCl gradient was applied in 2 stages from 0.0-0.5 M NaCl (50% purification buffer H, Section 7.5.2.8) over 4 CV followed by 0.5-1.0 M

NaCl (100% buffer H) over 0.1 CV. Fractions were collected (10 mL) and the absorbance was monitored at 280 nm. All fractions were analysed by SDS-PAGE and fractions containing the target protein were dialysed against 5 mM K_iPO_4, 5 mM βME.

7.5.3.5 *Method 5*

Protein was purified as in method 4 (section 7.5.3.4) with the exception that purification buffers I and J (section 7.5.2.9 and 7.5.2.10) were used in place of purification buffers G and H respectively. Fractions containing the desired protein from the Q-sepharose column were concentrated to a volume of ~10 mL using an Amicon ultrafiltration device. The concentrated solution was loaded onto a Superdex75 size exclusion column (320 mL CV) equilibrated with purification buffer M. The sample was eluted using buffer I (1.5 CV) and fractions collected (10 mL). The point of elution of the desired protein was determined using a DHFR activity assay and SDS-PAGE was used to determine the purity of the protein.

7.5.3.6 *Method 6*

Induced cells harvested after expression were resuspended in purification buffer A and lysed by sonication (5 second sonication period followed by 10 second rest period for a total of 15 minutes). The solution was heat shocked at 55 °C for 10 minutes before clarification by centrifugation at 4 °C, 24300 g for 30 minutes. The supernatant was loaded onto a DEAE anion exchange column (~60 mL CV). Flow through was collected and the column washed with purification buffer K (section 7.5.2.11) for 2 CV until baseline absorbance was reached. An NaCl gradient was applied in 2 stages from 0.0-0.5 M NaCl (50% purification buffer L, section 7.5.2.12) over 4 CV followed by 0.5-1.0 M NaCl (100% buffer L) over 0.1 CV. Fractions were collected (12 mL) and the absorbance was monitored at 280 nm. The point of elution of the desired protein was determined using a DHFR activity assay and SDS-PAGE was used to determine the purity of the protein.

Fractions containing the desired protein were concentrated to a volume of ~10 mL using an Amicon ultrafiltration device. The concentrated solution was loaded onto a

Superdex75 size exclusion column (320 mL CV) equilibrated with purification buffer M. The sample was eluted using buffer I (1.5 CV) and fractions collected (10 mL). The point of elution of the desired protein was determined using a DHFR activity assay and SDS-PAGE was used to determine the purity of the protein.

7.6 SDS-PAGE

7.6.1 Resolving gel (10%)

30% acrylamide/bis-acrylamide (3.3 mL), resolving buffer (2.5 mL, Section 7.6.4), 10% (w/v) SDS (0.1 mL) and deionised water (4.1 mL) were mixed together. Immediately prior to pouring, freshly made 10% ammonium persulfate (APS, 50 μL) and N,N,N',N'-tetramethylenediamine (TEMED, 20 μL) were added and the solution gently mixed to initiate polymerisation.

7.6.2 Stacking gel (4%)

30% acrylamide/bis-acrylamide (1.3 mL), stacking buffer (2.5 mL, Section 7.6.3), 10% (w/v) SDS (0.1 mL) and deionised water (6.1 mL) were mixed together. Immediately prior to pouring, freshly made 10% ammonium persulfate (APS, 50 μL) and N,N,N',N'-tetramethylenediamine (TEMED, 20 μL) were added and the solution gently mixed to initiate polymerisation.

7.6.3 SDS stacking buffer

Tris base (6 g) was added to 60 mL deionised water and the pH adjusted to 6.8. The final volume was taken up to 100 mL.

7.6.4 SDS resolving buffer

Tris base (27.23 g) was dissolved in 80 mL deionised water and the pH adjusted to 8.8. The final volume was taken to 150 mL.

7.6.5 Running buffer

Tris base (30.3 g), glycine (144.0 g) and SDS (10.0 g) was added to 900 mL deionised water. The final volume was taken up to 1 L with deionised water. Buffer was diluted 1:9 with deionised water immediately prior to use.

7.6.6 SDS-PAGE protocol

Monomer solutions were prepared (Section 7.6.1 and 7.6.2). When the resolving gel had been poured it was covered with a layer of isopropanol and allowed to polymerise. Isopropanol was then removed and the stacking gel poured and a comb inserted to create loading wells. Fraction samples (10 µL) were incubated with protein marker (10 µL) for 10 minutes at 85 °C and loaded into the gel. Running buffer (section 7.6.5) was added and the gel allowed to run at 150 V for 45 minutes.

7.7 Enzyme storage

QAPRT, NAMNT, NAD kinase, GTP cyclohydrolase I, PRPP synthetase and XGPRT were stored by lyophilising the enzymes. After purification and dialysis into 5 mM K_iPO_4, 5 mM βME the enzyme solutions were frozen in liquid nitrogen, lyophilised and stored in aliquots at -20 °C.

7.8 Enzyme activity assays

7.8.1 Assay buffers

7.8.1.1 *Assay buffer 1*

HEPES (2.4 g) and NaCl (0.6 g) were added to 150 mL of deionised water. The pH was adjusted to 8.0 using 6 M HCl and the total volume taken to 200 mL with deionised water to give a final concentration of 50 mM HEPES and 50 mM NaCl. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and used immediately.

7.8.1.2 *Assay buffer 2*

1 M K_2HPO_4 (12.3 mL, section 7.5.1.2), 1 M KH_2PO_4 (7.7 mL, section 7.5.1.1), NaCl (1.17 g) and 1 M $MgCl_2$ (2 mL) were added to 150 mL deionised water. The pH was adjusted to 7.0 and the total volume taken to 200 mL with deionised water to give a

final concentration of 50 mM K_iPO_4, 100 mM NaCl and 10 mM $MgCl_2$. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany).

7.8.2   L-Aspartate oxidase

LAO activity was assayed either via UV-Vis by coupling the reaction to horseradish peroxidase (HRP) and monitoring at 436 nm or by directly through the absorption of FAD and $FADH_2$ at 450 nm. For the coupled assay stock solutions of FAD (4 mg/mL), σ-dianisidine (5 mg/mL it EtOH) 3 mg HRP and aspartate (13 mg/mL) were made up in assay buffer 1 (section 7.8.1.1). 4 μL FAD stock solution, 20 μL σ-dianisidine stock solution and 100 μL aspartate stock solution were added to a cuvette to give final concentrations of 20 μM FAD, 0.01% v/v σ-dianisidine, and 10 mM aspartate. The volume was brought to 850 μL with assay buffer 1, before the addition of 50 μL HRP stock solution and 100 μL LAO (approx. 100 μM stock solution) to initiate the reaction. The absorption change at 436 nm was followed.

For the direct assay 100 μL aspartate stock solution and 4 μL FAD stock solution were added to a cuvette and the volume taken to 900 μL with assay buffer 1 (section 7.8.1.1) to give final concentrations of 100 mM aspartate and 20 μM FAD. 100 μL LAO (approx. 100 μM stock solution) was added to start the reaction and the absorption at 450 nm was measured.

7.8.3   NAD kinase

NAD kinase activity was assayed via UV-Vis by coupling the reaction to TbADH activity. Stock solution of NAD and ATP was made up (5 mg in 1 mL assay buffer 2 (section 7.8.1.2)) and 100 μL of each added to a cuvette. 10 μL TbADH (2 mg/mL) and 100 μL isopropanol was added and the volume taken up to 900 μL with assay buffer 2. 100 μL NAD kinase was added (typically 10 μM final concentration) and the increase in absorption at 340 nm monitored over 15 minutes.

7.8.4   GTP cyclohydrolase I

GTP cyclohydrolase I was assayed via UV-Vis spectroscopy. 0.2 μL GTP (100 mM) was added to 900 μL assay buffer 2 (section 7.8.1.2). 100 μL GTP cyclohydrolase

was added (typically 10 µM final concentration) and the increase in absorption at 330 nm was monitored.

7.8.5 PRPP synthetase

PRPP synthetase was assayed via UV-Vis spectroscopy by coupling the reaction to myokinase (MK), pyruvate kinase (PK) and lactose dehydrogenase activity. Stock solutions of ATP, ribose 5-phosphate, phosphoenolpyruvate (PEP) and NADH were made up (5 mg in 1 mL assay buffer 2 (section 7.8.1.2)) and 100 µL of each added to a cuvette. 5 µL PK (5 mg/mL), 2 µL MK ((NH_4)$_2$$SO_4$ suspension) 5 µL LDH (5 mg/mL) were added and the volume taken up to 900 µL with assay buffer 2. 100 µL of PRPP synthetase (typically 10 µM final concentration in cuvette) was added and the decrease in absorption at 340 nm monitored over 15-20 minutes.

7.8.6 DHFR

Steady state kinetics were performed on a Jasco V-660 spectrometer. The cuvettes used all have a pathlength of 10 mm. Assay buffer 2 (section 7.8.1.2) was incubated with the sample fraction (0.5-10 µL) and NADPH (100 µM). The reaction was started by the addition of DHF (100 µM). The oxidation of NADPH was monitored at 340 nm and a linear decrease in the absorbance was recorded over time. A linear fit was applied to the decrease in absorbance signal using the software provided (Spectra Analysis, Spectra Manager Version 203.02$^{©}$ JASCO corporation). Initial velocities were recorded up to 40 seconds and any fractions with an absorbance change >0.5 AU were pooled and concentrated to ~15 mL.

7.9 Enzyme concentration determination

7.9.1 Bradford Assay

7.9.1.1 Bradford reagent

Brilliant blue G250 (20 mg) was dissolved in 2 mL of ethanol and H_3PO_4 added (80%, 20 mL). The final volume was brought up to 200 mL with dH_2O and stored in the dark at 4 °C.

7.9.1.2 Bradford assay

A 1 mg mL^{-1} sample of bovine serum albumin (BSA) in dH$_2$O was used to prepare a range of concentrations from 0 to 100 µg mL^{-1} in 200 µL deionised water. Bradford reagent (1 mL, section 7.9.1.1)) was added and the absorbance measured from 400 to 600 nm using a Jasco V-660 spectrophotometer. The ratio of absorbances at 590 nm and 450 nm was calculated and used to plot a standard curve. The procedure was repeated for the purified proteins and the ratio of absorbances compared to the BSA standard to calculate the concentration of each sample.

7.9.1.3 UV-Vis assay

Filtered and degassed potassium phosphate (100 mM, pH 7) was used. The Beer-Lambert law was used to calculate the concentration (mg mL^{-1}) of all the protein solutions using a Jasco V-660 spectrophotometer:

$$A = \varepsilon.c.l$$

where A = absorbance at the given wavelength, ε = the extinction coefficient, c = concentration of the sample and l = the pathlength of the cuvette (cm). Average extinction coefficients for proteins at 205, 210, 215 and 220 nm (ε = 31, 20, 15 and 11 mL mg^{-1} cm^{-1} respectively) were used, and the concentration in mg mL^{-1} was divided by the molecular weight of the protein (~18000 g mol^{-1}) to give a molar concentration.

7.10 NADPH biosynthesis

7.10.1 Purification buffers

7.10.1.1 Purification buffer M

1 M K$_2$HPO$_4$ (9.6 mL, section 7.5.1.2) and 1 M KH$_2$PO$_4$ (0.6 mL, section 7.5.1.1) were added to 900 mL deionised water. The pH was adjusted to 7.0 and the total volume taken to 1 L with deionised water to give a final concentration of 10 mM K$_i$PO$_4$. The

solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at 4 °C.

7.10.1.2 Purification buffer N

1 M K_2HPO_4 (9.4 mL, section 7.5.1.2), 1 M KH_2PO_4 (0.6 mL, section 7.5.1.1) and NaCl (58.4 g) were added to 900 mL deionised water. The pH was adjusted to 7.0 and the total volume taken to 1 L with deionised water to give a final concentration of 10 mM K_iPO_4 and 1 M NaCl. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at 4 °C.

7.10.1.3 Purification buffer P

Tris base (2.4 g) was added to 900 mL deionised water. The pH was adjusted to pH 9 with 1 M NaOH and the final volume taken to 1 L to give 20 mM Tris. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at 4 °C.

7.10.1.4 Purification buffer Q

Tris base (2.4 g) and NaCl (58.4 g) were added to 900 mL deionised water. The pH was adjusted to pH 9 with 1 M NaOH and the final volume taken to 1 L to give 20 mM Tris and 1 M NaCl. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at 4 °C.

7.10.2 [4-^{13}C]-NADP$^+$ biosynthesis

7.10.2.1 [4-^{13}C]-Quinolinic acid

[1,6-^{13}C$_2$]-glucose (30 mg), ATP (10 mg) and PEP (100 mg) were dissolved in 2 mL of reaction buffer (50 mM HEPES, 0.2% v/v βME and 50 mM NaCl, pH 8). The pH was adjusted to pH 8 before the addition of hexokinase (3 mg), phosphoglucoisomerase (10 µL (NH$_4$)$_2$SO$_4$ suspension), phosphofructokinase (10 µL (NH$_4$)$_2$SO$_4$ suspension), aldolase (30 µL, (NH$_4$)$_2$SO$_4$ suspension), TIM (5 µL, (NH$_4$)$_2$SO$_4$ suspension) and PK (3 mg). The reaction was left at 37 °C for 6-12 h. Aspartate (30 mg) and FAD (5 mg) were added to 2 mL of reaction buffer and the pH adjusted to 8 before addition to the glucose reaction. The reaction solution was bubbled with argon for 1 h.

Immediately, freshly prepared QS was added (typically 12 mL of approx. 100 µM stock solution) and LAO (3 mL, approx. 100 µM stock solution), the reaction sealed and left at 30 °C for 24-48 h. The reaction was quenched by lowering the pH to 3, before the pH was readjusted to approx. 7. The solution was centrifuged to remove precipitate and passed through a 10000 MW spin column to remove any residual enzymes in solution before loading onto a C18 Gemini Semi-prep column equilibrated with 0.15 v/v trifluoroacetic acid (TFA) and eluted with MeCN, 0.1% v/v TFA. QA was eluted using the following gradient at 5 mL/min: 0-5 min 0% MeCN, 5-20 min 0-20% MeCN, 20-25 min 20-100% MeCN and 25-27 min 100% MeCN. ATP eluted at 7.5 min and QA eluted at 8.8 min. Yield: 1%. δ_H (600 MHz, D_2O): 8.61 (1H, s, broad), 7.98 (1H, dd, J_1 = 9.0, J_{13C-1H} = 165.0), 7.62 (1H, s, broad). δ_C (125 MHz, D_2O): 137.2.

7.10.2.2 [4-^{13}C]-Nicotinic acid adenine dinucleotide

Quinolinic acid (section 7.10.2.1), ribose-5-phosphate (60 mg), ATP (5 mg) and phosphoenolpyruvate (50 mg) were added to 1 mL buffer containing 50 mM K_iPO_4, 50 mM KCl and 10 mM $MgCl_2$. The pH was adjusted to pH 7 before the addition of pyruvate kinase (2 mg), myokinase (1 µL, $(NH_4)_2SO_4$ suspension), PRPP synthetase (1 mg) and QAPRT (5 mg) giving a total volume of 1.5 mL. The reaction was incubated at 37 °C whilst shaking for 3 hours. The reaction was quenched via heat shocking at 120 °C for 90 s and centrifuged at 16100 g for 15 minutes to remove precipitate. The supernatant was loaded onto a SAX10 (9 x 250 mm) anion exchange column equilibrated with purification buffer M (section 7.10.1.1) and eluted with purification buffer N (section 7.10.1.2). The following gradient was used to elute the product at a flow rate of 2 mL/min: 0-11 min 0% buffer N, 11-28 min 0-10% buffer N, 28-38 min 10-30% buffer N and 38-43 min 30-100% buffer N. NAMN eluted at 5 min, QA at 21.5 min and ATP at 37 min.

The eluted NAMN peak was readjusted to pH 7 before the addition of 1 M $MgCl_2$ and KCl to give a final concentration of 10 mM K_iPO_4, 10 mM $MgCl_2$ and 50 mM KCl. ATP (50 mg) was dissolved in 2 mL of 50 mM K_iPO_4, 50 mM KCl and 10 mM $MgCl_2$ and the pH readjusted to 7 before addition of NAMN solution. NAMNT (10 mg) was added and the reaction incubated at 37 °C for 3 h. The reaction was quenched via

heat shocking at 120 °C for 90 s and centrifuged at 16100 g for 15 minutes to remove precipitate. The supernatant was reloaded on the SAX10 column to remove excess ATP and the corresponding NAMN/NAAD peak immediately loaded onto a C18 Gemini Semiprep column equilibrated with 20 mM NH_4HCO_3 and eluted with MeCN. NAAD was eluted using the following gradient at 5 mL/min: 0-5 min 0% MeCN, 5-20 min 0-20% MeCN, 20-25 min 20-100% MeCN and 25-27 min 100% MeCN. NAMN eluted at ~3.5 min and NAAD eluted at 11.5 min. δ_H (500 MHz, D_2O): 9.03 (1H, s, broad), 8.89 (1H, t, J = 5.0), 8.63 (1H, dd, J_1 = 7.5, J_{13C-1H} = 172.5), 8.32 (1H, s), 8.03 (1H, s), 7.94 (1H, t, J_1 = 10.0, J_2 = 5.0), 5.93 (2H, m), 4.41 (2H, s, broad), 4.30-4.13 (8H, m).

7.10.2.3 [4-^{13}C]-NAD$^+$

ATP (50 mg) and NH_4Cl (27 mg) were dissolved in 1 mL buffer containing 50 mM K_iPO_4, 50 mM KCl and 10 mM $MgCl_2$. The pH was adjusted to pH 7 before the addition of NAAD (section 7.10.2.2) and 0.5 mL NAD synthetase (approx. 100 mM stock). The reaction was incubated at 37 °C for 3 hours before being quenched by heat shock at 120 °C for 90 s. The precipitate was removed via centrifugation and the supernatant loaded onto a C18 Gemini Semiprep column equilibrated with 20 mM NH_4HCO_3 and eluted with MeCN. NAD$^+$ was eluted using the following gradient at 5 mL/min: 0-5 min 0% MeCN, 5-20 min 0-20% MeCN, 20-25 min 20-100% MeCN and 25-27 min 100% MeCN. NAD$^+$ eluted at 13.6 min. Yield: 41% from quinolinic acid. δ_H (500 MHz, D_2O): 9.21 (1H, s, broad), 9.02 (1H, t, J = 7.5), 8.71 (1H, dd, J_1 = 10.0, J_{13C-1H} = 175.0), 8.31 (1H, s), 8.07 (1H, t, J_1 = 5.0, J_2 = 10), 8.05 (1H, s), 5.96 (1H, d, J = 5.0), 5.91 (1H, d, J = 5.0), 4.42-4.10 (10 H, m)

7.10.2.4 [4-^{13}C]-NADP$^+$

ATP (25 mg) was dissolved in 1 mL buffer containing 50 mM K_iPO_4, 50 mM KCl and 10 mM $MgCl_2$. The pH was adjusted to pH 7 before the addition of NAD$^+$ (section 7.10.2.3) and NAD kinase (approx. 5 mg). The reaction was incubated at 37 °C for 3 hours before being quenched by heat shock at 120 °C for 90 s. The supernatant was loaded onto a SAX10 (9 x 250 mm) anion exchange column equilibrated with purification buffer M (section 7.10.1.1) and eluted with purification buffer N

(section 7.10.1.2). The following gradient was used to elute the product at a flow rate of 2 mL/min: 0-11 min 0% buffer N, 11-28 min 0-10% buffer N, 28-38 min 10-30% buffer N and 38-43 min 30-100% buffer N. NADP+ eluted at and ATP at . Yield: 40%. δ_H (600 MHz, D_2O): 9.31 (1H, s, broad), 9.08 (1H, t, J = 6.0), 8.87 (1H, dd, J_1 = 6.0, J_{13C-1H} = 174.0), 8.45 (1H, s), 8.21 (1H, t, J = 6.0), 8.19 (1H, s), 6.12 (2H, J = 6.0), 5.01 (1H, m), 4.54-4.24 (9H, m).

7.10.3 (*R*)-[4-²H]-NADPH

Thermoanaerobacter brockii ADH (1 mg) and NADP⁺ (7 mg) were dissolved in purification buffer P (2 mL, Section 7.10.1.3) and d_8-²H-isopropanol (0.2 mL) was added. The reaction mixture was incubated at 40 °C for 10 minutes and loaded onto a SAX10 (4 x 250 mm) anion exchange column equilibrated with purification buffer P (section 7.10.1.3). An NaCl gradient was applied from 0.0-1.0 M (100% purification buffer Q, section 7.10.1.4) over 30 minutes. Absorbance at 260 and 340 nm was monitored and fractions collected when absorbance at 340 nm increased. A control run with NADP⁺ was performed first to enable a comparison showing that reduction of the substrate had occurred and the product had eluted.

7.10.4 (*S*)-[4-²H]-NADPH

(*S*)-[4-²H]-NADPH was produced and purified as (*R*)-[4-²H]-NADPH but glucose-1-D was used as the source of the deuterium label with glucose dehydrogenase rather than alcohol dehydrogenase as the enzyme.

7.10.5 [4-¹³C]-NADPH and (*R*)-[4-²H,¹³C]-NADPH

[4-¹³C]-NADPH and (*R*)-[4-²H,¹³C]-NADPH were produced and purified as (*R*)-[4-²H]-NADPH with the exception that [4-¹³C]-NADP⁺ (section 7.10.2.4) was the starting material and unlabelled isopropanol was used as the hydride source for [4-¹³C]-NADPH.

7.11 DHF synthesis and purification

7.11.1 Dihydrofolate

Folic acid (400 mg) was dissolved in a minimal volume of 5 M NaOH and added to 80 mL 10% aqueous ascorbic acid (pH 6.0). The pH was readjusted to 6 if any change was seen after the folic acid was added. The mixture was stirred under N_2 and the sodium dithionite (4.4 g) added. The reaction was cooled on ice for 5 minutes and the pH slowly lowered to 2.8 by the addition of 1 M HCl. The mixture was centrifuged at 4 °C, 24300 g for 10 minutes and the pellet re-suspended in ~20 mL of acetone, purged with N_2 and re-centrifuged. The washing step and centrifugation was repeated once more with acetone and then with diethyl ether. The product was dried with N_2 and stored at -20 °C in an N_2-purged, foil-wrapped container.

7.11.2 GMP biosynthesis

Ribose-5-phosphate (20 mg), ATP (5 mg) and phosphoenolpyruvate (50 mg) were added to 1 mL buffer containing 50 mM K_iPO_4, 50 mM KCl and 10 mM $MgCl_2$. The pH was adjusted to pH 7 before the addition of pyruvate kinase (2 mg), myokinase (1 µL ,$(NH_4)_2SO_4$ suspension), PRPP synthetase (1 mg) and XGPRT (5 mg) giving a total volume of 1.5 mL. The reaction was incubated at 37 °C whilst shaking for 3 hours. Guanine (15 mg) was dissolved in 1 M NaOH and added in 10 µL aliquots every 5 minutes until saturation occurred. The reaction was quenched via heat shocking at 120 °C for 90 s and centrifuged at 16100 g for 15 minutes to remove precipitate. The supernatant was loaded onto a SAX10 (9 x 250 mm) anion exchange column equilibrated with purification buffer M (Section 7.10.1.1) and eluted with purification buffer N (Section 7.10.1.2). The following gradient was used to elute the product at a flow rate of 2 mL/min: 0-11 min 0% buffer N, 11-28 min 0-10% buffer N, 28-38 min 10-30% buffer N and 38-43 min 30-100% buffer N. GMP eluted at 34 min and ATP at 37 min.

7.11.3 7,8-Dihydroneopterin triphosphate biosynthesis

GTP cyclohydrolase I (10 mg) and GTP (10 mg) were dissolved in 3 mL of 50 mM K_iPO_4, 50 mM KCl and 1 mM $ZnCl_2$ pH 7.5. The reaction was left for 6 hours at 37 °C

before the solution was put through a 10,000 MW filter to remove the enzyme. The supernatant was loaded onto a SAX10 (9 x 250 mm) anion exchange column equilibrated with purification buffer M (Section 7.10.1.1) and eluted with purification buffer N (Section 7.10.1.2). The following gradient was used to elute the product at a flow rate of 2 mL/min: 0-2 min 0% buffer N, 2-11 min 0-10% buffer N, 11-20 min 10-30% buffer N and 20-22 min 30-100% buffer N. 7,8-dihydroneopterin triphosphate eluted at 18 min and the oxidised product at 20 min.

7.11.4 6-Hydroxymethyl-7,8-dihydroneopterin biosynthesis

Dihydroneopterin triphosphate was dissolved in 3 mL of 50 mM K_iPO_4, 50 mM KCl and 1 mM $ZnCl_2$ pH 7.5. Bovine alkaline phosphatase (3 mg) and DHNA (1 mL) were added and the reaction incubated for 3 h at 37 °C. The solution was passed through a 10000 MW spin column to remove any enzyme before loading onto a C18 Gemini Semi-prep column equilibrated with 0.15 v/v trifluoroacetic acid (TFA) and eluted with MeCN, 0.1% v/v TFA. The oxidised products were eluted using the following gradient at 5 mL/min: 0-5 min 0% MeCN, 5-20 min 0-20% MeCN, 20-25 min 20-100% MeCN and 25-27 min 100% MeCN. Neopterin eluted at 12.6 min and 6-hydroxymethyl-7,8-neopterin eluted at 13.2 min.

7.12 Enzyme kinetics

7.12.1 Buffers

7.12.1.1 Phosphate buffer

1 M K_2HPO_4 (12.3 mL, section 7.5.1.2), 1 M KH_2PO_4 (7.7 mL, section 7.5.1.1), NaCl (1.17 g) and βME (140 μL) were added to 150 mL deionised water. The pH was adjusted to 7.0 and the total volume taken to 200 mL with deionised water to give a final concentration of 50 mM K_iPO_4, 100 mM NaCl and 10 mM βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at room temperature.

7.12.1.2 Deuterated phosphate buffer

1 M K_2HPO_4 (12.3 mL, section 7.5.1.2), 1 M KH_2PO_4 (7.7 mL, section 7.5.1.1) and NaCl (1.17 g) were added to 50 mL [$^2H_2$, 99.9%] H_2O. The solution was lyophilised and resuspended in 150 mL [$^2H_2$, 99.9%] H_2O and βME (70 μL) added. The pD was adjusted to 7.0 and the total volume taken to 200 mL with [$^2H_2$, 99.9%] H_2O to give a final concentration of 50 mM KiPO4, 100 mM NaCl and 10 mM βME. The buffer was filtered and degassed immediately prior to use via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at 4 °C.

7.12.1.3 MTEN buffer

Morpholinoethanesulfonic acid (MES, 1.952 g), Tris (0.6057 g), ethanolamine (1.5 mL), NaCl (1.168 g) and 10 mM βME (140 μL) were added to 150 mL deionised water. The pH was adjusted to the required pH (5-9) at the temperature of use and the total volume taken to 200 mL with deionised water to give a final concentration of 50 mM MES, 25 mM Tris, 25 mM ethanolamine, 100 mM NaCl and 10 mM βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at room temperature.

7.12.1.4 Deuterated MTEN buffer

Morpholinoethanesulfonic acid (MES, 1.952 g), Tris (0.6057 g), ethanolamine (1.5 mL), NaCl (1.168 g) and 10 mM βME (140 μL) were added to 50 mL [$^2H_2$, 99.9%] H_2O. The solution was lyophilised and resuspended in 150 mL [$^2H_2$, 99.9%] H_2O. The pD was adjusted to the required pD (5-9) at the temperature of use and the total volume taken to 200 mL with [$^2H_2$, 99.9%] H_2O to give a final concentration of 50 mM MES, 25 mM Tris, 25 mM ethanolamine, 100 mM NaCl and 10 mM βME. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at 4 °C.

7.12.2 Steady state kinetics

Steady state kinetics were performed on a Jasco V-660 spectrophotometer (Essex, UK) and the temperature was monitored in the cuvette using a Traceable® RTD Platinum Thermometer (Control Company, TX, USA) immediately before the

reaction was started. The reaction was monitored by following the oxidation of NADPH at 340 nm ($\varepsilon 340$ (cofactor + substrate) = 11,800 M^{-1} cm^{-1} (84)) over a 40 s time period.

7.12.2.1 K_M measurements

Typically, for $K_{M\ NADPH}$, enzyme (0.01-0.03 μM) was incubated with NADPH (5-300 μM) for 5 minutes at the required temperature before the addition of DHF (100 μM) to start the reaction. For $K_{M\ DHF}$ measurements, enzyme (0.01-0.03 μM) was incubated with NADPH (100 or 250 μM) for 5 minutes at the required temperature before the addition of DHF (0.5-100 μM) to start the reaction. Measurements were performed at 10, 20 and 35 °C using phosphate buffer (section 7.12.1.1).

7.12.3 Pre-steady state enzyme kinetics

Pre-steady state kinetics were performed on an Applied Photophysics stopped-flow spectrophotometer (Surrey, UK). The reaction was monitored via fluorescence resonance energy transfer (FRET) for EcDHFR and MpDHFR measurements. Tryptophan residues were excited at 292 nm, in turn exciting the NADPH at 340 nm and measuring the subsequent emission at 450 nm. Measurements involving BsDHFR were monitored via FRET or by following the oxidation of NADPH by absorbance at 340 nm.

7.12.3.1 Hydride transfer/primary and secondary KIE measurements

Enzyme and cofactors were incubated in separate syringes with enzyme (20 μM) and NADPH (8 μM) in one syringe and only DHF (200 μM) in the other syringe. The syringes were incubated at the required temperature for 10 minutes before starting the reaction by rapid mixing of the solutions in syringes. Measurements were performed between 5 and 40 °C using phosphate buffer (section 7.12.1.1) or MTEN buffer (section 7.12.1.3) for pH rate profiles. Primary KIE measurements were performed by repeating the experiment using (R)-[4-^{2}H]-NADPH instead of NADPH and comparing rates. α-secondary KIEs measurements were performed by repeating the experiment using (S)-[4-^{2}H]-NADPH and comparing rates.

7.12.3.2 Solvent isotope effect measurements

The reaction was set up as for hydride transfer measurements (section 7.12.3.1). Measurements were performed between 5 and 40 °C using phosphate buffer (section 7.12.1.1) and deuterated phosphate buffer (section 7.12.1.2) or MTEN buffer (section 7.12.1.3) and deuterated MTEN buffer (section 7.12.1.4) for pH rate profiles.

7.12.3.3 Heavy atom KIE measurements

Heavy atom KIE measurements were performed as for hydride transfer measurements (section 7.12.3.1) with the exception that $[4-^{13}C]$-labelled cofactors were used. Measurements were performed between 5 and 35 °C using phosphate buffer (section 7.12.1.1).

7.12.4 Heavy enzyme kinetic measurements

Heavy enzyme kinetic experiments were performed in an identical manner to previous steady state (section 7.12.2) and pre-steady state kinetics (section 7.12.3) with the exception that $^{13}C,^{15}N,^{2}H$-isotopically labelled enzyme was used.

7.13 Circular dichroism spectroscopy

7.13.1 CD Buffer

1 M K_2HPO_4 (9.4 mL, section 7.5.1.2), and 1 M KH_2PO_4 (0.6 mL, section 7.5.1.1) were added to 900 mL deionised water. The pH was adjusted to 7.0 and the total volume taken to 1 L with deionised water to give a final concentration of 10 mM K_iPO_4. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) immediately prior to use.

7.13.2 CD measurements

Circular dichroism experiments were performed on an Applied PhotoPhysics Chirascan spectrometer. Spectra of 10-20 µM protein in CD buffer (Section 7.13.1) were measured between 190 nm and 400 nm in 10 mm quartz cuvettes under N_2

with a 50 nm/min scan speed, 0.5 nm data pitch, 1 nm bandwidth and 0.5 s response time.

7.13.3 Mean residue ellipticity (MRE)

The spectra obtained from the CD spectrometer was converted to MRE using the following equation:

$$\Theta_{MRE} = \Theta/10.c.n.l$$

where Θ = CD signal in millidegrees, c = molar concentration of sample, n = number of peptide bonds and l = pathlength of the cuvette in cm.

7.14 Error calculations

7.14.1 Standard error of the mean

All errors are reported as the standard error of the mean (σ_M), which is defined as:

$$\sigma_M = \frac{\sigma}{\sqrt{n}}$$

where n is the number of samples and σ is the standard deviation, defined as:

$$\sigma = \sqrt{\frac{\sum(X - M)^2}{(n - 1)^2}}$$

where X is each value measured in the sample and M is the mean of the sample.

8 REFERENCES

1. Radzicka, A. and Wolfenden, R., (1995), A proficient enzyme. *Science,* **267**, 90-93
2. Fischer, E., (1894), Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.,* **27**, 2985-2993
3. Pauling, L., (1946), Molecular architecture and biological reactions. *Chem. Eng. News Arch.,* **24**, 1375-1377
4. Pauling, L., (1948), Nature of forces between large molecules of biological interest. *Nature,* **161**, 707-709
5. Koshland, D.E., (1958), Application of a theory of enzyme specificity to protein synthesis. *Proc. Nat. Acad. Sci. U.S.A.,* **44**, 98-104
6. Koshland, D.E., (1994), The key-lock theory and the induced fit theory. *Angew. Chem. Int. Ed.,* **33**, 2375-2378
7. Warshel, A., Sharma, P.K., Kato, M., Xiang, Y., Liu, H., and Olsson, M.H.M., (2006), Electrostatic basis for enzyme catalysis. *Chem. Rev.,* **106**, 3210-3235
8. Monod, J., Wyman, J., and Changeux, J.-P., (1965), On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.,* **12**, 88-118
9. Kaptein, R., Zuiderweg, E.R.P., Scheek, R.M., Boelens, R., and Vangunsteren, W.F., (1985), A protein structure from nuclear magnetic resonance data - lac repressor headpiece. *J. Mol. Biol.,* **182**, 179-182
10. Hilvert, D., (2013), Design of protein catalysts. *Annu. Rev. Biochem.,* **82**, 447-470
11. Jäckel, C. and Hilvert, D., (2010), Biocatalysts by evolution. *Curr. Opin. Biotechnol.,* **21**, 753-759
12. Kries, H., Blomberg, R., and Hilvert, D., (2013), *De novo* enzymes by computational design. *Curr. Opin. Chem. Biol.,* **17**, 221-228
13. Raynal, M., Ballester, P., Vidal-Ferran, A., and van Leeuwen, P.W.N.M., (2014), Supramolecular catalysis. Part 2: artificial enzyme mimics. *Chem. Soc. Rev.,* **43**, 1734-1787
14. McCammon, J.A. and Harvey, S.C., (1988), *Dynamics of proteins and nucleic acids*. New York: Cambridge University Press.
15. Cannon, W.R. and Benkovic, S.J., (1998), Solvation, reorganization energy, and biological catalysis. *J. Biol. Chem.,* **273**, 26257-26260
16. Benkovic, S.J. and Hammes-Schiffer, S., (2003), A perspective on enzyme catalysis. *Science,* **301**, 1196-1202
17. Nagel, Z.D. and Klinman, J.P., (2009), A 21(st) century revisionist's view at a turning point in enzymology. *Nat. Chem. Biol.,* **5**, 543-550
18. Schwartz, S.D. and Schramm, V.L., (2009), Enzymatic transition states and dynamic motion in barrier crossing. *Nat. Chem. Biol.,* **5**, 552-559
19. Pudney, C.R., Guerriero, A., Baxter, N.J., Johannissen, L.O., Waltho, J.P., Hay, S., and Scrutton, N.S., (2013), Fast protein motions are coupled to enzyme H-transfer reactions. *J. Am. Chem. Soc.,* **135**, 2512-2517
20. Cha, Y., Murray, C.J., and Klinman, J.P., (1989), Hydrogen tunneling in enzyme reactions. *Science,* **243**, 1325-1330
21. Kohen, A., Cannio, R., Bartolucci, S., and Klinman, J.P., (1999), Enzyme dynamics and hydrogen tunnelling in a thermophilic alcohol dehydrogenase. *Nature,* **399**, 496-499
22. Knapp, M.J., Rickert, K., and Klinman, J.P., (2002), Temperature-dependent isotope effects in soybean lipoxygenase-1: Correlating hydrogen tunneling with protein dynamics. *J. Am. Chem. Soc.,* **124**, 3865-3874
23. Agarwal, P.K., Billeter, S.R., and Hammes-Schiffer, S., (2002), Nuclear quantum effects and enzyme dynamics in dihydrofolate reductase catalysis. *J. Phys. Chem. B,* **106**, 3283-3293

24.   Maglia, G. and Allemann, R.K., (2003), Evidence for environmentally coupled hydrogen tunneling during dihydrofolate reductase catalysis. *J. Am. Chem. Soc.,* **125**, 13372-13373

25.   Basran, J., Sutcliffe, M.J., and Scrutton, N.S., (1999), Enzymatic H-transfer requires vibration-driven extreme tunneling. *Biochemistry,* **38**, 3218-3222

26.   Eyring, H., (1935), The activated complex in chemical reactions. *J. Chem. Phys.,* **3**, 107-115

27.   Garcia-Viloca, M., Gao, J., Karplus, M., and Truhlar, D.G., (2004), How enzymes work: Analysis by modern rate theory and computer simulations. *Science,* **303**, 186-195

28.   Melander, L.C.S. and Saunders, W.H., (1980), *Reaction rates of isotopic molecules*. New York: Wiley.

29.   Westheimer, F.H., (1961), The magnitude of the primary kinetic isotope effect for compounds of hydrogen and deuterium. *Chemical Reviews,* **61**, 265-273

30.   Bell, R.P., (1980), *The tunnel effect in chemistry*. New York: Routledge, Chapman & Hall, Incorporated.

31.   Harris, R.J., Meskys, R., Sutcliffe, M.J., and Scrutton, N.S., (2000), Kinetic studies of the mechanism of carbon-hydrogen bond breakage by the heterotetrameric sarcosine oxidase of *Arthrobacter* sp 1-IN. *Biochemistry,* **39**, 1189-1198

32.   Masgrau, L., Roujeinikova, A., Johannissen, L.O., Hothi, P., Basran, J., Ranaghan, K.E., Mulholland, A.J., Sutcliffe, M.J., Scrutton, N.S., and Leys, D., (2006), Atomic description of an enzyme reaction dominated by proton tunneling. *Science,* **312**, 237-241

33.   Basran, J., Sutcliffe, M.J., and Scrutton, N.S., (2001), Deuterium isotope effects during carbon-hydrogen bond cleavage by trimethylamine dehydrogenase - Implications for mechanism and vibrationally assisted hydrogen tunneling in wild-type and mutant enzymes. *J. Biol. Chem.,* **276**, 24581-24587

34.   Bandarian, V. and Reed, G.H., (2000), Isotope effects in the transient phases of the reaction catalyzed by ethanolamine ammonia-lyase: Determination of the number of exchangeable hydrogens in the enzyme–cofactor complex. *Biochemistry,* **39**, 12069-12075

35.   Agrawal, N., Hong, B., Mihai, C., and Kohen, A., (2004), Vibrationally enhanced hydrogen tunneling in the *Escherichia coli* thymidylate synthase catalyzed reaction. *Biochemistry,* **43**, 1998-2006

36.   Stojkovic, V. and Kohen, A., (2009), Enzymatic H transfers: Quantum tunneling and coupled motion from kinetic isotope effect studies. *Isr. J Chem.,* **49**, 163-173

37.   Anslyn, E.V. and Dougherty, D.A., (2006), *Modern physical organic chemistry*. Sausalito, California: University Science Books.

38.   Caratzoulas, S. and Schwartz, S.D., (2001), A computational method to discover the existence of promoting vibrations for chemical reactions in condensed phases. *J. Chem. Phys.,* **114**, 2910-2918

39.   Antoniou, D., Caratzoulas, S., Kalyanaraman, C., Mincer, J.S., and Schwartz, S.D., (2002), Barrier passage and protein dynamics in enzymatically catalyzed reactions. *Eur. J. Biochem.,* **269**, 3103-3112

40.   Sutcliffe, M.J. and Scrutton, N.S., (2002), A new conceptual framework for enzyme catalysis - Hydrogen tunneling coupled to enzyme dynamics in flavoprotein and quinoprotein enzymes. *Eur. J. Biochem.,* **269**, 3096-3102

41.   Knapp, M.J. and Klinman, J.P., (2002), Environmentally coupled hydrogen tunneling - Linking catalysis to dynamics. *Eur. J. Biochem.,* **269**, 3113-3121

42.   Klinman, J.P., (2010), A new model for the origin of kinetic hydrogen isotope effects. *J. Phys. Org. Chem.,* **23**, 606-612

43. Pu, J., Gao, J., and Truhlar, D.G., (2006), Multidimensional tunneling, recrossing, and the transmission coefficient for enzymatic reactions. *Chem. Rev.,* **106**, 3140-3169

44. Truhlar, D.G., (2006), *Variational transition state theory and multidimensional tunneling for simple and complex reactions in the gas phase, solids, liquids and enzymes*, in *Isotope Effect in Chemistry and Biology*, A. Kohen and H.H. Limbach, Editors. CRC/Taylor and Francis: Boca Raton.

45. Truhlar, D.G., Gao, J.L., Alhambra, C., Garcia-Viloca, M., Corchado, J., Sanchez, M.L., and Villa, J., (2002), The incorporation of quantum effects in enzyme kinetics modeling. *Acc. Chem. Res.,* **35**, 341-349

46. Alhambra, C., Corchado, J., Sanchez, M.L., Garcia-Viloca, M., Gao, J., and Truhlar, D.G., (2001), Canonical variational theory for enzyme kinetics with the protein mean force and multidimensional quantum mechanical tunneling dynamics. Theory and application to liver alcohol dehydrogenase. *J. Phys. Chem. B,* **105**, 11326-11340

47. Truhlar, D.G., (2010), Tunneling in enzymatic and nonenzymatic hydrogen transfer reactions. *J. Phys. Org. Chem.,* **23**, 660-676

48. Garcia-Viloca, M., Alhambra, C., Truhlar, D.G., and Gao, J.L., (2003), Hydride transfer catalyzed by xylose isomerase: Mechanism and quantum effects. *J. Comput. Chem.,* **24**, 177-190

49. Villa, J. and Warshel, A., (2001), Energetics and dynamics of enzymatic reactions. *J. Phys. Chem. B,* **105**, 7887-7907

50. Kamerlin, S.C.L. and Warshel, A., (2010), At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis? *Proteins Struct. Funct. Bioinf.,* **78**, 1339-1375

51. Adamczyk, A.J., Cao, J., Kamerlin, S.C.L., and Warshel, A., (2011), Catalysis by dihydrofolate reductase and other enzymes arises from electrostatic preorganization, not conformational motions. *Proc. Nat. Acad. Sci. U.S.A.,* **108**, 14115-14120

52. Pisliakov, A.V., Cao, J., Kamerlin, S.C.L., and Warshel, A., (2009), Enzyme millisecond conformational dynamics do not catalyze the chemical step. *Proc. Nat. Acad. Sci. U.S.A.,* **106**, 17359-17364

53. McGeagh, J.D., Ranaghan, K.E., and Mulholland, A.J., (2011), Protein dynamics and enzyme catalysis: Insights from simulations. *Biochim. Biophys. Acta, Proteins Proteomics,* **1814**, 1077-1092

54. Glowacki, D.R., Harvey, J.N., and Mulholland, A.J., (2012), Taking Ockham's razor to enzyme dynamics and catalysis. *Nat. Chem.,* **4**, 169-176

55. Olsson, M.H.M., Parson, W.W., and Warshel, A., (2006), Dynamical contributions to enzyme catalysis: Critical tests of a popular hypothesis. *Chem. Rev.,* **106**, 1737-1756

56. Blakley, R.L., Ramasastry, B.V., and McDougall, B.M., (1963), Biosynthesis of thymidylic acid V. Hydrogen isotope studies with dihydrofolic reductase and thymidylate synthetase. *J. Biol. Chem.,* **238**, 3075-&

57. Charlton, P.A., Young, D.W., Birdsall, B., Feeney, J., and Roberts, G.C.K., (1979), Stereochemistry of reduction of folic acid using dihydrofolate reductase. *J. Chem. Soc., Chem. Commun.,* 922-924

58. Charlton, P.A., Young, D.W., Birdsall, B., Feeney, J., and Roberts, G.C.K., (1985), Stereochemistry of reduction of the vitamin folic acid by dihydrofolate reductase. *J. Chem. Soc., Perkin Trans. 1,* 1349-1353

59. Nelson, D.D.L., Lehninger, A.L., and Cox, M.M., (2013), *Lehninger principles of biochemistry*. New York: W.H. Freeman.

60. McMurry, J. and Begley, T.P., (2005), *The organic chemistry of biological pathways*. Englewood: Roberts and Company Publishers.

61. O'Neil, R.H., Lilien, R.H., Donald, B.R., Stroud, R.M., and Anderson, A.C., (2003), Phylogenetic classification of protozoa based on the structure of the linker domain in the bifunctional enzyme, dihydrofolate reductase-thymidylate synthase. *J. Biol. Chem.,* **278**, 52980-52987

62. Myllykallio, H., Lipowski, G., Leduc, D., Filee, J., Forterre, P., and Liebl, U., (2002), An alternative flavin-dependent mechanism for thymidylate synthesis. *Science,* **297**, 105-107

63. Myllykallio, H., Leduc, D., Filee, J., and Liebl, U., (2003), Life without dihydrofolate reductase FolA. *Trends Microbiol.,* **11**, 220-223

64. Giladi, M., Bitan-Banin, G., Mevarech, M., and Ortenberg, R., (2002), Genetic evidence for a novel thymidylate synthase in the halophilic archaeon *Halobacterium salinarum* and in *Campylobacter jejuni*. *FEMS Microbiol. Lett.,* **216**, 105-109

65. Bolin, J.T., Filman, D.J., Matthews, D.A., Hamlin, R.C., and Kraut, J., (1982), Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 A resolution I. General features and binding of methotrexate. *J. Biol. Chem.,* **257**, 13650-13652

66. Matthews, D.A., Bolin, J.T., Burridge, J.M., Filman, D.J., Volz, K.W., Kaufman, B.T., Beddell, C.R., Champness, J.N., Stammers, D.K., and Kraut, J., (1985), Refined crystal structures of *Escherichia coli* and chicken liver dihydrofolate reductase containing bound trimethoprim. *J. Biol. Chem.,* **260**, 381-391

67. Bystroff, C. and Kraut, J., (1991), Crystal structure of unliganded *Escherichia coli* dihydrofolate reductase: ligand-induced conformational changes and cooperativity in binding. *Biochemistry,* **30**, 2227-2239

68. Bystroff, C., Oatley, S.J., and Kraut, J., (1990), Crystal structures of *Escherichia coli* dihydrofolate reductase: the NADP$^+$ holoenzyme and the folate-NADP$^+$ ternary complex. Substrate binding and a model for the transition-state. *Biochemistry,* **29**, 3263-3277

69. Sawaya, M.R. and Kraut, J., (1997), Loop and subdomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: Crystallographic evidence. *Biochemistry,* **36**, 586-603

70. Osborne, M.J., Venkitakrishnan, R.P., Dyson, H.J., and Wright, P.E., (2003), Diagnostic chemical shift markers for loop conformation and substrate and cofactor binding in dihydrofolate reductase complexes. *Protein Sci.,* **12**, 2230-2238

71. Boehr, D.D., McElheny, D., Dyson, H.J., and Wright, P.E., (2006), The dynamic energy landscape of dihydrofolate reductase catalysis. *Science,* **313**, 1638-1642

72. Rod, T.H. and Brooks, C.L., (2003), How dihydrofolate reductase facilitates protonation of dihydrofolate. *J. Am. Chem. Soc.,* **125**, 8718-8719

73. Fierke, C.A. and Benkovic, S.J., (1989), Probing the functional role of threonine-113 of *Escherichia coli* dihydrofolate reductase for its effect on turnover efficiency, catalysis, and binding. *Biochemistry,* **28**, 478-486

74. Warren, M.S., Brown, K.A., Farnum, M.F., Howell, E.E., and Kraut, J., (1991), Investigation of the functional role of tryptophan-22 in *Escherichia coli* dihydrofolate reductase by site-directed mutagenesis. *Biochemistry,* **30**, 11092-11103

75. Howell, E.E., Villafranca, J.E., Warren, M.S., Oatley, S.J., and Kraut, J., (1986), Functional role of aspartic acid-27 in dihydrofolate reductase revealed by mutagenesis. *Science,* **231**, 1123-1128

76. Adams, J.A., Fierke, C.A., and Benkovic, S.J., (1991), The function of amino acid residues contacting the nicotinamide ring of NADPH in dihydrofolate reductase from *Escherichia coli*. *Biochemistry,* **30**, 11046-11054

77.   Miller, G.P. and Benkovic, S.J., (1998), Stretching exercises - flexibility in dihydrofolate reductase catalysis. *Chem. Biol.,* **5**, R105-R113

78.   Hecht, D., Tran, J., and Fogel, G.B., (2011), Structural-based analysis of dihydrofolate reductase evolution. *Mol. Phylogenet. Evol.,* **61**, 212-230

79.   McTigue, M.A., Davies, J.F., Kaufman, B.T., and Kraut, J., (1992), Crystal structure of chicken liver dihydrofolate reductase complexed with NADP$^+$ and biopterin. *Biochemistry,* **31**, 7264-7273

80.   Shrimpton, P. and Allemann, R.K., (2002), Role of water in the catalytic cycle of *E. coli* dihydrofolate reductase. *Protein Sci.,* **11**, 1442-1451

81.   Chen, Y.Q., Kraut, J., Blakley, R.L., and Callender, R., (1994), Determination by Raman spectroscopy of the pK_a of N5 of dihydrofolate bound to dihydrofolate reductase - mechanistic implications. *Biochemistry,* **33**, 7021-7026

82.   Meiering, E.M. and Wagner, G., (1995), Detection of long-lived bound water molecules in complexes of human dihydrofolate reductase with methotrexate and NADPH. *J. Mol. Biol.,* **247**, 294-308

83.   Cummins, P.L. and Gready, J.E., (2001), Energetically most likely substrate and active-site protonation sites and pathways in the catalytic mechanism of dihydrofolate reductase. *J. Am. Chem. Soc.,* **123**, 3418-3428

84.   Stone, S.R. and Morrison, J.F., (1982), Kinetic mechanism of the reaction catalyzed by dihydrofolate reductase from *Escherichia coli*. *Biochemistry,* **21**, 3757-3765

85.   Stone, S.R. and Morrison, J.F., (1984), Catalytic mechanism of the dihydrofolate reductase reaction as determined by pH studies. *Biochemistry,* **23**, 2753-2758

86.   Fierke, C.A., Johnson, K.A., and Benkovic, S.J., (1987), Construction and evaluation of the kinetic scheme associated with dihydrofolate reductase from *Escherichia coli*. *Biochemistry,* **26**, 4085-4092

87.   Swanwick, R.S., Maglia, G., Tey, L., and Allemann, R.K., (2006), Coupling of protein motions and hydrogen transfer during catalysis by *Escherichia coli* dihydrofolate reductase. *Biochem. J.,* **394**, 259-265

88.   Sikorski, R.S., Wang, L., Markham, K.A., Rajagopalan, P.T.R., Benkovic, S.J., and Kohen, A., (2004), Tunneling and coupled motion in the *Escherichia coli* dihydrofolate reductase catalysis. *J. Am. Chem. Soc.,* **126**, 4778-4779

89.   Francis, K., Stojkovic, V., and Kohen, A., (2013), Preservation of protein dynamics in dihydrofolate reductase evolution. *J. Biol. Chem.,* **288**, 35961-35968

90.   Osborne, M.J., Schnell, J., Benkovic, S.J., Dyson, H.J., and Wright, P.E., (2001), Backbone dynamics in dihydrofolate reductase complexes: Role of loop flexibility in the catalytic mechanism. *Biochemistry,* **40**, 9846-9859

91.   Schnell, J.R., Dyson, H.J., and Wright, P.E., (2004), Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu. Rev. Biophys. Biomol. Struct.,* **33**, 119-140

92.   McElheny, D., Schnell, J.R., Lansing, J.C., Dyson, H.J., and Wright, P.E., (2005), Defining the role of active-site loop fluctuations in dihydrofolate reductase catalysis. *Proc. Nat. Acad. Sci. U.S.A.,* **102**, 5032-5037

93.   Epstein, D.M., Benkovic, S.J., and Wright, P.E., (1995), Dynamics of the dihydrofolate reductase folate complex: Catalytic sites and regions known to undergo conformational change exhibit diverse dynamical features. *Biochemistry,* **34**, 11037-11048

94.   Cameron, C.E. and Benkovic, S.J., (1997), Evidence for a functional role of the dynamics of glycine-121 of *Escherichia coli* dihydrofolate reductase obtained from kinetic analysis of a site-directed mutant. *Biochemistry,* **36**, 15792-15800

95.   Wang, L., Goodey, N.M., Benkovic, S.J., and Kohen, A., (2006), Coordinated effects of distal mutations on environmentally coupled tunneling in dihydrofolate reductase. *Proc. Nat. Acad. Sci. U.S.A.,* **103**, 15753-15758

96.   Singh, P., Sen, A., Francis, K., and Kohen, A., (2014), Extension and limits of the network of coupled motions correlated to hydride transfer in dihydrofolate reductase. *J. Am. Chem. Soc.,* **136**, 2575-2582

97.   Adams, J., Johnson, K., Matthews, R., and Benkovic, S.J., (1989), Effects of distal point-site mutations on the binding and catalysis of dihydrofolate reductase from *Escherichia coli*. *Biochemistry,* **28**, 6611-6618

98.   Farnum, M.F., Magde, D., Howell, E.E., Hirai, J.T., Warren, M.S., Grimsley, J.K., and Kraut, J., (1991), Analysis of hydride transfer and cofactor fluorescence decay in mutants of dihydrofolate reductase - possible evidence for participation of enzyme molecular motions in catalysis. *Biochemistry,* **30**, 11567-11579

99.   Murphy, D.J. and Benkovic, S.J., (1989), Hydrophobic interactions via mutants of *Escherichia coli* dihydrofolate reductase - separation of binding and catalysis. *Biochemistry,* **28**, 3025-3031

100.   Miller, G.P. and Benkovic, S.J., (1998), Strength of an interloop hydrogen bond determines the kinetic pathway in catalysis by *Escherichia coli* dihydrofolate reductase. *Biochemistry,* **37**, 6336-6342

101.   Wang, L., Goodey, N.M., Benkovic, S.J., and Kohen, A., (2006), The role of enzyme dynamics and tunnelling in catalysing hydride transfer: studies of distal mutants of dihydrofolate reductase. *Philos. Trans. R. Soc. London, Ser. B,* **361**, 1307-1315

102.   Radkiewicz, J.L. and Brooks, C.L., (2000), Protein dynamics in enzymatic catalysis: Exploration of dihydrofolate reductase. *J. Am. Chem. Soc.,* **122**, 225-231

103.   Agarwal, P.K., Billeter, S.R., Rajagopalan, P.T.R., Benkovic, S.J., and Hammes-Schiffer, S., (2002), Network of coupled promoting motions in enzyme catalysis. *Proc. Nat. Acad. Sci. U.S.A.,* **99**, 2794-2799

104.   Loveridge, E.J. and Allemann, R.K., (2011), Effect of pH on hydride transfer by Escherichia coli dihydrofolate reductase. *Chembiochem,* **12**, 1258-1262

105.   Loveridge, E.J., Tey, L.-H., and Allemann, R.K., (2010), Solvent effects on catalysis by Escherichia coli dihydrofolate reductase. *J. Am. Chem. Soc.,* **132**, 1137-1143

106.   Bhabha, G., Lee, J., Ekiert, D.C., Gam, J., Wilson, I.A., Dyson, H.J., Benkovic, S.J., and Wright, P.E., (2011), A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis. *Science,* **332**, 234-238

107.   Swanwick, R.S., Shrimpton, P.J., and Allemann, R.K., (2004), Pivotal role of Gly 121 in dihydrofolate reductase from *Escherichia coli*: The altered structure of a mutant enzyme may form the basis of its diminished catalytic performance. *Biochemistry,* **43**, 4119-4127

108.   Kamerlin, S.C.L. and Warshel, A., (2010), An analysis of all the relevant facts and arguments indicates that enzyme catalysis does not involve large contributions from nuclear tunneling. *J. Phys. Org. Chem.,* **23**, 677-684

109.   Weikl, T.R. and Boehr, D.D., (2012), Conformational selection and induced changes along the catalytic cycle of Escherichia coli dihydrofolate reductase. *Proteins Struct. Funct. Bioinf.,* **80**, 2369-2383

110.   Loveridge, E.J., Behiry, E.M., Guo, J., and Allemann, R.K., (2012), Evidence that a 'dynamic knockout' in *Escherichia coli* dihydrofolate reductase does not affect the chemical step of catalysis. *Nat. Chem.,* **4**, 292-297

111.   Liu, C.T., Wang, L., Goodey, N.M., Hanoian, P., and Benkovic, S.J., (2013), Temporally overlapped but uncoupled motions in dihydrofolate reductase catalysis. *Biochemistry,* **52**, 5332-5334

112. van den Bedem, H., Bhabha, G., Yang, K., Wright, P.E., and Fraser, J.S., (2013), Automated identification of functional dynamic contact networks from X-ray crystallography. *Nat. Methods,* **10**, 896-902

113. Loveridge, E.J., Tey, L.-H., Behiry, E.M., Dawson, W.M., Evans, R.M., Whittaker, S.B.M., Guenther, U.L., Williams, C., Crump, M.P., and Allemann, R.K., (2011), The role of large-scale motions in catalysis by dihydrofolate reductase. *J. Am. Chem. Soc.,* **133**, 20561-20570

114. Kipp, D.R., Silva, R.G., and Schramm, V.L., (2011), Mass-dependent bond vibrational dynamics influence catalysis by HIV-1 protease. *J. Am. Chem. Soc.,* **133**, 19358-19361

115. Silva, R.G., Murkin, A.S., and Schramm, V.L., (2011), Femtosecond dynamics coupled to chemical barrier crossing in a Born-Oppenheimer enzyme. *Proc. Nat. Acad. Sci. U.S.A.,* **108**, 18661-18665

116. Toney, M.D., Castro, J.N., and Addington, T.A., (2013), Heavy-enzyme kinetic isotope effects on proton transfer in alanine racemase. *J. Am. Chem. Soc.,* **135**, 2509-2511

117. Luk, L.Y.P., Javier Ruiz-Pernia, J., Dawson, W.M., Roca, M., Joel Loveridge, E., Glowacki, D.R., Harvey, J.N., Mulholland, A.J., Tunon, I., Moliner, V., and Allemann, R.K., (2013), Unraveling the role of protein dynamics in dihydrofolate reductase catalysis. *Proc. Nat. Acad. Sci. U.S.A.,* **110**, 16344-16349

118. Javier Ruiz-Pernia, J., Luk, L.Y.P., Garcia-Meseguer, R., Marti, S., Loveridge, E.J., Tunon, I., Moliner, V., and Allemann, R.K., (2013), Increased dynamic effects in a catalytically compromised variant of Escherichia coli dihydrofolate reductase. *J. Am. Chem. Soc.,* **135**, 18689-18696

119. Huber, R., Langworthy, T.A., Konig, H., Thomm, M., Woese, C.R., Sleytr, U.B., and Stetter, K.O., (1986), *Thermotoga maritima* sp-nov represents a new genus of unique extremely thermophilic eubacteria growing up to 90-degrees-C. *Arch. Microbiol.,* **144**, 324-333

120. Maglia, G., Javed, M.H., and Allemann, R.K., (2003), Hydride transfer during catalysis by dihydrofolate reductase from *Thermotoga maritima*. *Biochem. J.,* **374**, 529-535

121. Dams, T., Auerbach, G., Bader, G., Jacob, U., Ploom, T., Huber, R., and Jaenicke, R., (2000), The crystal structure of dihydrofolate reductase from *Thermotoga maritima*: Molecular features of thermostability. *J. Mol. Biol.,* **297**, 659-672

122. Dams, T., Bohm, G., Auerbach, G., Bader, G., Schurig, H., and Jaenicke, R., (1998), Homo-dimeric recombinant dihydrofolate reductase from *Thermotoga maritima* shows extreme intrinsic stability. *Biol. Chem.,* **379**, 367-371

123. Dams, T. and Jaenicke, R., (1999), Stability and folding of dihydrofolate reductase from the hyperthermophilic bacterium *Thermotoga maritima*. *Biochemistry,* **38**, 9169-9178

124. Loveridge, E.J., Behiry, E.M., Swanwick, R.S., and Allemann, R.K., (2009), Different reaction mechanisms for mesophilic and thermophilic dihydrofolate reductases. *J. Am. Chem. Soc.,* **131**, 6926-6927

125. Pang, J.Y., Pu, J.Z., Gao, J.L., Truhlar, D.G., and Allemann, R.K., (2006), Hydride transfer reaction catalyzed by hyperthermophilic dihydrofolate reductase is dominated by quantum mechanical tunneling and is promoted by both inter- and intramonomeric correlated motions. *J. Am. Chem. Soc.,* **128**, 8015-8023

126. Loveridge, E.J., Evans, R.M., and Allemann, R.K., (2008), Solvent effects on environmentally coupled hydrogen tunnelling during catalysis by dihydrofolate reductase from *Thermotoga maritima*. *Chem. Eur. J.,* **14**, 10782-10788

127. Loveridge, E.J. and Allemann, R.K., (2010), The temperature dependence of the kinetic isotope effects of dihydrofolate reductase from *Thermotoga maritima* is influenced by intersubunit interactions. *Biochemistry,* **49**, 5390-5396

128. Guo, J., Loveridge, E.J., Luk, L.Y.P., and Allemann, R.K., (2013), Effect of dimerization on dihydrofolate reductase catalysis. *Biochemistry,* **52**, 3881-3887

129. Donk, P.J., (1920), A highly resistant thermophilic organism. *J. Bacteriol.,* **5**, 373-4

130. Nazina, T.N., Tourova, T.P., Poltaraus, A.B., Novikova, E.V., Grigoryan, A.A., Ivanova, A.E., Lysenko, A.M., Petrunyaka, V.V., Osipov, G.A., Belyaev, S.S., and Ivanov, M.V., (2001), Taxonomic study of aerobic thermophilic bacilli: descriptions of *Geobacillus* subterraneus gen. nov., sp nov and *Geobacillus uzenensis* sp nov from petroleum reservoirs and transfer of *Bacillus stearothermophilus*, *Bacillus thermocatenulatus*, *Bacillus thermoleovorans*, *Bacillus kaustophilus*, *Bacillus thermoglucosidasius* and *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus*, *G. thermocatenulatus*, *G. thermoleovorans*, *G. kaustophilus*, *G. thermoglucosidasius* and *G. thermodenitrificans*. *Int. J. Syst. Evol. Microbiol.,* **51**, 433-446

131. Oyeyemi, O.A., Sours, K.M., Lee, T., Resing, K.A., Ahn, N.G., and Klinman, J.P., (2010), Temperature dependence of protein motions in a thermophilic dihydrofolate reductase and its relationship to catalytic efficiency. *Proc. Nat. Acad. Sci. U.S.A.,* **107**, 10074-10079

132. Kim, H.S., Damo, S.M., Lee, S.Y., Wemmer, D., and Klinman, J.P., (2005), Structure and hydride transfer mechanism of a moderate thermophilic dihydrofolate reductase from *Bacillus stearothermophilus* and comparison to its mesophilic and hyperthermophilic homologues. *Biochemistry,* **44**, 11428-11439

133. Meinhold, L., Clement, D., Tehei, M., Daniel, R., Finney, J.L., and Smith, J.C., (2008), Protein dynamics and stability: The distribution of atomic fluctuations in thermophilic and mesophilic dihydrofolate reductase derived using elastic incoherent neutron scattering. *Biophys. J.,* **94**, 4812-4818

134. Oyeyemi, O.A., Sours, K.M., Lee, T., Kohen, A., Resing, K.A., Ahn, N.G., and Klinman, J.P., (2011), Comparative hydrogen-deuterium exchange for a mesophilic vs thermophilic dihydrofolate reductase at 25 degrees c: Identification of a single active site region with enhanced flexibility in the mesophilic protein. *Biochemistry,* **50**, 8251-8260

135. Evans, R.M., Behiry, E.M., Tey, L.-H., Guo, J., Loveridge, E.J., and Allemann, R.K., (2010), Catalysis by dihydrofolate reductase from the psychropiezophile Moritella profunda. *Chembiochem,* **11**, 2010-2017

136. Guo, J., (2014) Thermophilicity and catalytic efficiency in dihydrofolate reductase, *Ph.D Thesis*, Cardiff University.

137. Allemann, R.K., Evans, R.M., and Loveridge, E.J., (2009), Probing coupled motions in enzymatic hydrogen tunnelling reactions. *Biochem. Soc. Trans.,* **37**, 349-353

138. Guo, J., Luk, L.Y.P., Loveridge, E.J., and Allemann, R.K., (2014), Thermal adaptation of dihydrofolate reductase from the moderate thermophile *Geobacillus stearothermophilus*. *Biochemistry,* **53**, 2855-2863

139. Xu, Y., Nogi, Y., Kato, C., Liang, Z.Y., Ruger, H.J., De Kegel, D., and Glansdorff, N., (2003), *Moritella profunda* sp nov and *Moritella abyssi* sp nov., two psychropiezophilic organisms isolated from deep Atlantic sediments. *Int. J. Syst. Evol. Microbiol.,* **53**, 533-538

140. Hay, S., Evans, R.M., Levy, C., Loveridge, E.J., Wang, X., Leys, D., Allemann, R.K., and Scrutton, N.S., (2009), Are the catalytic properties of enzymes from piezophilic organisms pressure adapted? *Chembiochem,* **10**, 2348-2353

141. Coque, T.M., Singh, K.V., Weinstock, G.M., and Murray, B.E., (1999), Characterization of dihydrofolate reductase genes from trimethoprim-susceptible and trimethoprim-resistant strains of *Enterococcus faecalis*. *Antimicrob. Agents Chemother.,* **43**, 141-147

142. Huovinen, P., Sundstrom, L., Swedberg, G., and Skold, O., (1995), Trimethoprim and sulfonamide resistance. *Antimicrob. Agents Chemother.,* **39**, 279-289

143. Xu, Y., Feller, G., Gerday, C., and Glansdorff, N., (2003), *Moritella* cold-active dihydrofolate reductase: Are there natural limits to optimization of catalytic efficiency at low temperature? *J. Bacteriol.,* **185**, 5519-5526

144. Dawson, W.M., (2010) Catalysis at the extreme, *MSc Dissertation*, Cardiff University.

145. Loveridge, E.J., Dawson, W.M., Evans, R.M., Sobolewska, A., and Allemann, R.K., (2011), Reduced susceptibility of Moritella profunda dihydrofolate reductase to trimethoprim is not due to glutamate 28. *Protein J.,* **30**, 546-548

146. Behiry, E.M., (2013) Dihydrofolate reductase and the physical basis of enzyme catalysis, *Ph.D Thesis*, Cardiff University.

147. Cook, P.F., (1991), *Enzyme mechanism from isotope effects*. Boca Raton: CRC Press.

148. Sen, A. and Kohen, A., (2009), *Quantum effects in enzyme kinetics*, in *Quantum Tunnelling in Enzyme-Catalysed Reactions*, R.K. Allemann and N.S. Scrutton, Editors. p. 161-178.

149. Pu, J.Z., Ma, S.H., Garcia-Viloca, M., Gao, J.L., Truhlar, D.G., and Kohen, A., (2005), Nonperfect synchronization of reaction center rehybridization in the transition state of the hydride transfer catalyzed by dihydrofolate reductase. *J. Am. Chem. Soc.,* **127**, 14879-14886

150. Cook, P.F., Oppenheimer, N.J., and Cleland, W.W., (1981), Secondary deuterium and ^{15}N isotope effects in enzyme catalyzed reactions - chemical mechanism of liver alcohol dehydrogenase. *Biochemistry,* **20**, 1817-1825

151. Cook, P.F., Blanchard, J.S., and Cleland, W.W., (1980), Primary and secondary deuterium isotope effects on equilibrium constants for enzyme catalyzed reactions. *Biochemistry,* **19**, 4853-4858

152. Huskey, W.P. and Schowen, R.L., (1983), Reaction coordinate tunneling in hydride transfer reactions. *J. Am. Chem. Soc.,* **105**, 5704-5706

153. Pudney, C.R., Hay, S., Sutcliffe, M.J., and Scrutton, N.S., (2006), alpha-Secondary isotope effects as probes of "tunneling-ready" configurations in enzymatic H-tunneling: Insight from environmentally coupled tunneling models. *J. Am. Chem. Soc.,* **128**, 14053-14058

154. Basran, J., Harris, R.J., Sutcliffe, M.J., and Scrutton, N.S., (2003), H-tunneling in the multiple H-transfers of the catalytic cycle of morphinone reductase and in the reductive half-reaction of the homologous pentaerythritol tetranitrate reductase. *J. Biol. Chem.,* **278**, 43973-43982

155. Hay, S., Pang, J.Y., Monaghan, P.J., Wang, X., Evans, R.M., Sutcliffe, M.J., Allemann, R.K., and Scrutton, N.S., (2008), Secondary kinetic isotope effects as probes of environmentally-coupled enzymatic hydrogen tunneling reactions. *Chemphyschem,* **9**, 1536-1539

156. Wolfenden, R., (2003), Thermodynamic and extrathermodynamic requirements of enzyme catalysis. *Biophys. Chem.,* **105**, 559-572

157. Bigeleisen, J. and Wolfsberg, M., (1958), Theoretical and experimental aspects of isotope effects in chemical kinetics. *Adv. Chem. Phys.,* **1**, 15-76

158. Schramm, V.L., (1999), Enzymatic transition-state analysis and transition-state analogs. *Methods Enzymol.,* **308**, 301-355

159. Berti, P.J., (1999), Determining transition states from kinetic isotope effects. *Methods Enzymol.,* **308**, 355-397

160. Berti, P.J. and Schramm, V.L., (1997), Transition state structure of the solvolytic hydrolysis of NAD(+). *J. Am. Chem. Soc.,* **119**, 12069-12078

161. Kipp, D.R., Hirschi, J.S., Wakata, A., Goldstein, H., and Schramm, V.L., (2012), Transition states of native and drug-resistant HIV-1 protease are the same. *Proc. Nat. Acad. Sci. U.S.A.,* **109**, 6543-6548

162. Evans, G.B., Furneaux, R.H., Lenz, D.H., Painter, G.F., Schramm, V.L., Singh, V., and Tyler, P.C., (2005), Second generation transition state analogue inhibitors of human 5 '-methylthioadenosine phosphorylase. *J. Med. Chem.,* **48**, 4679-4689

163. Gutierrez, J.A., Luo, M., Singh, V., Li, L., Brown, R.L., Norris, G.E., Evans, G.B., Furneaux, R.H., Tyler, P.C., Painter, G.F., Lenz, D.H., and Schramm, V.L., (2007), Picomolar inhibitors as transition-state probes of 5 '-methylthioadenosine nucleosidases. *ACS Chem. Biol.,* **2**, 725-734

164. Singh, V. and Schramm, V.L., (2006), Transition-state structure of human 5 '-methylthioadenosine phosphorylase. *J. Am. Chem. Soc.,* **128**, 14691-14696

165. Burgos, E.S., Vetticatt, M.J., and Schramm, V.L., (2013), Recycling nicotinamide. The transition-state structure of human nicotinamide phosphoribosyltransferase. *J. Am. Chem. Soc.,* **135**, 3485-3493

166. Parkin, D.W., (1991), *Methods for the determination of competitive and noncompetitive kinetic isotope effects.* Enzyme Mechanism from Isotope Effects, ed. P.F. Cook. Boca Raton: CRC Press.

167. O'Leary, M.H., (1980), Determination of heavy-atom isotope effects on enzyme catalyzed reactions. *Methods Enzymol.,* **64**, 83-104

168. Bahnson, B.J. and Anderson, V.E., (1991), Crotonase catalyzed beta-elimination is concerted: A double isotope effect study. *Biochemistry,* **30**, 5894-5906

169. Singleton, D.A. and Thomas, A.A., (1995), High-precision simultaneous determination of multiple small kinetic isotope effects at natural abundance. *J. Am. Chem. Soc.,* **117**, 9357-9358

170. Parkin, D.W., Leung, H.B., and Schramm, V.L., (1984), Synthesis of nucleotides with specific radiolabels in ribose - primary C-14 and secondary H-3 kinetic isotope effects on acid-catalyzed glycosidic bond hydrolysis of AMP, dAMP, and inosine. *J. Biol. Chem.,* **259**, 9411-9417

171. Parkin, D.W. and Schramm, V.L., (1987), Catalytic and allosteric mechanism of AMP nucleosidase from primary, beta-secondary, and multiple heavy-atom kinetic isotope effects. *Biochemistry,* **26**, 913-920

172. Kline, P.C. and Schramm, V.L., (1993), Purine nucleoside phosphorylase - catalytic mechanism and transition-state analysis of the arsenolysis reaction. *Biochemistry,* **32**, 13212-13219

173. Miles, R.W., Tyler, P.C., Furneaux, R.H., Bagdassarian, C.K., and Schramm, V.L., (1998), One-third-the-sites transition-state inhibitors for purine nucleoside phosphorylase. *Biochemistry,* **37**, 8615-8621

174. Lewandowicz, A. and Schramm, V.L., (2004), Transition state analysis for human and *Plasmodium falciparum* purine nucleoside phosphorylases. *Biochemistry,* **43**, 1458-1468

175. Lewandowicz, A., Tyler, P.C., Evans, G.B., Furneaux, R.H., and Schramm, V.L., (2003), Achieving the ultimate physiological goal in transition state analogue inhibitors for purine nucleoside phosphorylase. *J. Biol. Chem.,* **278**, 31465-31468

176. Evans, G.B., Furneaux, R.H., Lewandowicz, A., Schramm, V.L., and Tyler, P.C., (2003), Synthesis of second-generation transition state analogues of human purine nucleoside phosphorylase. *J. Med. Chem.,* **46**, 5271-5276

177. Roth, B. and Stammers, D.K., (1992), *Drug interactions with target enzymes of known structure*, in *The design of drugs to macromolecular targets*, C.R. Beddell, Editor. Wiley & Sons . Ltd: New York.

178. Huennekens, F.M., (1994), The methotrexate story: A paradigm for development of cancer chemotherapeutic agents. *Adv. Enzyme Reg.,* **34**, 397-419

179. Farber, S., Diamond, L.K., Mercer, R.D., Sylvester, R.F., and Wolff, J.A., (1948), Temporary remissions in acute leukemia in children produced by folic acid antagonist, 4-aminopteroyl-glutamic acid (aminopterin). *New Engl. J. Med.,* **238**, 787-793

180. Klareskog, L., van der Heijde, D., de Jager, J.P., Gough, A., Kalden, J., Malaise, M., Mola, E.M., Pavelka, K., Sany, J., Settas, L., Wajdula, J., Pedersen, R., Fatenejad, S., Sanda, M., and Investigators, T.S., (2004), Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. *Lancet,* **363**, 675-681

181. Rajagopalan, P.T.R., Zhang, Z.Q., McCourt, L., Dwyer, M., Benkovic, S.J., and Hammes, G.G., (2002), Interaction of dihydrofolate reductase with methotrexate: Ensemble and single-molecule kinetics. *Proc. Nat. Acad. Sci. U.S.A.,* **99**, 13481-13486

182. Hazarika, M., White, R.M., Johnson, J.R., and Pazdur, R., (2004), FDA drug approval summaries: Pemetrexed (Alimta (R)). *Oncologist,* **9**, 482-488

183. Hitchings, G.H., (1989), Selective inhibitors of dihydrofolate reductase (Nobel lecture). *Angew. Chem. Int. Ed.,* **28**, 879-885

184. Kovalevskaya, N.V., Smurnyi, E.D., Birdsall, B., Feeney, J., and Pol'shakov, V.I., (2007), Structural factors determining the binding selectivity of antibacterial drug trimethoprim to dihydrofolate reductase. *Pharm. Chem. J.,* **41**, 350-353

185. Falco, E.A., Goodwin, L.G., Hitchings, G.H., Rollo, I.M., and Russell, P.B., (1951), 2-4-diaminopyrimidines - A new series of antimalarials. *Brit. J. Pharmacol. Chemother.,* **6**, 185-200

186. Singh, J., Ray, A.P., Basu, P.C., and Misra, B.G., (1953), Pyrimethamine in treatment of *Vivax malaria* - preliminary report. *Brit. M. J.,* **1**, 1260

187. Singh, P., Kaur, M., and Sachdeva, S., (2012), Mechanism inspired development of rationally designed dihydrofolate reductase inhibitors as anticancer agents. *J. Med. Chem.,* **55**, 6381-6390

188. Tanaka, T. and Knox, W.E., (1959), Nature and mechanism of the tryptophan pyrrolase (peroxidase-oxidase) reaction of P*seudomonas* and of rat liver. *J. Biol. Chem.,* **234**, 1162-1170

189. Moore, G.P. and Sullivan, D.T., (1975), Characterization of multiple forms of kynurenine formidase in *Drosophila melanogaster*. *Biochim. Biophys. Acta,* **397**, 468-477

190. Breton, J., Avanzi, N., Magagnin, S., Covini, N., Magistrelli, G., Cozzi, L., and Isacchi, A., (2000), Functional characterization and mechanism of action of recombinant human kynurenine 3-hydroxylase. *Eur. J. Biochem.,* **267**, 1092-1099

191. Phillips, R.S., Sundararaju, B., and Koushik, S.V., (1998), The catalytic mechanism of kynureninase from *Pseudomonas fluorescens*: Evidence for transient quinonoid and ketimine intermediates from rapid-scanning stopped-flow spectrophotometry. *Biochemistry,* **37**, 8783-8789

192. Cesura, A.M., AlberatiGiani, D., Buchli, R., Broger, C., Kohler, C., Vilbois, F., Lahm, H.W., Heitz, M.P., and Malherbe, P., (1996), *Molecular characterisation of kynurenine pathway enzymes - 3-hydroxyanthranilic-acid dioxygenase and kynurenine aminotransferase*, in *Recent Advances in Tryptophan Research:*

Tryptophan and Serotonin Pathways, G.A. Filippini, C.V.L. Costa, and A. Bertazzo, Editors. p. 477-483.

193. Foster, A.C., White, R.J., and Schwarcz, R., (1986), Synthesis of quinolinic acid by 3-hydroxyanthranilic acid oxygenase in rat brain tissue *in vitro*. *J. Neurochem.,* **47**, 23-30

194. Begley, T.P., Kinsland, C., Mehl, R.A., Osterman, A., and Dorrestein, P., (2001), The biosynthesis of nicotinamide adenine dinucleotides in bacteria. *Vitam. Horm.,* **61**, 103-119

195. Kurnasov, O., Goral, V., Colabroy, K., Gerdes, S., Anantha, S., Osterman, A., and Begley, T.P., (2003), NAD biosynthesis: Identification of the tryptophan to quinolinate pathway in bacteria. *Chem. Biol.,* **10**, 1195-1204

196. Nasu, S., Wicks, F.D., and Gholson, R.K., (1982), L-Aspartate oxidase, a newly discovered enzyme of *Escherichia coli*, is the B protein of quinolinate synthetase. *J. Biol. Chem.,* **257**, 626-632

197. Tedeschi, G., Negri, A., Mortarino, M., Ceciliani, F., Simonic, T., Faotto, L., and Ronchi, S., (1996), L-Aspartate oxidase from *Escherichia coli* .2. Interaction with C-4 dicarboxylic acids and identification of a novel L-aspartate:fumarate oxidoreductase activity. *Eur. J. Biochem.,* **239**, 427-433

198. Nasu, S. and Gholson, R.K., (1981), Replacement of the B-protein requirement of the *Escherichia coli* quinolinate synthetase system by chemically generated iminoaspartate. *Biochem. Biophys. Res. Commun.,* **101**, 533-539

199. Wicks, F.D., Sakakibara, S., Gholson, R.K., and Scott, T.A., (1977), Mode of condensation of aspartic acid and dihydroxyacetone phosphate in quinolinate synthesis in *Escherichia coli*. *Biochim. Biophys. Acta,* **500**, 213-216

200. Cicchillo, R.M., Tu, L., Stromberg, J.A., Hoffart, L.M., Krebs, C., and Booker, S.J., (2005), Escherichia coli quinolinate synthetase does indeed Harbor a 4Fe-4S cluster. *J. Am. Chem. Soc.,* **127**, 7310-7311

201. Ollagnier-de Choudens, S., Loiseau, L., Sanakis, Y., Barras, F., and Fontecave, M., (2005), Quinolinate synthetase, an iron–sulfur enzyme in NAD biosynthesis. *FEBS Lett.,* **579**, 3737-3743

202. Suzuki, N., Carlson, J., Griffith, G., and Gholson, R.K., (1973), Studies on *de novo* biosynthesis of NAD in *Escherichia coli* .5. Properties of quinolinic acid synthetase system. *Biochim. Biophys. Acta,* **304**, 309-315

203. Chan, A., Clemancey, M., Mouesca, J.-M., Amara, P., Hamelin, O., Latour, J.-M., and de Choudens, S.O., (2012), Studies of inhibitor binding to the 4Fe-4S cluster of quinolinate synthase. *Angew. Chem. Int. Ed.,* **51**, 7711-7714

204. Flint, D.H. and Allen, R.M., (1996), Iron-sulfur proteins with nonredox functions. *Chem. Rev.,* **96**, 2315-2334

205. Mattevi, A., (2006), A close look at NAD biosynthesis. *Nat. Struct. Mol. Biol.,* **13**, 563-564

206. Bi, J., Wang, H., and Xie, J., (2011), Comparative genomics of NAD(P) biosynthesis and novel antibiotic drug targets. *J. Cell. Physiol.,* **226**, 331-340

207. Bello, Z., Stitt, B., and Grubmeyer, C., (2010), Interactions at the 2 and 5 positions of 5-phosphoribosyl pyrophosphate are essential in *Salmonella typhimurium* quinolinate phosphoribosyltransferase. *Biochemistry,* **49**, 1377-1387

208. Bhatia, R. and Calvo, K.C., (1996), The sequencing, expression, purification, and steady-state kinetic analysis of quinolinate phosphoribosyl transferase from *Escherichia coli*. *Arch. Biochem. Biophys.,* **325**, 270-278

209. Tao, W., Grubmeyer, C., and Blanchard, J.S., (1996), Transition state structure of *Salmonella typhimurium* orotate phosphoribosyltransferase. *Biochemistry,* **35**, 14-21

210. Rozenberg, A. and Lee, J.K., (2008), Theoretical studies of the quinolinic acid to nicotinic acid mononucleotide transformation. *J. Org. Chem.,* **73**, 9314-9319

211. Mehl, R.A., Kinsland, C., and Begley, T.P., (2000), Identification of the *Escherichia coli* nicotinic acid mononucleotide adenylyltransferase gene. *J. Bacteriol.,* **182**, 4372-4374

212. Ozment, C., Barchue, J., DeLucas, L.J., and Chattopadhyay, D., (1999), Structural study of *Escherichia coli* NAD synthetase: Overexpression, purification, crystallization, and preliminary crystallographic analysis. *J. Struct. Biol.,* **127**, 279-282

213. De Ingeniis, J., Kazanov, M.D., Shatalin, K., Gelfand, M.S., Osterman, A.L., and Sorci, L., (2012), Glutamine versus ammonia utilization in the NAD synthetase family. *PLoS One,* **7**,

214. Kawai, S., Mori, S., Mukai, T., Hashimoto, W., and Murata, K., (2001), Molecular characterization of *Escherichia coli* NAD kinase. *Eur. J. Biochem.,* **268**, 4359-4365

215. Kawai, S. and Murata, K., (2008), Structure and function of NAD kinase and NADP phosphatase: Key enzymes that regulate the intracellular balance of NAD(H) and NADP(H). *Biosci. Biotechnol. Biochem.,* **72**, 919-930

216. Flachmann, R., Kunz, N., Seifert, J., Gutlich, M., Wientjes, F.J., Laufer, A., and Gassen, H.G., (1988), Molecular biology of pyridine nucleotide biosynthesis in *Escherichia coli*. Cloning and characterization of quinolinate synthesis genes *nadA* and *nadB*. *Eur. J. Biochem.,* **175**, 221-228

217. Allibert, P., Willison, J.C., and Vignais, P.M., (1987), Complementation of nitrogen-regulatory (*ntr*-like) mutations in *Rhodobacter capsulatus* by an *Escherichia coli gene*: cloning and sequencing of the gene and characterization of the gene product. *J. Bacteriol.,* **169**, 260-271

218. Young, D.W., (1986), The biosynthesis of the vitamins thiamin, riboflavin, and folic acid. *Nat. Prod. Rep.,* **3**, 395-419

219. Hanson, A.D. and Gregory, J.F., III, (2011), Folate biosynthesis, turnover, and transport in plants. *Annu. Rev. Plant Biol.,* **62**, 105-125

220. Illarionova, V., Eisenreich, W., Fischer, M., Haussmann, C., Romisch, W., Richter, G., and Bacher, A., (2002), Biosynthesis of tetrahydrofolate - Stereochemistry of dihydroneopterin aldolase. *J. Biol. Chem.,* **277**, 28841-28847

221. Burg, A.W. and Brown, G.M., (1968), Biosynthesis of folic acid VIII. Purification and properties of enzyme that catalyzes production of formate from carbon atom 8 of guanosine triphosphate. *J. Biol. Chem.,* **243**, 2349-&

222. Shiota, T., Palumbo, M.P., and Tsai, L., (1967), A chemically prepared formamidopyrimidine derivative of guanosine triphosphate as a possible intermediate in pteridine biosynthesis. *J. Biol. Chem.,* **242**, 1961-&

223. Bracher, A., Eisenreich, W., Schramek, N., Ritz, H., Gotze, E., Herrmann, A., Gutlich, M., and Bacher, A., (1998), Biosynthesis of pteridines - NMR studies on the reaction mechanisms of GTP cyclohydrolase I, pyruvoyltetrahydropterin synthase, and sepiapterin reductase. *J. Biol. Chem.,* **273**, 28132-28141

224. Morrison, S.D., Roberts, S.A., Zegeer, A.M., Montfort, W.R., and Bandarian, V., (2008), A new use for a familiar fold: The X-ray crystal structure of GTP-Bound GTP cyclohydrolase III from *Methanocaldococcus jannaschii* reveals a two metal ion catalytic mechanism. *Biochemistry,* **47**, 230-242

225. Suzuki, Y. and Brown, G.M., (1974), Biosynthesis of folic acid .12. Purification and properties of dihydroneopterin triphosphate pyrophospho-hydrolase. *J. Biol. Chem.,* **249**, 2405-2410

226. Klaus, S.M.J., Wegkamp, A., Sybesma, W., Hugenholtz, J., Gregory, J.F., and Hanson, A.D., (2005), A nudix enzyme removes pyrophosphate from dihydroneopterin

triphosphate in the folate synthesis pathway of bacteria and plants. *J. Biol. Chem.,* **280**, 5274-5280

227. Haussmann, C., Rohdich, F., Schmidt, E., Bacher, A., and Richter, G., (1998), Biosynthesis of pteridines in *Escherichia coli*. Structural and mechanistic similarity of dihydroneopterin-triphosphate epimerase and dihydroneopterin aldolase. *J. Biol. Chem.,* **273**, 17418-17424

228. Talarico, T.L., Dev, I.K., Dallas, W.S., Ferone, R., and Ray, P.H., (1991), Purification and partial characterization of 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase and 7,8-dihydropteroate synthase from *Escherichia coli* MC4100. *J. Bacteriol.,* **173**, 7029-7032

229. Bermingham, A., Bottomley, J.R., Primrose, W.U., and Derrick, J.P., (2000), Equilibrium and kinetic studies of substrate binding to 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Escherichia coli*. *J. Biol. Chem.,* **275**, 17962-17967

230. Maeda, H. and Dudareva, N., (2012), The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annu. Rev. Plant Biol.,* **63**, 73-105

231. Bushby, S.R.M. and Hitching.Gh, (1968), Trimethoprim: A sulphonamide potentiator. *Brit. J. Pharmacol.,* **33**, 72-90

232. Bognar, A.L., Osborne, C., Shane, B., Singer, S.C., and Ferone, R., (1985), Folylpoly-gamma-glutamate synthetase-dihydrofolate synthetase. Cloning and high expression of the *Escherichia coli folC* gene and purification and properties of the gene product. *J. Biol. Chem.,* **260**, 5625-5630

233. Lee, S., Ahn, C., Park, E., Hwang, D.S., and Yim, J., (2002), Biochemical characterization of oligomerization of *Escherichia coli* GTP cyclohydrolase I. *J. Biochem. Mol. Biol.,* **35**, 255-261

234. Talarico, T.L., Ray, P.H., Dev, I.K., Merrill, B.M., and Dallas, W.S., (1992), Cloning, sequence analysis, and overexpression of *Escherichia coli folK*, the gene coding for 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase. *J. Bacteriol.,* **174**, 5971-5977

235. Vos, S., de Jersey, J., and Martin, J.L., (1997), Crystal structure of *Escherichia coli* xanthine phosphoribosyltransferase. *Biochemistry,* **36**, 4125-4134

236. Luk, L.Y.P., Loveridge, E.J., and Allemann, R.K., (2014), Different dynamical effects in mesophilic and hyperthermophilic dihydrofolate reductases. *J. Am. Chem. Soc.,* **136**, 6862-6865

237. Hammes-Schiffer, S. and Benkovic, S.J., (2006), Relating protein motion to catalysis. *Annu. Rev. Biochem.,* **75**, 519-541

238. Glowacki, D.R., Harvey, J.N., and Mulholland, A.J., (2012), Protein dynamics and enzyme catalysis: the ghost in the machine? *Biochem. Soc. Trans.,* **40**, 515-521

239. Brown, K.L. and Li, J., (1998), Activation parameters for the carbon-cobalt bond homolysis of coenzyme B-12 induced by the B-12-dependent ribonucleotide reductase from Lactobacillus leichmannii. *J. Am. Chem. Soc.,* **120**, 9466-9474

240. Pu, J.Z., Ma, S.H., Gao, J.L., and Truhlar, D.G., (2005), Small temperature dependence of the kinetic isotope effect for the hydride transfer reaction catalyzed by *Escherichia coli* dihydrofolate reductase. *J. Phys. Chem. B,* **109**, 8551-8556

241. Świderek, K., Javier Ruiz-Pernía, J., Moliner, V., and Tuñón, I., (2014), Heavy enzymes; experimental and computational insights in enzyme dynamics. *Curr. Opin. Chem. Biol.,* **21**, 11-18

242. Blumenst.M and Raftery, M.A., (1972), ^{31}P and ^{13}C nuclear magnetic resonance studies of nicotinamide adenine dinucleotide and related compounds. *Biochemistry,* **11**, 1643-1648

243. Oberfrank, M., Hull, W.E., and Retey, J., (1984), Synthesis and properties of (4-^{13}C)NAD$^+$ - observation of its binding to yeast alcohol dehydrogenase by ^{13}C-NMR spectroscopy. *Eur. J. Biochem.,* **140**, 157-161

244. Klepp, J., Fallertmuller, A., Grimm, K., Hull, W.E., and Retey, J., (1990), Mechanism of action of urocanase - specific ^{13}C-labeling of the prosthetic NAD$^+$ and revision of the structure of its adduct with imidazolylpropionate. *Eur. J. Biochem.,* **192**, 669-676

245. Bryson, T.A., Wisowaty, J.C., Dunlap, R.B., Fisher, R.R., and Ellis, P.D., (1974), Biological probes I. Carbon-6 labeled nicotinamide. *J. Org. Chem.,* **39**, 1158-1160

246. Bryson, T.A., Wisowaty, J.C., Dunlap, R.B., Fisher, R.R., and Ellis, P.D., (1974), Biological probes II. Ring labeled nicotinamide. *J. Org. Chem.,* **39**, 3436-3438

247. Bryson, T.A., Donelson, D.M., Dunlap, R.B., Fisher, R.R., and Ellis, P.D., (1976), Biological probes III. Methods for carbon-4 and carbon-5 labeling in nicotinamide. *J. Org. Chem.,* **41**, 2066-2067

248. Rising, K.A. and Schramm, V.L., (1994), Enzymatic synthesis of NAD$^+$ with the specific incorporation of atomic labels. *J. Am. Chem. Soc.,* **116**, 6531-6536

249. Jauch, R., Humm, A., Huber, R., and Wahl, M.C., (2005), Structures of *Escherichia coli* NAD synthetase with substrates and products reveal mechanistic rearrangements. *J. Biol. Chem.,* **280**, 15131-15140

250. Wubbolts, M.G., Terpstra, P., Vanbeilen, J.B., Kingma, J., Meesters, H.A.R., and Witholt, B., (1990), Variation of cofactor levels in *Escherichia coli* - sequence-analysis and expression of the *pncB* gene encoding nicotinic acid phosphoribosyltransferase. *J. Biol. Chem.,* **265**, 17665-17672

251. Stephens, P.J., Thomson, A.J., Dunn, J.B.R., Keiderling, T.A., Rawlings, J., Rao, K.K., and Hall, D.O., (1978), Circular dichroism and magnetic circular dichroism of iron-sulfur proteins. *Biochemistry,* **17**, 4770-4778

252. Sweeney, W.V. and Rabinowitz, J.C., (1980), Proteins containing 4Fe-4S clusters - an overview. *Annu. Rev. Biochem.,* **49**, 139-161

253. Jeong, S.S. and Gready, J.E., (1994), A method of preparation and purification of (4*R*)-deuterated-reduced nicotinamide adenine dinucleotide phosphate. *Anal. Biochem.,* **221**, 273-277

254. Griffith, G.R., Chandler, J.L.R., and Gholson, R.K., (1975), Studies on *de novo* biosynthesis of NAD in *Escherichia coli* - separation of *nadB* gene product from *nadA* gene product and its purification. *Eur. J. Biochem.,* **54**, 239-245

255. Schultheisz, H.L., Szymczyna, B.R., Scott, L.G., and Williamson, J.R., (2008), Pathway engineered enzymatic *de novo* purine nucleotide synthesis. *ACS Chem. Biol.,* **3**, 499-511

256. Tolbert, T.J. and Williamson, J.R., (1996), Preparation of specifically deuterated RNA for NMR studies using a combination of chemical and enzymatic synthesis. *J. Am. Chem. Soc.,* **118**, 7929-7940

257. Tolbert, T.J. and Williamson, J.R., (1997), Preparation of specifically deuterated and ^{13}C-labeled RNA for NMR studies using enzymatic synthesis. *J. Am. Chem. Soc.,* **119**, 12100-12108

258. Willison, J.C. and Tissot, G., (1994), The *Escherichia coli efg* gene and the *Rhodobacter capsulatus adgA* gene code for NH$_3$-dependent NAD synthetase. *J. Bacteriol.,* **176**, 3400-3402

259. Stockman, B.J., Lodovice, I.J., Fisher, D.A., McColl, A.S., and Xie, Z., (2007), A nuclear magnetic resonance-based functional assay for nicotinamide adenine dinucleotide synthetase. *J. Biomol. Screening,* **12**, 457-463

260. Dickinson, F.M. and Monger, G.P., (1973), Study of kinetics and mechanism of yeast alcohol dehydrogenase with a variety of substrates. *Biochem. J.,* **131**, 261-270

261. Schultheisz, H.L., Szymczyna, B.R., and Williamson, J.R., (2009), Enzymatic synthesis and structural characterization of ^{13}C, ^{15}N-poly(ADP-ribose). *J. Am. Chem. Soc.,* **131**, 14571-14578

262. Ottolina, G., Riva, S., Carrea, G., Danieli, B., and Buckmann, A.F., (1989), Enzymatic synthesis of 4*R*-^{2}H-NAD(P)H and 4*S*-^{2}H-NAD(P)H and determination of the stereospecificity of 7-alpha-hydroxysteroid and 12-alpha-hydroxysteroid dehydrogenase. *Biochim. Biophys. Acta,* **998**, 173-178

263. Meyer, M.P. and Klinman, J.P., (2011), Investigating Inner-Sphere Reorganization via Secondary Kinetic Isotope Effects in the C-H Cleavage Reaction Catalyzed by Soybean Lipoxygenase: Tunneling in the Substrate Backbone as Well as the Transferred Hydrogen. *J. Am. Chem. Soc.,* **133**, 430-439

264. Lin, H., Zhao, Y., Ellingson, B.A., Pu, J.Z., and Truhlar, D.G., (2005), Temperature dependence of carbon-13 kinetic isotope effects of importance to global climate change. *J. Am. Chem. Soc.,* **127**, 2830-2831

265. Villa, M., Senent, M.L., and Carvajal, M., (2013), Highly correlated *ab initio* study of the low frequency modes of propane and various monosubstituted isotopologues containing D and ^{13}C. *Phys. Chem. Chem. Phys.,* **15**, 10258-10269

266. Gonzalez-James, O.M., Zhang, X., Datta, A., Hrovat, D.A., Borden, W.T., and Singleton, D.A., (2010), Experimental evidence for heavy-atom tunneling in the ring-opening of cyclopropylcarbinyl radical from intramolecular C-12/C-13 kinetic isotope effects. *J. Am. Chem. Soc.,* **132**, 12548-12549

267. Cowart, M., Falzone, C.J., and Benkovic, S.J., (1994), A new synthesis of double-labeled 7,9-^{13}C$_2$ folic acid. *J. Labelled Compd. Radiopharm.,* **34**, 67-71

268. Maunder, P., Finglas, P.M., Mallet, A.I., Mellon, F.A., Razzaque, M.A., Ridge, B., Vahteristo, L., and Witthoft, C., (1999), The synthesis of folic acid, multiply labelled with stable isotopes, for bio-availability studies in human nutrition. *J. Chem. Soc., Perkin Trans. 1*, 1311-1323

269. Roland, S., Ferone, R., Harvey, R.J., Styles, V.L., and Morrison, R.W., (1979), Characteristics and significance of sulfonamides as substrates for *Escherichia coli* dihydropteroate synthase. *J. Biol. Chem.,* **254**, 337-345

270. Suzuki, T., Kurita, H., and Ichinose, H., (2004), GTP cyclohydrolase I utilizes metal-free GTP as its substrate. *Eur. J. Biochem.,* **271**, 349-355

271. Schramm, V.L., (2011), Enzymatic transition states, transition-state analogs, dynamics, thermodynamics, and lifetimes. *Annu. Rev. Biochem.,* **80**, 703-732

9   APPENDICES

9.1 Appendix 1 - Sequencing data

Sequencing was performed by Eurofins (Germany). Sequencing results are shown below for each gene (FASTA sequence), and the alignment score from a BLAST alignment to confirm the correct gene sequence (or corresponding amino acid sequence) had been correctly cloned. Where genes were significantly longer than the reliable trace from N-terminus sequencing, typically above 900 bp, sequencing was repeated from the C-terminus to confirm the correct sequence of the gene.

9.1.1 *nad* gene family sequencing data

9.1.1.1 *nadA*

```
>113590_pET28anadA1_T7
CACAAACCGTGCGTCTTCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCAT
CATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGAGCGTAATGTTTGATCCAGACACGGCG
ATTTATCCTTTCCCCCCGAAGCCGACGCCGTTAAGCATTGATGAAAAAGCGTATTACCGCGAGAAGATAAAACGT
CTGCTAAAAGAACGTAATGCGGTGATGGTTGCCCACTACTATACCGATCCCGAAATTCAACAACTGGCAGAAGAA
ACCGGTGGCTGTATTTCTGATTCTCTGGAAATGGCGCGCTTCGGTGCAAAGCATCCCGCTTCTACTTTGTTAGTC
GCTGGGGTGAGATTTATGGGAGAAACCGCCAAAATTCTCAGTCCGGAAAAAACAATTCTGATGCCGACACTTCAG
GCTGAATGTTCACTGGATCTCGGCTGCCCTGTTGAAGAATTTAACGCATTTTGCGATGCCCATCCCGATCGTACT
GTCGTCGTCTACGCCAACACTTCTGCTGCGGTAAAAGCGCGCGCAGATTGGGTGGTAACTTCAAGCATTGCCGTC
GAACTTATTGATCATCTTGATAGTTTGGGTGAAAAAATCATCTGGGCACCCGACAAACATCTGGGGCGTTACGTG
CAAAAACAGACGGGTGGAGACATTCTATGCTGGCAGGGTGCCTGTATTGTGCATGATGAATTTAAGACTCAGGCG
TTAACCCGCTTGCAAGAAGAATACCCGGATGCTGCCATACTGGTGCATCCAGAATCACCACAAGCTATTGTCGAT
ATGGCGGATGCGGTCGGTTCCACCAGTCAACTGATCGCTGCTGCGAAAACATTGCCACATCAGAGGCTTATTGTG
GCAACCGATCGGGGTATTTTCTACAAAATGCAGCAGCGGTGCCAGATAAAGAGTTACTGGAAGCACCAACCGCAG
TGAGGGTGCAACCTGCCGCAGCTGCGCGCATTGTCCGTGGATGGCCATGAATGGTCTTCAGGCCATCGCAGAGGC
ATTAGAACAGGAAGGAAGCAATCACGAGGTTCATGTTGATGAAA
```

Escherichia coli genome assembly EcRV308Chr ,chromosome : 1
Sequence ID: emb|LM995446.1| Length: 4585620 Number of Matches: 1

Range 1: 745868 to 746843 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1739 bits(1928)	0.0	973/976(99%)	2/976(0%)	Plus/Plus

Features: quinolinate synthase

```
Query  121    ATGAGCGTAATGTTTGATCCAGACACGGCGATTTATCCTTTCCCCCCGAAGCCGACGCCG  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  745868 ATGAGCGTAATGTTTGATCCAGACACGGCGATTTATCCTTTCCCCCCGAAGCCGACGCCG  745927

Query  181    TTAAGCATTGATGAAAAAGCGTATTACCGCGAGAAGATAAAACGTCTGCTAAAAGAACGT  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  745928 TTAAGCATTGATGAAAAAGCGTATTACCGCGAGAAGATAAAACGTCTGCTAAAAGAACGT  745987

Query  241    AATGCGGTGATGGTTGCCCACTACTATACCGATCCCGAAATTCAACAACTGGCAGAAGAA  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  745988 AATGCGGTGATGGTTGCCCACTACTATACCGATCCCGAAATTCAACAACTGGCAGAAGAA  746047

Query  301    ACCGGTGGCTGTATTTCTGATTCTCTGGAAATGGCGCGCTTCGGTGCAAAGCATCCCGCT  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  746048 ACCGGTGGCTGTATTTCTGATTCTCTGGAAATGGCGCGCTTCGGTGCAAAGCATCCCGCT  746107

Query  361    TCTACTTTGTTAGTCGCTGGGGTGAGATTTATGGGAGAAACCGCCAAAATTCTCAGTCCG  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  746108 TCTACTTTGTTAGTCGCTGGGGTGAGATTTATGGGAGAAACCGCCAAAATTCTCAGTCCG  746167
```

```
Query   421    GAAAAAACAATTCTGATGCCGACACTTCAGGCTGAATGTTCACTGGATCTCGGCTGCCCT    480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746168 GAAAAAACAATTCTGATGCCGACACTTCAGGCTGAATGTTCACTGGATCTCGGCTGCCCT    746227

Query   481    GTTGAAGAATTTAACGCATTTTGCGATGCCCATCCCGATCGTACTGTCGTCGTCTACGCC    540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746228 GTTGAAGAATTTAACGCATTTTGCGATGCCCATCCCGATCGTACTGTCGTCGTCTACGCC    746287

Query   541    AACACTTCTGCTGCGGTAAAAGCGCGCGCAGATTGGGTGGTAACTTCAAGCATTGCCGTC    600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746288 AACACTTCTGCTGCGGTAAAAGCGCGCGCAGATTGGGTGGTAACTTCAAGCATTGCCGTC    746347

Query   601    GAACTTATTGATCATCTTGATAGTTTGGGTGAAAAAATCATCTGGGCACCCGACAAACAT    660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746348 GAACTTATTGATCATCTTGATAGTTTGGGTGAAAAAATCATCTGGGCACCCGACAAACAT    746407

Query   661    CTGGGGCGTTACGTGCAAAAACAGACGGGTGGAGACATTCTATGCTGGCAGGGTGCCTGT    720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746408 CTGGGGCGTTACGTGCAAAAACAGACGGGTGGAGACATTCTATGCTGGCAGGGTGCCTGT    746467

Query   721    ATTGTGCATGATGAATTTAAGACTCAGGCGTTAACCCGCTTGCAAGAAGAATACCCGGAT    780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746468 ATTGTGCATGATGAATTTAAGACTCAGGCGTTAACCCGCTTGCAAGAAGAATACCCGGAT    746527

Query   781    GCTGCCATACTGGTGCATCCAGAATCACCACAAGCTATTGTCGATATGGCGGATGCGGTC    840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746528 GCTGCCATACTGGTGCATCCAGAATCACCACAAGCTATTGTCGATATGGCGGATGCGGTC    746587

Query   841    GGTTCCACCAGTCAACTGATCGCTGCTGCGAAAACATTGCCACATCAGAGGCTTATTGTG    900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746588 GGTTCCACCAGTCAACTGATCGCTGCTGCGAAAACATTGCCACATCAGAGGCTTATTGTG    746647

Query   901    GCAACCGATCGGGGTATTTTCTACAAAATGCAGCA-GCGGTGCCAGATAAAGAGTTACTG    959
               |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Sbjct   746648 GCAACCGATCGGGGTATTTTCTACAAAATGCAGCAGGCGGTGCCAGATAAAGAGTTACTG    746707

Query   960    GAAGCACCAACCGCA-GTGAGGGTGCAACCTGCCGCAGCTGCGCGCATTGTCCGTGGATG    1018
               ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746708 GAAGCACCAACCGCAGGTGAGGGTGCAACCTGCCGCAGCTGCGCGCATTGTCCGTGGATG    746767

Query   1019   GCCATGAATGGTCTTCAGGCCATCGCAGAGGCATTAGAACAGGAAGGAAGCAATCACGAG    1078
               ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746768 GCCATGAATGGCCTTCAGGCCATCGCAGAGGCATTAGAACAGGAAGGAAGCAATCACGAG    746827

Query   1079   GTTCATGTTGATGAAA    1094
               ||||||||||||||||
Sbjct   746828 GTTCATGTTGATGAAA    746843
```

>113781_pET28anadA1_T7term
AAGGTAGCTAGCTTCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGTGCGGCC
GCAAGCTTATTATCCACGTAGTGTAGCCGCAAAATCCAGCATACGATTGAGCGGCACCAGCGCCCTCTCTCGCAG
CCTTTCATCAACATGAACCTCGTGATTGCTTCCTTCCTGTTCTAATGCCTCTGCGATGGCCTGAAGGCCATTCAT
GGCCATCCACGGACAATGCGCGCAGCTGCGGCAGGTTGCACCCTCACCTGCGGTTGGTGCTTCCAGTAACTCTTT
ATCTGGCACCGCCTGCTGCATTTTGTAGAAAATACCCCGATCGGTTGCCACAATAAGCCTCTGATGTGGCAATGT
TTTCGCAGCAGCGATCAGTTGACTGGTGGAACCGACCGCATCCGCCATATCGACAATAGCTTGTGGTGATTCTGG
ATGCACCAGTATGGCAGCATCCGGGTATTCTTCTTGCAAGCGGGTTAACGCCTGAGTCTTAAATTCATCATGCAC
AATACAGGCACCCTGCCAGCATAGAATGTCTCCACCCGTCTGTTTTTGCACGTAACGCCCCAGATGTTTGTCGGG
TGCCCAGATGATTTTTTCACCCAAACTATCAAGATGATCAATAAGTTCGACGGCAATGCTTGAAGTTACCACCCA
ATCTGCGCGCGCTTTTACCGCAGCAGAAGTGTTGGCGTAGACGACGACAGTACGATCGGGATGGGCATCGCAAAA
TGCGTTAAATTCTTCAACAGGGCAGCCGAGATCCAGTGAACATTCAGCCTGAAGTGTCGGCATCAGAATTGTTTT
TTCCGGACTGAGAATTTTGGCGGTTTCTCCCATAAATCTCACCCCAGCGACTAACAAAGTAGAAGCGGGATGCTT
TGCACCGAAGCGCGCCATTTCCAGAGAATCAGAAATACAGCCACCGGTTTCTTCTGCCAGTTGTTGAATTTCGGG
ATCGGTATAGTAGTGGGCAACCATCACCGCATTACGTTCTTTAGCAGACGTTTTATCTCTCGCGGTAATACGCTT
TTTCATCATGCTTAACGGCGTCGCTTCGGGGGGGAAAGGATAAAATCGGCCGGTGTC

Escherichia coli genome assembly EcRV308Chr ,chromosome : 1
Sequence ID: emb|LM995446.1| Length: 4585620 Number of Matches: 1

Range 1: 745899 to 746911 GenBank Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1795 bits(1990)	0.0	1009/1013(99%)	4/1013(0%)	Plus/Minus

Features: quinolinate synthase

```
Query   85     TTATCCACGTAGTGTAGCCGCAAAATCCAGCATACGATTGAGCGGCACCAGCGCCCTCTC   144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746911  TTATCCACGTAGTGTAGCCGCAAAATCCAGCATACGATTGAGCGGCACCAGCGCCCTCTC   746852

Query   145    TCGCAGCCTTTCATCAACATGAACCTCGTGATTGCTTCCTTCCTGTTCTAATGCCTCTGC   204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746851  TCGCAGCCTTTCATCAACATGAACCTCGTGATTGCTTCCTTCCTGTTCTAATGCCTCTGC   746792

Query   205    GATGGCCTGAAGGCCATTCATGGCCATCCACGGACAATGCGCGCAGCTGCGGCAGGTTGC   264
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746791  GATGGCCTGAAGGCCATTCATGGCCATCCACGGACAATGCGCGCAGCTGCGGCAGGTTGC   746732

Query   265    ACCCTCACCTGCGGTTGGTGCTTCCAGTAACTCTTTATCTGGCACCGCCTGCTGCATTTT   324
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746731  ACCCTCACCTGCGGTTGGTGCTTCCAGTAACTCTTTATCTGGCACCGCCTGCTGCATTTT   746672

Query   325    GTAGAAAATACCCCGATCGGTTGCCACAATAAGCCTCTGATGTGGCAATGTTTTCGCAGC   384
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746671  GTAGAAAATACCCCGATCGGTTGCCACAATAAGCCTCTGATGTGGCAATGTTTTCGCAGC   746612

Query   385    AGCGATCAGTTGACTGGTGGAACCGACCGCATCCGCCATATCGACAATAGCTTGTGGTGA   444
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746611  AGCGATCAGTTGACTGGTGGAACCGACCGCATCCGCCATATCGACAATAGCTTGTGGTGA   746552

Query   445    TTCTGGATGCACCAGTATGGCAGCATCCGGGTATTCTTCTTGCAAGCGGGTTAACGCCTG   504
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746551  TTCTGGATGCACCAGTATGGCAGCATCCGGGTATTCTTCTTGCAAGCGGGTTAACGCCTG   746492

Query   505    AGTCTTAAATTCATCATGCACAATACAGGCACCCTGCCAGCATAGAATGTCTCCACCCGT   564
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746491  AGTCTTAAATTCATCATGCACAATACAGGCACCCTGCCAGCATAGAATGTCTCCACCCGT   746432

Query   565    CTGTTTTTGCACGTAACGCCCCAGATGTTTGTCGGGTGCCCAGATGATTTTTTCACCCAA   624
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746431  CTGTTTTTGCACGTAACGCCCCAGATGTTTGTCGGGTGCCCAGATGATTTTTTCACCCAA   746372

Query   625    ACTATCAAGATGATCAATAAGTTCGACGGCAATGCTTGAAGTTACCACCCAATCTGCGCG   684
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746371  ACTATCAAGATGATCAATAAGTTCGACGGCAATGCTTGAAGTTACCACCCAATCTGCGCG   746312

Query   685    CGCTTTTACCGCAGCAGAAGTGTTGGCGTAGACGACGACAGTACGATCGGGATGGGCATC   744
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746311  CGCTTTTACCGCAGCAGAAGTGTTGGCGTAGACGACGACAGTACGATCGGGATGGGCATC   746252

Query   745    GCAAAATGCGTTAAATTCTTCAACAGGGCAGCCGAGATCCAGTGAACATTCAGCCTGAAG   804
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746251  GCAAAATGCGTTAAATTCTTCAACAGGGCAGCCGAGATCCAGTGAACATTCAGCCTGAAG   746192

Query   805    TGTCGGCATCAGAATTGTTTTTTCCGGACTGAGAATTTTGGCGGTTTCTCCCATAAATCT   864
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746191  TGTCGGCATCAGAATTGTTTTTTCCGGACTGAGAATTTTGGCGGTTTCTCCCATAAATCT   746132

Query   865    CACCCCAGCGACTAACAAAGTAGAAGCGGGATGCTTTGCACCGAAGCGCGCCATTTCCAG   924
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746131  CACCCCAGCGACTAACAAAGTAGAAGCGGGATGCTTTGCACCGAAGCGCGCCATTTCCAG   746072

Query   925    AGAATCAGAAATACAGCCACCGGTTTCTTCTGCCAGTTGTTGAATTTCGGGATCGGTATA   984
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   746071  AGAATCAGAAATACAGCCACCGGTTTCTTCTGCCAGTTGTTGAATTTCGGGATCGGTATA   746012

Query   985    GTAGTGGGCAACCATCACCGCATTACGTTC-TTTAGCAGACGTTTTATC-TCTCGCGGTA   1042
               ||||||||||||||||||||||||||||||| |||||||||||||||| | ||||||||||
Sbjct   746011  GTAGTGGGCAACCATCACCGCATTACGTTCTTTTAGCAGACGTTTTATCTTCTCGCGGTA   745952
```

```
Query   1043    ATACGCTTTTTCATC-ATGCTTAACGGCGTC-GCTTCGGGGGGAAAGGATAAA   1093
                |||||||||||||||| ||||||||||||||| ||||||||||||||||||||
Sbjct   745951  ATACGCTTTTTCATCAATGCTTAACGGCGTCGGCTTCGGGGGGAAAGGATAAA   745899
```

9.1.1.2 nadB

>113593_pET28anadB1_T7
CCCTTTAATACCTTCTTCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCAT
CATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGAATACTCTCCCTGAACATTCATGTGAC
GTGTTGATTATCGGTAGCGGCGCAGCCGGACTTTCACTGGCGCTACGCCTGGCTGACCAGCATCAGGTCATCGTT
CTAAGTAAAGGCCCGGTAACGGAAGGTTCAACATTTTATGCCCAGGGCGGTATTGCCGCCGTGTTTGATGAAACT
GACAGCATTGACTCGCATGTGGAAGACACATTGATTGCCGGGGCTGGTATTTGCGATCGCCATGCAGTTGAATTT
GTCGCCAGCAATGCACGATCCTGTGTGCAATGCCTAATCGACCAGGGGGTGTTATTTGATACCCACATTCAGCCG
AATGGCGAAGAAAGCTACCATCTGACCCGTGAAGGTGGACATAGTCACCGTCGTATTCTTCATGCCGCCGACGCA
ACTGGTAGAGAAGTAGAAACCACGCTGGTGAGCAAGGCGCTGAACCATCCGAATATTCGCGTGCTGGAGCGCAGC
AACGCGGTTGATCTGATTATTTCTGACAAAATTGGCCTGCCGGGCACGCGACGGGTTGTTGGCGCGTGGGTATGG
AACCGAAATAAAGAAAAGGTGGAAACCTGCCACGCAAAAGCGGTGGTGCTGGCAACCGGCGGTGCGTCGAAGGTT
TATCAGTACACCACCAATCCGGATATTTCTTCTGGCGATGGCATTGCTATGGCGTGCGCGCAGCTGCCGGGTTGC
CAATCTCGAATTTAATCAGTTCCACCCTACCGCGCTGTATCACCCCCAGGCACGCAATTTCCTGTTAACGGAAGC
ACTGCGCGGCGAAGCGCTTATCTCAAGCGCCCGGATGGCACGCGTTTTATGCCCGATTTTGATGTGCGCGGCGAA
CTGGGCCCGCGCGATATGTCGCCCGCGCATTGACCATGAATGAAACGCCTCGGCGCGGATTGTATGTCTGATATC
AGCATAAGCCGCTGATTTATCGCAGCATTTCCCGATGATTATGAAAATGCTCGGCTGGGGATGATCTCACACAGA
CGGTAC

Escherichia coli strain ST540, complete genome
Sequence ID: gb|CP007391.1| Length: 4875682 Number of Matches: 1

Range 1: 1272111 to 1273132 GenBank Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1689 bits(1872)	0.0	999/1022(98%)	19/1022(1%)	Plus/Minus

Features: L-aspartate oxidase

```
Query   121      ATGAATACTCTCCCTGAACATTCATGTGACGTGTTGATTATCGGTAGCGGCGCAGCCGGA   180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1273132  ATGAATACTCTCCCTGAACATTCATGTGACGTGTTGATTATCGGTAGCGGCGCAGCCGGA   1273073

Query   181      CTTTCACTGGCGCTACGCCTGGCTGACCAGCATCAGGTCATCGTTCTAAGTAAAGGCCCG   240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1273072  CTTTCACTGGCGCTACGCCTGGCTGACCAGCATCAGGTCATCGTTCTAAGTAAAGGCCCG   1273013

Query   241      GTAACGGAAGGTTCAACATTTTATGCCCAGGGCGGTATTGCCGCCGTGTTTGATGAAACT   300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1273012  GTAACGGAAGGTTCAACATTTTATGCCCAGGGCGGTATTGCCGCCGTGTTTGATGAAACT   1272953

Query   301      GACAGCATTGACTCGCATGTGGAAGACACATTGATTGCCGGGGCTGGTATTTGCGATCGC   360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1272952  GACAGCATTGACTCGCATGTGGAAGACACATTGATTGCCGGGGCTGGTATTTGCGATCGC   1272893

Query   361      CATGCAGTTGAATTTGTCGCCAGCAATGCACGATCCTGTGTGCAATGCCTAATCGACCAG   420
                 ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Sbjct   1272892  CATGCAGTTGAATTTGTCGCCAGCAATGCACGATCCTGTGTGCAATGGCTAATCGACCAG   1272833

Query   421      GGGGTGTTATTTGATACCCACATTCAGCCGAATGGCGAAGAAAGCTACCATCTGACCCGT   480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1272832  GGGGTGTTATTTGATACCCACATTCAGCCGAATGGCGAAGAAAGCTACCATCTGACCCGT   1272773

Query   481      GAAGGTGGACATAGTCACCGTCGTATTCTTCATGCCGCCGACGCAACTGGTAGAGAAGTA   540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1272772  GAAGGTGGACATAGTCACCGTCGTATTCTTCATGCCGCCGACGCAACTGGTAGAGAAGTA   1272713

Query   541      GAAACCACGCTGGTGAGCAAGGCGCTGAACCATCCGAATATTCGCGTGCTGGAGCGCAGC   600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1272712  GAAACCACGCTGGTGAGCAAGGCGCTGAACCATCCGAATATTCGCGTGCTGGAGCGCAGC   1272653

Query   601      AACGCGGTTGATCTGATTATTTCTGACAAAATTGGCCTGCCGGGCACGCGACGGGTTGTT   660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  1272652  AACGCGGTTGATCTGATTATTTCTGACAAAATTGGCCTGCCGGGCACGCGACGGGTTGTT  1272593
Query  661      GGCGCGTGGGTATGGAACCGAAATAAAGAAAAGGTGGAAACCTGCCACGCAAAAGCGGTG  720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1272592  GGCGCGTGGGTATGGAACCGAAATAAAGAAAAGGTGGAAACCTGCCACGCAAAAGCGGTG  1272533

Query  721      GTGCTGGCAACCGGCGGTGCGTCGAAGGTTTATCAGTACACCACCAATCCGGATATTTCT  780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1272532  GTGCTGGCAACCGGCGGTGCGTCGAAGGTTTATCAGTACACCACCAATCCGGATATTTCT  1272473

Query  781      TCTGGCGATGGCATTGCTATGGCGT-GCGCGCA-GCTGCCGGGTTGCCAATCTCGAATTT  838
                ||||||||||||||||||||||||| ||||||| ||||||||||||||||||||||||||
Sbjct  1272472  TCTGGCGATGGCATTGCTATGGCGTGGCGCGCAGGCTGCCGGGTTGCCAATCTCGAATTT  1272413

Query  839      AATCAGTTCCACCCTACCGCGCTGTATCACCCCCAGGCACGCAATTTCCTGTTAACGGAA  898
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1272412  AATCAGTTCCACCCTACCGCGCTGTATCACCCCCAGGCACGCAATTTCCTGTTAACGGAA  1272353

Query  899      GCACTGCGCGGCGAA-GCGCTTATCTCAAGCGCCCGGATGGCACGCGTTTTATGCCCGAT  957
                |||||||||||||||  |||||||||||||||||||||||||||||||||||||||||||
Sbjct  1272352  GCACTGCGCGGCGAAGGCGCTTATCTCAAGCGCCCGGATGGCACGCGTTTTATGCCCGAT  1272293

Query  958      TTTGATGTGCGCGGCGAACTGGGCCCGCGCGATA-TGTCGCCCGCG-CATTGACCATG-A  1014
                ||||||||||||||||||||||||  |||||||| |||||||||||| |||||||||| |
Sbjct  1272292  TTTGATGTGCGCGGCGAACTGGCCCCGCGCGATATTGTCGCCCGCGCCATTGACCATGAA  1272233

Query  1015     ATGAAACGCCTCGGCGCGGATTGTATGT---CTGATATCAG-CATAAG-CCGCTGATTT-  1068
                ||||||||||||||||||||||||||||     |||||||| ||||||  ||||||||||
Sbjct  1272232  ATGAAACGCCTCGGCGCGGATTGTATGTTCCTTGATATCAGCCATAAGCCCGCTGATTTT  1272173

Query  1069     -ATCG-CAGCATTTCCCGATGA-TTATGAAAA--TGCTC-GGCTGGGGA-TGATCTCACA  1121
                 |||  ||||||||||||||||  |||||||||  |||||  |||||||||  ||||||||
Sbjct  1272172  ATTCGCCAGCATTTCCCGATGATTTATGAAAAATTGCTCGGGCTGGGGATTGATCTCACA  1272113

Query  1122     CA  1123
                ||
Sbjct  1272112  CA  1272111
```

\>113782_pET28anadB1_T7term

```
AAAAGCCTAGCTTCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGTGCGGCCG
CAAGCTTATTATCTGTTTATGTAATGATTGCCGGGAGAAAGGATCGATGGACCGGAATGCGTGAGCAGTTCCGGG
TAATCCAGCGTAAAATGCAGGCCACGACTCTCCTTGCGCATCATCGCACAGCGAACAATCAACTCGGCAACCTGT
ACCAGATTACGCAGCTCCAGCAAATTATTTGAGACGCGGAAATGGGCGTAATATTCGTCTATTTCTTGTTGGAGC
ATGGTTATCCGCCGCAGGGCGCGTTCCAGGCGCTTCGTTGTGCGCACAATGCCAACGTAATCCCACATAAACAGA
CGTAGCTCGTGCCAGTTATGCTGAATTACTACCCGTTCGTCAGGGTTCTCAACGCGGCTTTCATCCCACGGCGGT
AACGTACTGATGTCGTGGGCATAAGGCATACGTCTGGTGATATCTTCCGCCGCCGACCAGCCATAGACCAGACAC
TCCAGCAATGAATTCGAGGCCATGCGGTTAGCGCCGTGTAAGCCGGTATAACTCACCTCGCCAATGGCATACAAG
CCCTCGACGTCCGTACGCCCATGATCATCAACCATTACACCACCGCAGGTATAGTGTGCAGCAGGCACAATCCGC
ACCGGTTCTTGTGTGAGATCAATCCCCAGCCCGAGCAATTTTTCATAAATCATCGGGAAACGCTGGCGAATAAAA
TCAGCGGGCTTATGGCTGATATCAAGGAACATACGATCCCGCGCCGAGGCGTTTCATTTTATTGGTCAATGGCGC
GGGCTACAATATCGCGCGGGTCCAATTCTGTCTCCCCGCATCAAAATCGTTTCTAAAA
```

Escherichia coli strain ST540, complete genome
Sequence ID: gb|CP007391.1| Length: 4875682 Number of Matches: 1

Range 1: 1271502 to 1272296 GenBank Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1346 bits(1492)	0.0	780/799(98%)	4/799(0%)	Plus/Plus

Features: tRNA (adenine-N6)-methyltransferase
L-aspartate oxidase

```
Query  76       CAAGCTTATTATCTGTTTATGTAATGATTGCCGGGAGAAAGGATCGATGGACCGGAATGC  135
                || |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1271502  CAGGCTTTTTATCTGTTTATGTAATGATTGCCGGGAGAAAGGATCGATGGACCGGAATGC  1271561

Query  136      GTGAGCAGTTCCGGGTAATCCAGCGTAAAATGCAGGCCACGACTCTCCTTGCGCATCATC  195
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1271562  GTGAGCAGTTCCGGGTAATCCAGCGTAAAATGCAGGCCACGACTCTCCTTGCGCATCATC  1271621
```

```
Query  196       GCACAGCGAACAATCAACTCGGCAACCTGTACCAGATTACGCAGCTCCAGCAAATTATTT  255
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1271622   GCACAGCGAACAATCAACTCGGCAACCTGTACCAGATTACGCAGCTCCAGCAAATTATTT  1271681

Query  256       GAGACGCGGAAATGGGCGTAATATTCGTCTATTTCTTGTTGGAGCATGGTTATCCGCCGC  315
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1271682   GAGACGCGGAAATGGGCGTAATATTCGTCTATTTCTTGTTGGAGCATGGTTATCCGCCGC  1271741

Query  316       AGGGCGCGTTCCAGGCGCTTCGTTGTGCGCACAATGCCAACGTAATCCCACATAAACAGA  375
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1271742   AGGGCGCGTTCCAGGCGCTTCGTTGTGCGCACAATGCCAACGTAATCCCACATAAACAGA  1271801

Query  376       CGTAGCTCGTGCCAGTTATGCTGAATTACTACCCGTTCGTCAGGGTTCTCAACGCGGCTT  435
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1271802   CGTAGCTCGTGCCAGTTATGCTGAATTACTACCCGTTCGTCAGGGTTCTCAACGCGGCTT  1271861

Query  436       TCATCCCACGGCGGTAACGTACTGATGTCGTGGGCATAAGGCATACGTCTGGTGATATCT  495
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1271862   TCATCCCACGGCGGTAACGTACTGATGTCGTGGGCATAAGGCATACGTCTGGTGATATCT  1271921

Query  496       TCCGCCGCCGACCAGCCATAGACCAGACACTCCAGCAATGAATTCGAGGCCATGCGGTTA  555
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1271922   TCCGCCGCCGACCAGCCATAGACCAGACACTCCAGCAATGAATTCGAGGCCATGCGGTTA  1271981

Query  556       GCGCCGTGTAAGCCGGTATAACTCACCTCGCCAATGGCATACAAGCCCTCGACGTCCGTA  615
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1271982   GCGCCGTGTAAGCCGGTATAACTCACCTCGCCAATGGCATACAAGCCCTCGACGTCCGTA  1272041

Query  616       CGCCCATGATCATCAACCATTACACCACCGCAGGTATAGTGTGCAGCAGGCACAATCCGC  675
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Sbjct  1272042   CGCCCATGATCATCAACCATTACACCACCGCAGGTATAGTGTGCAGCAGGCACAATCGGT  1272101

Query  676       ACCGGTTCTTGTGTGAGATCAATCCCCAGCCCGAGCAATTTTTCATAAATCATCGGGAAA  735
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1272102   ACCGGTTCTTGTGTGAGATCAATCCCCAGCCCGAGCAATTTTTCATAAATCATCGGGAAA  1272161

Query  736       CGCTGGCGAATAAAATCAGCGGGCTTATGGCTGATATCAAGGAACATACGATCCCGCGCC  795
                  ||||||||||||||||||||||||||||||||||||||||||||||||| ||  |||||||
Sbjct  1272162   TGCTGGCGAATAAAATCAGCGGGCTTATGGCTGATATCAAGGAACATACAAT-CCGCGCC  1272220

Query  796       GAGGCGTTTCATTTTATTGGTCAATGGCGCGGGCTACAATATCGCGCGGGTCCAATTCTG  855
                 |||||||||||||| | |||||||||||||||| |||||||||||| ||| |||
Sbjct  1272221   GAGGCGTTTCATTTCA-TGGTCAATGGCGCGGGCGACAATATCGCGCGGGGCCAGTTC--  1272277

Query  856       TCTCCCCGCATCAAAATCG  874
                  |  | | ||||||||||||
Sbjct  1272278   GCCGCGCACATCAAAATCG  1272296
```

9.1.1.3   *nadC*

```
>113597_pET28anadC2_T7
AGGTTAACCGGTACATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCC
ATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGCCGCCTCGCCGCTATAACCCTGACA
CCCGACGTGACGAGCTGCTGGAACGCATTAATCTCGATATCCCCGGCGCGGTGGCCCAGGCGCTGCGGGAAGATT
TAGGCGGAACAGTCGATGCCAACAATGATATTACGGCAAAACTTTTACCGGAAAATTCTCGCTCTCATGCCACGG
TGATCACCCGCGAGAATGGCGTCTTTTGCGGCAAACGCTGGGTTGAAGAGGTGTTTATTCAACTGGCAGGCGACG
ATGTCACCATAATCTGGCATGTGGATGACGGCGATGTCATCAATGCCAATCAATCCTTGTTCGAACTTGAAGGCC
CATCCCGCGTGCTGTTAACGGGCGAACGCACTGCGCTTAATTTTGTGCAAACCCTTTCAGGAGTTGCCAGTAAGG
TACGCCACTATGTCGAATTGCTGGAAGGCACCAACACGCAGTTGTTGGATACGCGCAAAACCTTACCCGGCCTGC
GTTCAGCTCTGAAATACGCGGTACTTTGCGGCGGCGGAGCGAATCACCGTCTGGGGCTTTCTGATGCCTTCCTGA
TCAAAGAAAACCATATTATTGCCTCCGGCTCAGTGCGCCAGGCGGTCGAAAAAGCGTCCTGGCTGCACCCGGATG
CGCCAGTAGAAGTCGAAGTAGAGAATCTGGAAGAACTTGATGAAGCCCTGAAAGCAGGAGCCGATATCATCATGC
TGGATAACTTCGAAACAGAACAGATGCGCGAAGCCGTCAAACGCACCAACGGCAAGGCGCTACTGGAAGTGTCTG
GCAACGTCACTGACAAAACACTGCGTGAATTTGCCGAAACGGGCGTGGACTTTATCTCCGTCGTGCGCTAACTAA
ACACGTACAAGCACTCGACCTTTCAATGCGTTTCGCTAATAAGCTTGCGGCCGCACCTCGAGCACACACACACAC
TGAGATCCGGCTGCTAACAAAGCCCGAAAGAGCTGAGTGCTGCTGCCACCGCTGAGCATACTAGCATAACCCTTT
GGGGCCTCTAACGGTCTTGGAGGG
```

Escherichia coli genome assembly EcRV308Chr ,chromosome : 1
Sequence ID: emb|LM995446.1| Length: 4585620 Number of Matches: 1

Range 1: 117750 to 118645 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1600 bits(1774)	0.0	894/896(99%)	2/896(0%)	Plus/Minus

Features: quinolinate phosphoribosyltransferase (decarboxylating)

```
Query   123     ATGCCGCCTCGCCGCTATAACCCTGACACCCGACGTGACGAGCTGCTGGAACGCATTAAT   182
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   118645  ATGCCGCCTCGCCGCTATAACCCTGACACCCGACGTGACGAGCTGCTGGAACGCATTAAT   118586

Query   183     CTCGATATCCCCGGCGCGGTGGCCCAGGCGCTGCGGGAAGATTTAGGCGGAACAGTCGAT   242
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   118585  CTCGATATCCCCGGCGCGGTGGCCCAGGCGCTGCGGGAAGATTTAGGCGGAACAGTCGAT   118526

Query   243     GCCAACAATGATATTACGGCAAAACTTTTACCGGAAAATTCTCGCTCTCATGCCACGGTG   302
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   118525  GCCAACAATGATATTACGGCAAAACTTTTACCGGAAAATTCTCGCTCTCATGCCACGGTG   118466

Query   303     ATCACCCGCGAGAATGGCGTCTTTTGCGGCAAACGCTGGGTTGAAGAGGTGTTTATTCAA   362
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   118465  ATCACCCGCGAGAATGGCGTCTTTTGCGGCAAACGCTGGGTTGAAGAGGTGTTTATTCAA   118406

Query   363     CTGGCAGGCGACGATGTCACCATAATCTGGCATGTGGATGACGGCGATGTCATCAATGCC   422
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   118405  CTGGCAGGCGACGATGTCACCATAATCTGGCATGTGGATGACGGCGATGTCATCAATGCC   118346

Query   423     AATCAATCCTTGTTCGAACTTGAAGGCCCATCCCGCGTGCTGTTAACGGGCGAACGCACT   482
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   118345  AATCAATCCTTGTTCGAACTTGAAGGCCCATCCCGCGTGCTGTTAACGGGCGAACGCACT   118286

Query   483     GCGCTTAATTTTGTGCAAACCCTTTCAGGAGTTGCCAGTAAGGTACGCCACTATGTCGAA   542
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   118285  GCGCTTAATTTTGTGCAAACCCTTTCAGGAGTTGCCAGTAAGGTACGCCACTATGTCGAA   118226

Query   543     TTGCTGGAAGGCACCAACACGCAGTTGTTGGATACGCGCAAAACCTTACCCGGCCTGCGT   602
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   118225  TTGCTGGAAGGCACCAACACGCAGTTGTTGGATACGCGCAAAACCTTACCCGGCCTGCGT   118166

Query   603     TCAGCTCTGAAATACGCGGTACTTTGCGGCGGCGGAGCGAATCACCGTCTGGGGCTTTCT   662
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   118165  TCAGCTCTGAAATACGCGGTACTTTGCGGCGGCGGAGCGAATCACCGTCTGGGGCTTTCT   118106

Query   663     GATGCCTTCCTGATCAAAGAAAACCATATTATTGCCTCCGGCTCAGTGCGCCAGGCGGTC   722
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   118105  GATGCCTTCCTGATCAAAGAAAACCATATTATTGCCTCCGGCTCAGTGCGCCAGGCGGTC   118046

Query   723     GAAAAAGCGTCCTGGCTGCACCCGGATGCGCCAGTAGAAGTCGAAGTAGAGAATCTGGAA   782
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   118045  GAAAAAGCGTCCTGGCTGCACCCGGATGCGCCAGTAGAAGTCGAAGTAGAGAATCTGGAA   117986

Query   783     GAACTTGATGAAGCCCTGAAAGCAGGAGCCGATATCATCATGCTGGATAACTTCGAAACA   842
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   117985  GAACTTGATGAAGCCCTGAAAGCAGGAGCCGATATCATCATGCTGGATAACTTCGAAACA   117926

Query   843     GAACAGATGCGCGAAGCCGTCAAACGCACCAACGGCAAGGCGCTACTGGAAGTGTCTGGC   902
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   117925  GAACAGATGCGCGAAGCCGTCAAACGCACCAACGGCAAGGCGCTACTGGAAGTGTCTGGC   117866

Query   903     AACGTCACTGACAAAACACTGCGTGAATTTGCCGAAACGGGCGTGGACTTTATCTCCGTC   962
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   117865  AACGTCACTGACAAAACACTGCGTGAATTTGCCGAAACGGGCGTGGACTTTATCTCCGTC   117806

Query   963     -GTGCGCTAACTAAACACGTACAAGCACTCGACCTTTCAATGCG-TTTCGCTAATA   1016
                |||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Sbjct   117805  GGTGCGCTAACTAAACACGTACAAGCACTCGACCTTTCAATGCGTTTTCGCTAATA   117750
```

>113783_pET28anadC2_T7term
TCATGCGGAGCGCTAGCTTCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGTG
CGGCCGCAAGCTTATTAGCGAAAACGCATTGAAAGGTCGAGTGCTTGTACGTGTTTAGTTAGCGCACCGACGGAG
ATAAAGTCCACGCCCGTTTCGGCAAATTCACGCAGTGTTTTGTCAGTGACGTTGCCAGACACTTCCAGTAGCGCC
TTGCCGTTGGTGCGTTTGACGGCTTCGCGCATCTGTTCTGTTTCGAAGTTATCCAGCATGATGATATCGGCTCCT
GCTTTCAGGGCTTCATCAAGTTCTTCCAGATTCTCTACTTCGACTTCTACTGGCGCATCCGGGTGCAGCCAGGAC
GCTTTTTCGACCGCCTGGCGCACTGAGCCGGAGGCAATAATATGGTTTTCTTTGATCAGGAAGGCATCAGAAAGC
CCCAGACGGTGATTCGCTCCGCCGCCGCAAAGTACCGCGTATTTCAGAGCTGAACGCAGGCCGGGTAAGGTTTTG
CGCGTATCCAACAACTGCGTGTTGGTGCCTTCCAGCAATTCGACATAGTGGCGTACCTTACTGGCAACTCCTGAA
AGGGTTTGCACAAAATTAAGCGCAGTGCGTTCGCCCGTTAACAGCACGCGGGATGGGCCTTCAAGTTCGAACAAG
GATTGATTGGCATTGATGACATCGCCGTCATCCACATGCCAGATTATGGTGACATCGTCGCCTGCCAGTTGAATA
AACACCTCTTCAACCCAGCGTTTGCCGCAAAAGACGCCATTCTCGCGGGTGATCACCGTGGCATGAGAGCGAGAA
TTTTCCGGTAAAAGTTTTGCCGTAATATCATTGTTGGCATCGACTGTTCCGCCTAAATCTTCCCGCAGCGCCTGG
GCCACCGCGCCGGGGATATCGAGATTAATGCGTTCCAGCAGCTCGTCACGTCGGGTGTCAGGTTATAGCGGCGAG
GCGGCATATGGCTGCCGCGCGGCACCAGGCCGCTGCTGTGATGATGATGATGATGGCTGCTGCCCATGGTATATC
TCCTTCTTAAAGTTAAACAAAATAT

Escherichia coli genome assembly EcRV308Chr ,chromosome : 1
Sequence ID: emb|LM995446.1| Length: 4585620 Number of Matches: 1

Range 1: 117750 to 118645 GenBank Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1608 bits(1782)	0.0	895/896(99%)	1/896(0%)	Plus/Plus

Features: quinolinate phosphoribosyltransferase (decarboxylating)

```
Query  88      TATTAGCGAAAACGCATTGAAAGGTCGAGTGCTTGTACGTGTTTAGTTAGCGCACCGACG  147
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  117750  TATTAGCGAAAACGCATTGAAAGGTCGAGTGCTTGTACGTGTTTAGTTAGCGCACCGACG  117809

Query  148     GAGATAAAGTCCACGCCCGTTTCGGCAAATTCACGCAGTGTTTTGTCAGTGACGTTGCCA  207
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  117810  GAGATAAAGTCCACGCCCGTTTCGGCAAATTCACGCAGTGTTTTGTCAGTGACGTTGCCA  117869

Query  208     GACACTTCCAGTAGCGCCTTGCCGTTGGTGCGTTTGACGGCTTCGCGCATCTGTTCTGTT  267
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  117870  GACACTTCCAGTAGCGCCTTGCCGTTGGTGCGTTTGACGGCTTCGCGCATCTGTTCTGTT  117929

Query  268     TCGAAGTTATCCAGCATGATGATATCGGCTCCTGCTTTCAGGGCTTCATCAAGTTCTTCC  327
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  117930  TCGAAGTTATCCAGCATGATGATATCGGCTCCTGCTTTCAGGGCTTCATCAAGTTCTTCC  117989

Query  328     AGATTCTCTACTTCGACTTCTACTGGCGCATCCGGGTGCAGCCAGGACGCTTTTTCGACC  387
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  117990  AGATTCTCTACTTCGACTTCTACTGGCGCATCCGGGTGCAGCCAGGACGCTTTTTCGACC  118049

Query  388     GCCTGGCGCACTGAGCCGGAGGCAATAATATGGTTTTCTTTGATCAGGAAGGCATCAGAA  447
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  118050  GCCTGGCGCACTGAGCCGGAGGCAATAATATGGTTTTCTTTGATCAGGAAGGCATCAGAA  118109

Query  448     AGCCCCAGACGGTGATTCGCTCCGCCGCCGCAAAGTACCGCGTATTTCAGAGCTGAACGC  507
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  118110  AGCCCCAGACGGTGATTCGCTCCGCCGCCGCAAAGTACCGCGTATTTCAGAGCTGAACGC  118169

Query  508     AGGCCGGGTAAGGTTTTGCGCGTATCCAACAACTGCGTGTTGGTGCCTTCCAGCAATTCG  567
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  118170  AGGCCGGGTAAGGTTTTGCGCGTATCCAACAACTGCGTGTTGGTGCCTTCCAGCAATTCG  118229

Query  568     ACATAGTGGCGTACCTTACTGGCAACTCCTGAAAGGGTTTGCACAAAATTAAGCGCAGTG  627
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  118230  ACATAGTGGCGTACCTTACTGGCAACTCCTGAAAGGGTTTGCACAAAATTAAGCGCAGTG  118289

Query  628     CGTTCGCCCGTTAACAGCACGCGGGATGGGCCTTCAAGTTCGAACAAGGATTGATTGGCA  687
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  118290  CGTTCGCCCGTTAACAGCACGCGGGATGGGCCTTCAAGTTCGAACAAGGATTGATTGGCA  118349

Query  688     TTGATGACATCGCCGTCATCCACATGCCAGATTATGGTGACATCGTCGCCTGCCAGTTGA  747
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  118350  TTGATGACATCGCCGTCATCCACATGCCAGATTATGGTGACATCGTCGCCTGCCAGTTGA  118409
```

```
Query    748     ATAAACACCTCTTCAACCCAGCGTTTGCCGCAAAAGACGCCATTCTCGCGGGTGATCACC    807
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    118410  ATAAACACCTCTTCAACCCAGCGTTTGCCGCAAAAGACGCCATTCTCGCGGGTGATCACC    118469

Query    808     GTGGCATGAGAGCGAGAATTTTCCGGTAAAAGTTTTGCCGTAATATCATTGTTGGCATCG    867
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    118470  GTGGCATGAGAGCGAGAATTTTCCGGTAAAAGTTTTGCCGTAATATCATTGTTGGCATCG    118529

Query    868     ACTGTTCCGCCTAAATCTTCCCGCAGCGCCTGGGCCACCGCGCCGGGGATATCGAGATTA    927
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    118530  ACTGTTCCGCCTAAATCTTCCCGCAGCGCCTGGGCCACCGCGCCGGGGATATCGAGATTA    118589

Query    928     ATGCGTTCCAGCAGCTCGTCACGTCGGGTGTCA-GGTTATAGCGGCGAGGCGGCAT      982
                 |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Sbjct    118590  ATGCGTTCCAGCAGCTCGTCACGTCGGGTGTCAGGGTTATAGCGGCGAGGCGGCAT      118645
```

9.1.1.4 *nadD*

```
>109348_pET28anaD1_T7
CGTTGAACCGAACATTCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATC
ATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGACTGGTGGACAGCAAATGG
GTCGCGGATCCATGAAATCTTTACAGGCTCTGTTTGGCGGCACCTTTGATCCGGTGCACTATGGTCATCTAAAAC
CCGTGGAAACGCTGGCGAATTTGATTGGTCTGACGCGGGTCACAATCATCCCTAATAATGTTCCTCCGCATCGTC
CCCAGCCGGAAGCGAACAGCGTGCAGCGTAAACACATGCTTGAACTGGCGATTGCCGACAAGCCATTATTTACTC
TTGATGAACGCGAGCTAAAGCGCAATGCCCCCTCTTACACTGCGCAAACACTGAAAGAGTGGCGGCAGGAACAAG
GACCGGACGTGCCGCTGGCGTTTATTATTGGTCAGGATTCACTGCTGACCTTTCCGACCTGGTACGAATACGAAA
CGATACTCGACAATGCACATTTGATCGTCTGTCGGCGTCCAGGTTACCCACTTGAAATGGCGCAACCGCAATACC
AGCAATGGCTGGAAGATCATTTGACACATAACCCGGAAGATCTTCACCTTCAGCCTGCCGGTAAAATTTATCTGG
CTGAAACGCCGTGGTTTAACATCTCGGCGACCATCATCCGCGAACGTTTGCAAAACGGTGAATCATGTGAGGATT
TATTGCCGGAACCGGTACTGACTTACATTAACCAACAAGGCTTGTATCGCTGATAAGCTTGCGGCCGCACTCGAG
CACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCT
GAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATA
TCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGA
CCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTTCCTTTCTCGCCACGTTCGCCGGC
TTCCCGTCAGCTCTAAATCGGGGGGCTCCTTTAGGTTTCGATTAGTGCCTTACGGCACCTCGACCCCAAAACCTT
GATTAGGGTGATCTCACGTTATTGGCATCGGCCCTGGATAGAAACGGTTTTTCGCCCTTTTGAAACGTTTGTGGA
GATTTC
```

Escherichia coli genome assembly EcRV308Chr ,chromosome : 1
Sequence ID: emb|LM995446.1| Length: 4585620 Number of Matches: 1

Range 1: 633712 to 634355 GenBank Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1162 bits(1288)	0.0	644/644(100%)	0/644(0%)	Plus/Minus

Features: nicotinate-mononucleotide adenylyltransferase

```
Query    162     ATGAAATCTTTACAGGCTCTGTTTGGCGGCACCTTTGATCCGGTGCACTATGGTCATCTA    221
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    634355  ATGAAATCTTTACAGGCTCTGTTTGGCGGCACCTTTGATCCGGTGCACTATGGTCATCTA    634296

Query    222     AAACCCGTGGAAACGCTGGCGAATTTGATTGGTCTGACGCGGGTCACAATCATCCCTAAT    281
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    634295  AAACCCGTGGAAACGCTGGCGAATTTGATTGGTCTGACGCGGGTCACAATCATCCCTAAT    634236

Query    282     AATGTTCCTCCGCATCGTCCCCAGCCGGAAGCGAACAGCGTGCAGCGTAAACACATGCTT    341
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    634235  AATGTTCCTCCGCATCGTCCCCAGCCGGAAGCGAACAGCGTGCAGCGTAAACACATGCTT    634176

Query    342     GAACTGGCGATTGCCGACAAGCCATTATTTACTCTTGATGAACGCGAGCTAAAGCGCAAT    401
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    634175  GAACTGGCGATTGCCGACAAGCCATTATTTACTCTTGATGAACGCGAGCTAAAGCGCAAT    634116

Query    402     GCCCCCTCTTACACTGCGCAAACACTGAAAGAGTGGCGGCAGGAACAAGGACCGGACGTG    461
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    634115  GCCCCCTCTTACACTGCGCAAACACTGAAAGAGTGGCGGCAGGAACAAGGACCGGACGTG    634056

Query    462     CCGCTGGCGTTTATTATTGGTCAGGATTCACTGCTGACCTTTCCGACCTGGTACGAATAC    521
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    634055  CCGCTGGCGTTTATTATTGGTCAGGATTCACTGCTGACCTTTCCGACCTGGTACGAATAC    633996
```

```
Query   522    GAAACGATACTCGACAATGCACATTTGATCGTCTGTCGGCGTCCAGGTTACCCACTTGAA   581
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   633995 GAAACGATACTCGACAATGCACATTTGATCGTCTGTCGGCGTCCAGGTTACCCACTTGAA   633936

Query   582    ATGGCGCAACCGCAATACCAGCAATGGCTGGAAGATCATTTGACACATAACCCGGAAGAT   641
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   633935 ATGGCGCAACCGCAATACCAGCAATGGCTGGAAGATCATTTGACACATAACCCGGAAGAT   633876

Query   642    CTTCACCTTCAGCCTGCCGGTAAAATTTATCTGGCTGAAACGCCGTGGTTTAACATCTCG   701
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   633875 CTTCACCTTCAGCCTGCCGGTAAAATTTATCTGGCTGAAACGCCGTGGTTTAACATCTCG   633816

Query   702    GCGACCATCATCCGCGAACGTTTGCAAAACGGTGAATCATGTGAGGATTTATTGCCGGAA   761
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   633815 GCGACCATCATCCGCGAACGTTTGCAAAACGGTGAATCATGTGAGGATTTATTGCCGGAA   633756

Query   762    CCGGTACTGACTTACATTAACCAACAAGGCTTGTATCGCTGATA    805
               ||||||||||||||||||||||||||||||||||||||||||||
Sbjct   633755 CCGGTACTGACTTACATTAACCAACAAGGCTTGTATCGCTGATA    633712
```

9.1.1.5 nadE

>109352_pET28anadE2_T7
TTTTTAAACGGTTATTCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATC
ATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGACTGGTGGACAGCAAATGG
GTCGCGGATCCATGACATTGCAACAACAAATAATAAAGGCGCTGGGCGCAAAACCGCAGATTAATGCTGAAGAGG
AAATTCGTCGTAGTGTCGATTTTCTGAAAAGCTACCTGCAAACTTATCCGTTCATTAAATCACTGGTGCTCGGGA
TCAGCGGCGGTCAGGACTCCACGCTTGCCGGAAAGCTGTGCCAGATGGCGATTAATGAGCTGCGCCTGGAAACCG
GCAACGAATCACTGCAATTTATTGCCGTACGCCTGCCCTATGGTGTTCAGGCCGACGAACAAGATTGCCAGGATG
CCATTGCCTTTATTCAACCGGATCGCGTATTAACCGTTAATATCAAGGGCGCGGTATTGGCCAGCGAACAGGCAT
TGCGGGAAGCAGGCATTGAACTGAGCGATTTTGTCCGTGGCAATGAAAAAGCGCGTGAGCGGATGAAAGCACAAT
ATAGCATTGCGGGTATGACCAGCGGTGTCGTGGTGGGCACCGATCATGCAGCAGAAGCCATTACCGGATTCTTCA
CTAAATATGGTGACGGCGGTACGGACATTAATCCGCTGTATCGTCTCAACAAACGTCAGGGTAAACAGTTACTGG
CGGCATTAGCTTGCCCGGAACACCTTTATAAGAAAGCGCCAACGGCCGATCTGGAAGATGATCGCCCTTCTCTGC
CAGATGAAGTGGCACTCGGCGTGACCTATGACAATATCGACGACTATCTGGAAAGGGAAAAACGTACCTCAACAG
GTCGCCAGAACAATAGAGAACTGGTATCTGAAAACCGAACATAAACGCCGTCCGCCAATTACCGTTTTCGATGAT
TTCTGGAAAAGTAATAATTTGCACATATTGGTTGTGTGAAAAGCTTGCGGCCGCACTCGAGCACCACCACCACCA
CCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAGCTGAGTTGGCTGCTGCCACCGCTGAGCATACTAGCATA
CCCCTGGGGCCTCTAACGGGTCTGGAGGGTTTTTGCTGAAGGAGACTTATATCGATGGCGATGGGACCGCGCCTG
TACGTCCATACCGCGAGTGTGTGTACGCGCAGCGGGACGCTACACATGTCCAAGCGCCCTA

NAD synthetase, NH3/glutamine-dependent [Escherichia coli]
Sequence ID: ref|WP_001403659.1| Length: 275 Number of Matches: 2
▷ See 1 more title(s)

Range 1: 1 to 239 GenPept Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
489 bits(1260)	5e-170	Compositional matrix adjust.	239/239(100%)	239/239(100%)	0/239(0%)	+3

```
Query   162  MTLQQQIIKALGAKPQINAEEEIRRSVDFLKSYLQTYPFIKSLVLGISGGQDSTLAGKLC   341
             MTLQQQIIKALGAKPQINAEEEIRRSVDFLKSYLQTYPFIKSLVLGISGGQDSTLAGKLC
Sbjct   1    MTLQQQIIKALGAKPQINAEEEIRRSVDFLKSYLQTYPFIKSLVLGISGGQDSTLAGKLC   60

Query   342  QMAINELRLETGNESLQFIAVRLPYGVQADEQDCQDAIAFIQPDRVLTVNIKGAVLASEQ   521
             QMAINELRLETGNESLQFIAVRLPYGVQADEQDCQDAIAFIQPDRVLTVNIKGAVLASEQ
Sbjct   61   QMAINELRLETGNESLQFIAVRLPYGVQADEQDCQDAIAFIQPDRVLTVNIKGAVLASEQ   120

Query   522  ALREAGIELSDFVRGNEKARERMKAQYSIAGMTSGVVVGTDHAAEAITGFFTKYGDGGTD   701
             ALREAGIELSDFVRGNEKARERMKAQYSIAGMTSGVVVGTDHAAEAITGFFTKYGDGGTD
Sbjct   121  ALREAGIELSDFVRGNEKARERMKAQYSIAGMTSGVVVGTDHAAEAITGFFTKYGDGGTD   180

Query   702  INPLYRLNKRQGKQLLAALACPEHLYKKAPTADLEDDRPSLPDEVALGVTYDNIDDYLE   878
             INPLYRLNKRQGKQLLAALACPEHLYKKAPTADLEDDRPSLPDEVALGVTYDNIDDYLE
Sbjct   181  INPLYRLNKRQGKQLLAALACPEHLYKKAPTADLEDDRPSLPDEVALGVTYDNIDDYLE   239
```

9.1.1.6   nadK

```
>109964_pNCOnadK1
AGGTAGCAAATCTTTCGCTGATCTCATGGTCATGTTTCACCAAAGAACACCACGATCACGACACCACCTTTGAAG
AAATTATTATTGGTTCCGCGGCATCTGGTCTGGTGCCACGCGGATCCATGAATAATCATTTCAAGTGTATTGGCA
TTGTGGGACACCCACGGCACCCCACTGCACTGACAACACATGAAATGCTCTACCGCTGGCTGTGCACAAAAGGTT
ACGAGGTCATCGTTGAGCAACAAATCGCTCACGAACTGCAACTGAAGAATGTGAAAACTGGCACGCTCGCGGAGA
TTGGGCAACTAGCTGATCTCGCGGTAGTCGTTGGTGGCGACGGTAATATGCTGGGCGCGGCACGCACACTCGCCC
GTTACGATATTAAAGTTATTGGAATCAACCGTGGCAACCTGGGTTTCCTGACTGACCTTGACCCCGATAACGCCC
AGCAACAGTTAGCCGATGTGCTGGAAGGCCACTACATCAGCGAGAAACGTTTTTTGCTGGAAGCGCAAGTCTGTC
AGCAAGATTGCCAGAAACGCATCAGCACCGCGATAAATGAAGTGGTGCTTCATCCAGGCAAAGTGGCGCATATGA
TTGAGTTCGAAGTGTATATCGACGAGATCTTTGCGTTTTCTCAGCGATCTGATGGACTAATTATTTCGACGCCAA
CAGGCTCCACCGCCTATTCCCTCTCTGCAGGCGGTCCTATTCTGACCCCCTCTCTGGATGCGATTACCCTGGTGC
CCATGTTCCCGCATACGTTGTCAGCACGACCACTGGTCATAAACAGCAGCAGCACGATCCGTCTGCGTTTTTCGC
ATCGCCGTAACGACCTGAAATCAGTTGCGACAGCCAGATAGCACTGCCGATTCAGGAAGGTGAAGATGTCCTGAT
TCGTCGCTGTGATTACCATCTGAATCTGATTCATCCGAAAGATTACAGTTATTTCAACACATTAAGCACCAAGCT
CGGCTGGTCAAAAAATTATTTCTAATAAGCTTAATTAGCTGAGCTTGGACTCTGTTGATAGATCCAGTAATGACC
TCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGTGGGG
```

Escherichia coli genome assembly EcRV308Chr ,chromosome : 1

Sequence ID: emb|LM995446.1| Length: 4585620 Number of Matches: 1

Range 1: 2693359 to 2694238 GenBank Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1566 bits(1736)	0.0	876/880(99%)	1/880(0%)	Plus/Plus

Features: NAD kinase

```
Query   123      ATGAATAATCATTTCAAGTGTATTGGCATTGTGGGACACCCACGGCACCCCACTGCACTG  182
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2693359  ATGAATAATCATTTCAAGTGTATTGGCATTGTGGGACACCCACGGCACCCCACTGCACTG  2693418

Query   183      ACAACACATGAAATGCTCTACCGCTGGCTGTGCACAAAAGGTTACGAGGTCATCGTTGAG  242
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2693419  ACAACACATGAAATGCTCTACCGCTGGCTGTGCACAAAAGGTTACGAGGTCATCGTTGAG  2693478

Query   243      CAACAAATCGCTCACGAACTGCAACTGAAGAATGTGAAAACTGGCACGCTCGCGGAGATT  302
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2693479  CAACAAATCGCTCACGAACTGCAACTGAAGAATGTGAAAACTGGCACGCTCGCGGAGATT  2693538

Query   303      GGGCAACTAGCTGATCTCGCGGTAGTCGTTGGTGGCGACGGTAATATGCTGGGCGCGGCA  362
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2693539  GGGCAACTAGCTGATCTCGCGGTAGTCGTTGGTGGCGACGGTAATATGCTGGGCGCGGCA  2693598

Query   363      CGCACACTCGCCCGTTACGATATTAAAGTTATTGGAATCAACCGTGGCAACCTGGGTTTC  422
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2693599  CGCACACTCGCCCGTTACGATATTAAAGTTATTGGAATCAACCGTGGCAACCTGGGTTTC  2693658

Query   423      CTGACTGACCTTGACCCCGATAACGCCCAGCAACAGTTAGCCGATGTGCTGGAAGGCCAC  482
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2693659  CTGACTGACCTTGACCCCGATAACGCCCAGCAACAGTTAGCCGATGTGCTGGAAGGCCAC  2693718

Query   483      TACATCAGCGAGAAACGTTTTTTGCTGGAAGCGCAAGTCTGTCAGCAAGATTGCCAGAAA  542
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2693719  TACATCAGCGAGAAACGTTTTTTGCTGGAAGCGCAAGTCTGTCAGCAAGATTGCCAGAAA  2693778

Query   543      CGCATCAGCACCGCGATAAATGAAGTGGTGCTTCATCCAGGCAAAGTGGCGCATATGATT  602
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2693779  CGCATCAGCACCGCGATAAATGAAGTGGTGCTTCATCCAGGCAAAGTGGCGCATATGATT  2693838

Query   603      GAGTTCGAAGTGTATATCGACGAGATCTTTGCGTTTTCTCAGCGATCTGATGGACTAATT  662
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2693839  GAGTTCGAAGTGTATATCGACGAGATCTTTGCGTTTTCTCAGCGATCTGATGGACTAATT  2693898

Query   663      ATTTCGACGCCAACAGGCTCCACCGCCTATTCCCTCTCTGCAGGCGGTCCTATTCTGACC  722
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2693899  ATTTCGACGCCAACAGGCTCCACCGCCTATTCCCTCTCTGCAGGCGGTCCTATTCTGACC  2693958

Query   723      CCCTCTCTGGATGCGATTACCCTGGTGCCCATGTTCCCGCATACGTTGTCAGCACGACCA  782
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2693959  CCCTCTCTGGATGCGATTACCCTGGTGCCCATGTTCCCGCATACGTTGTCAGCACGACCA  2694018
```

```
Query   783       CTGGTCATAAACAGCAGCAGCACGATCCGTCTGCGTTTTTCGCATCGCCGTAACGACCT-  841
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2694019   CTGGTCATAAACAGCAGCAGCACGATCCGTCTGCGTTTTTCGCATCGCCGTAACGACCTG  2694078

Query   842       GAAATCAGTTGCGACAGCCAGATAGCACTGCCGATTCAGGAAGGTGAAGATGTCCTGATT  901
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2694079   GAAATCAGTTGCGACAGCCAGATAGCACTGCCGATTCAGGAAGGTGAAGATGTCCTGATT  2694138

Query   902       CGTCGCTGTGATTACCATCTGAATCTGATTCATCCGAAAGATTACAGTTATTTCAACACA  961
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   2694139   CGTCGCTGTGATTACCATCTGAATCTGATTCATCCGAAAGATTACAGTTATTTCAACACA  2694198

Query   962       TTAAGCACCAAGCTCGGCTGGTCAAAAAATTATTTCTAAT    1001
                  |||||||||||||||||||||||||||||| |  ||||||||
Sbjct   2694199   TTAAGCACCAAGCTCGGCTGGTCAAAAAAATTATTCTAAT    2694238
```

9.1.1.7 pncB

>109345_pET28apncB1_T7
ATGATTGCCGTCATTCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCA
TCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGACTGGTGGACAGCAAATGGG
TCGCGGATCCATGACACAATTCGCTTCTCCTGTTCTGCACTCGTTGCTGGATACAGATGCTTATAAGTTGCATAT
GCAGCAAGCCGTGTTTCATCACTATTACGATGTGCATGTCGCGGCGGAGTTTCGTTGCCGAGGTGACGATCTGCT
GGGTATTTATGCCGATGCTATTCGTGAACAGGTTCAGGCGATGCAGCACCTGCGCCTGCAGGATGATGAATATCA
GTGGCTTTCTGCCCTGCCTTTCTTTAAGGCCGACTATCTTAACTGGTTACGCGAGTTCCGCTTTAACCCGGAACA
AGTCACCGTGTCCAACGATAATGGCAAGCTGGATATTCGTTTAAGCGGCCCGTGGCGTGAAGTCATCCTCTGGGA
AGTTCCTTTGCTGGCGGTTATCAGTGAAATGGTACATCGCTATCGCTCACCGCAGGCCGACGTTGCGCAAGCCCT
CGACACGCTGGAAAGCAAATTAGTCGACTTCTCGGCGTTAACCGCCGGTCTTGATATGTCGCGCTTCCATCTGAT
GGATTTTGGCACCCGTCGCCGTTTTTCTCGCGAAGTACAAGAAACCATCGTTAAGCGTCTGCAACAGGAATCCTG
GTTTGTGGGCACCAGCAACTACGATCTGGCGCGTCGGCTTTCCCTCACGCCGATGGGAACACAGGCACACGAATG
GTTCCAGGCACATCAGCAAATCAGCCCGGATCTAGCCAACAGCCAGCGAGCTGCACTTGCTGCCTGGCTGGAAGA
GTATCCCGACCAACTTGGCATTGCATTAACCGACTGCATCACTATGGATGCTTTCCTGCGTGATTTCGGTGTCGA
GTTCGCTAGTCGGTATCAGGCCTGCGTCATGACTCTGGCGACCCGGTTGAATGGGGTGAAAAGCCATTGCACATT
ATGAAAAGCTGGGAATGATCCACAGAGTAAAACGCTGGGTTTCTCCTGACATCTGGATTTACGCAAGCGGTTGAG
CTATACCGCCAACTTCTCTTCCCGCGTGCAATTAGGTTTGATTGGGAACTCGGCCTGAACTGGCGAATATCCCCA
GGTTAAGCCTGAATTTGCCATTAAGTGGTCAGGTTACGTTAATCGGGTGCGAACTTCTTGAACAGGCCTGTGCAA
AAACCTATCTCTGC

Escherichia coli genome assembly EcRV308Chr ,chromosome : 1
Sequence ID: emb|LM995446.1| Length: 4585620 Number of Matches: 1

Range 1: 954344 to 955483 GenBank Graphics

Score	Expect	Identities	Gaps	Strand
1792 bits(1986)	0.0	1103/1150(96%)	27/1150(2%)	Plus/Minus

Features: nicotinate phosphoribosyltransferase

```
Query   153      GCGGATCCATGACACAATTCGCTTCTCCTGTTCTGCACTCGTTGCTGGATACAGATGCTT  212
                 ||| | | | |||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   955483   GCGCACCTATGACACAATTCGCTTCTCCTGTTCTGCACTCGTTGCTGGATACAGATGCTT  955424

Query   213      ATAAGTTGCATATGCAGCAAGCCGTGTTTCATCACTATTACGATGTGCATGTCGCGGCGG  272
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   955423   ATAAGTTGCATATGCAGCAAGCCGTGTTTCATCACTATTACGATGTGCATGTCGCGGCGG  955364

Query   273      AGTTTCGTTGCCGAGGTGACGATCTGCTGGGTATTTATGCCGATGCTATTCGTGAACAGG  332
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   955363   AGTTTCGTTGCCGAGGTGACGATCTGCTGGGTATTTATGCCGATGCTATTCGTGAACAGG  955304

Query   333      TTCAGGCGATGCAGCACCTGCGCCTGCAGGATGATGAATATCAGTGGCTTTCTGCCCTGC  392
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   955303   TTCAGGCGATGCAGCACCTGCGCCTGCAGGATGATGAATATCAGTGGCTTTCTGCCCTGC  955244

Query   393      CTTTCTTTAAGGCCGACTATCTTAACTGGTTACGCGAGTTCCGCTTTAACCCGGAACAAG  452
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  955243  CTTTCTTTAAGGCCGACTATCTTAACTGGTTACGCGAGTTCCGCTTTAACCCGGAACAAG  955184

Query  453     TCACCGTGTCCAACGATAATGGCAAGCTGGATATTCGTTTAAGCGGCCCGTGGCGTGAAG  512
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  955183  TCACCGTGTCCAACGATAATGGCAAGCTGGATATTCGTTTAAGCGGCCCGTGGCGTGAAG  955124

Query  513     TCATCCTCTGGGAAGTTCCTTTGCTGGCGGTTATCAGTGAAATGGTACATCGCTATCGCT  572
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  955123  TCATCCTCTGGGAAGTTCCTTTGCTGGCGGTTATCAGTGAAATGGTACATCGCTATCGCT  955064

Query  573     CACCGCAGGCCGACGTTGCGCAAGCCCTCGACACGCTGGAAAGCAAATTAGTCGACTTCT  632
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  955063  CACCGCAGGCCGACGTTGCGCAAGCCCTCGACACGCTGGAAAGCAAATTAGTCGACTTCT  955004

Query  633     CGGCGTTAACCGCCGGTCTTGATATGTCGCGCTTCCATCTGATGGATTTTGGCACCCGTC  692
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  955003  CGGCGTTAACCGCCGGTCTTGATATGTCGCGCTTCCATCTGATGGATTTTGGCACCCGTC  954944

Query  693     GCCGTTTTTCTCGCGAAGTACAAGAAACCATCGTTAAGCGTCTGCAACAGGAATCCTGGT  752
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954943  GCCGTTTTTCTCGCGAAGTACAAGAAACCATCGTTAAGCGTCTGCAACAGGAATCCTGGT  954884

Query  753     TTGTGGGCACCAGCAACTACGATCTGGCGCGTCGGCTTTCCCTCACGCCGATGGGAACAC  812
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954883  TTGTGGGCACCAGCAACTACGATCTGGCGCGTCGGCTTTCCCTCACGCCGATGGGAACAC  954824

Query  813     AGGCACACGAATGGTTCCAGGCACATCAGCAAATCAGCCCGGATCTAGCCAACAGCCAGC  872
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954823  AGGCACACGAATGGTTCCAGGCACATCAGCAAATCAGCCCGGATCTAGCCAACAGCCAGC  954764

Query  873     GAGCTGCACTTGCTGCCTGGCTGGAAGAGTATCCCGACCAACTTGGCATTGCATTAACCG  932
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954763  GAGCTGCACTTGCTGCCTGGCTGGAAGAGTATCCCGACCAACTTGGCATTGCATTAACCG  954704

Query  933     ACTGCATCACTATGGATGCTTTCCTGCGTGATTTCGGTGTCGAGTTCGCTAGTCGGTATC  992
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954703  ACTGCATCACTATGGATGCTTTCCTGCGTGATTTCGGTGTCGAGTTCGCTAGTCGGTATC  954644

Query  993     A-GGCCTGCGTCATGACTCTGGCGACCCGGTTGAATGGGGTG-AAAAGCCATTGCACATT  1050
               | ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Sbjct  954643  AGGGCCTGCGTCATGACTCTGGCGACCCGGTTGAATGGGGTGAAAAAGCCATTGCACATT  954584

Query  1051   ATGAAAAGCTGGGAA-TGATCCACAGAGTAAAACGCTGGGTTTCTCCTGAC-ATCTGGAT  1108
               ||||||||||||||| ||||||||||||||||||||||| ||||| |||||| ||||||||
Sbjct  954583  ATGAAAAGCTGGGAATTGATCCACAGAGTAAAACGCTGGTTTTCT-CTGACAATCTGGAT  954525

Query  1109   TTACGC-AAGCGGTTGAGCTATACCGCCAACTTCTCTTCCCGCGTGCAATTAGGTTT---  1164
               |||||| ||||||||||||||||||||||| ||||||||||||||||||||| ||||
Sbjct  954524  TTACGCAAAGCGGTTGAGCTATACCGCC-ACTTCTCTTCCCGCGTGCAATTAAGTTTTGG  954466

Query  1165   GATTGGGAACTCGGCCTGAACTGGCGAATAT-CCCCAGGTTAA--GCCTGAAT-TTGCCA  1220
                ||||||| |||| |||||| || ||| |||| ||||||| ||    |||||||| || ||
Sbjct  954465  TATTGGG-ACTC-GCCTGACCT-GCG-ATATCCCCCAGGTAAAACCCCTGAATATTGTCA  954410

Query  1221   TTAAGTGGTCAG---GTTACGTTAATCGGGT-GCG-AACTTCTTGAACAGGCCTGTGCAA  1275
               |||||| |   || |  || ||| ||| | ||| ||| ||||| || |||| |||| ||||
Sbjct  954409  TTAAGTTGGTAGAGTGTAACGGTAAACCGGTGGCGAAACTTTCTG-ACAGCCCTG-GCAA  954352

Query  1276   AAACCTATCT  1285
               ||  ||||||
Sbjct  954351  AA--CTATCT  954344
```

>200911_109431_pET28apncB1_T7-TERM
TTGGTGCTCGTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAAGC
TTTCGGAATTAAAGGTAAAAAAAGGGAGCCGATTAGCTCCCTGATGATATTAACTGGCTTTTTTAATATGCGGAA
GGTCGAACGCTTTGCGCAGCGCCCGAACAAACGCTTTATCATGGCAGATAGTTTTGCCAGGGCTGTCAGAAAGTT
TCGCCACCGGTTTACCGTTACACTCTACCAACTTAATGACAATATTCAGGGGTTTTACCTGGGGGGATATCGCAGG
TCAGGCGAGTCCCAATACCAAAACTTAATTGCACGCGGGAAGAGAAGTGGCGGTATAGCTCAACCGCTTTGCGTA
AATCCAGATTGTCAGAGAAAACCAGCGTTTTACTCTGTGGATCAATTCCCAGCTTTTCATAATGTGCAATGGCTT
TTTCACCCCATTCAACCGGGTCGCCAGAGTCATGACGCAGGCCCTGATACCGACTAGCGAACTCGACACCGAAAT
CACGCAGGAAAGCATCCATAGTGATGCAGTCGGTTAATGCAATGCCAAGTTGGTCGGGATACTCTTCCAGCCAGG
CAGCAAGTGCAGCTCGCTGGCTGTTGGCTAGATCCGGGCTGATTTGCTGATGTGCCTGGAACCATTCGTGTGCCT
GTGTTCCCATCGGCGTGAGGGAAAGCCGACGCGCCAGATCGTAGTTGCTGGTGCCCACAAACCAGGATTCCTGTT
GCAGACGCTTAACGATGGTTTCTTGTACTTCGCGAGAAAAACGGCGACGGGTGCCAAAATCCATCAGATGGAAGC
GCGACATATCAAGACCGGCGGTTAACGCCGAGAAGTCGACTAATTTGCTTTCCAGCGTGTCGAGGGCTTGCGCAA
CGTCGGCCTGCGGTGAGCGATAGCGATGTACCATTTCACTGATAACCGCCAGCAAAGGAACTTCCCAGAGGATGA
CTTCACGCCACGGGCCGCTTAAACGATATCCAGCTTGCCATTATCGTTGGGACACGGTGACTTGTTCCGGGTTAA
AGCGGACTCGCGTAGCAGTTAGATAGTCGCTTAAAGAAACGCAGGCAGAAGCCACTGATATTCATCATCTGCAGC
GCAGGTGCTGCATCGCCCTGACTGATCACGATAGCATCGGCATAAATACCCAGCGAATCGTCACCTCGCACGAAC
TTCGGCTCGCGCGAACAATTCCACACTCGCGTG

Escherichia coli genome assembly EcRV308Chr ,chromosome : 1
Sequence ID: emb|LM995446.1| Length: 4585620 Number of Matches: 1

Range 1: 954217 to 955352 GenBank Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1909 bits(2116)	0.0	1116/1138(98%)	15/1138(1%)	Plus/Plus

Features: nicotinate phosphoribosyltransferase

```
Query  69      CGCAAGCTTTCGGAATTAAAGGTaaaaaaaGGGAGCCGATTAGCTCCCTGATGATATTAA  128
               || ||| ||||||||||||||||       |||||||||||||||||||||||||||||||
Sbjct  954217  CGAAAGATTTCGGAATTAAAGGTAAAAAAAGGGAGCCGATTAGCTCCCTGATGATATTAA  954276

Query  129     CTGGCTTTTTTAATATGCGGAAGGTCGAACGCTTTGCGCAGCGCCCGAACAAACGCTTTA  188
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954277  CTGGCTTTTTTAATATGCGGAAGGTCGAACGCTTTGCGCAGCGCCCGAACAAACGCTTTA  954336

Query  189     TCATGGCAGATAGTTTTGCCAGGGCTGTCAGAAAGTTTCGCCACCGGTTTACCGTTACAC  248
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954337  TCATGGCAGATAGTTTTGCCAGGGCTGTCAGAAAGTTTCGCCACCGGTTTACCGTTACAC  954396

Query  249     TCTACCAACTTAATGACAATATTCAGGGGTTTTACCTGGGGGATATCGCAGGTCAGGCGA  308
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954397  TCTACCAACTTAATGACAATATTCAGGGGTTTTACCTGGGGGATATCGCAGGTCAGGCGA  954456

Query  309     GTCCCAATACCAAAACTTAATTGCACGCGGGAAGAGAAGTGGCGGTATAGCTCAACCGCT  368
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954457  GTCCCAATACCAAAACTTAATTGCACGCGGGAAGAGAAGTGGCGGTATAGCTCAACCGCT  954516

Query  369     TTGCGTAAATCCAGATTGTCAGAGAAAACCAGCGTTTTACTCTGTGGATCAATTCCCAGC  428
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954517  TTGCGTAAATCCAGATTGTCAGAGAAAACCAGCGTTTTACTCTGTGGATCAATTCCCAGC  954576

Query  429     TTTTCATAATGTGCAATGGCTTTTTCACCCCATTCAACCGGGTCGCCAGAGTCATGACGC  488
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954577  TTTTCATAATGTGCAATGGCTTTTTCACCCCATTCAACCGGGTCGCCAGAGTCATGACGC  954636

Query  489     AGGCCCTGATACCGACTAGCGAACTCGACACCGAAATCACGCAGGAAAGCATCCATAGTG  548
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954637  AGGCCCTGATACCGACTAGCGAACTCGACACCGAAATCACGCAGGAAAGCATCCATAGTG  954696

Query  549     ATGCAGTCGGTTAATGCAATGCCAAGTTGGTCGGGATACTCTTCCAGCCAGGCAGCAAGT  608
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954697  ATGCAGTCGGTTAATGCAATGCCAAGTTGGTCGGGATACTCTTCCAGCCAGGCAGCAAGT  954756

Query  609     GCAGCTCGCTGGCTGTTGGCTAGATCCGGGCTGATTTGCTGATGTGCCTGGAACCATTCG  668
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  954757  GCAGCTCGCTGGCTGTTGGCTAGATCCGGGCTGATTTGCTGATGTGCCTGGAACCATTCG  954816
```

```
Query   669    TGTGCCTGTGTTCCCATCGGCGTGAGGGAAAGCCGACGCGCCAGATCGTAGTTGCTGGTG   728
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   954817 TGTGCCTGTGTTCCCATCGGCGTGAGGGAAAGCCGACGCGCCAGATCGTAGTTGCTGGTG   954876

Query   729    CCCACAAACCAGGATTCCTGTTGCAGACGCTTAACGATGGTTTCTTGTACTTCGCGAGAA   788
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   954877 CCCACAAACCAGGATTCCTGTTGCAGACGCTTAACGATGGTTTCTTGTACTTCGCGAGAA   954936

Query   789    AAACGGCGACGGGTGCCAAAATCCATCAGATGGAAGCGCGACATATCAAGACCGGCGGTT   848
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   954937 AAACGGCGACGGGTGCCAAAATCCATCAGATGGAAGCGCGACATATCAAGACCGGCGGTT   954996

Query   849    AACGCCGAGAAGTCGACTAATTTGCTTTCCAGCGTGTCGAGGGCTTGCGCAACGTCGGCC   908
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   954997 AACGCCGAGAAGTCGACTAATTTGCTTTCCAGCGTGTCGAGGGCTTGCGCAACGTCGGCC   955056

Query   909    TGCGGTGAGCGATAGCGATGTACCATTTCACTGATAACCGCCAGCAAAGGAACTTCCCAG   968
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   955057 TGCGGTGAGCGATAGCGATGTACCATTTCACTGATAACCGCCAGCAAAGGAACTTCCCAG   955116

Query   969    AGGATGACTTCACGCCACGGGCCGCTTAAACG-ATATCCAGCTTGCCATTATCGTTGGGA   1027
               ||||||||||||||||||||||||||||||||| |||||||||||||||||||||| |||
Sbjct   955117 AGGATGACTTCACGCCACGGGCCGCTTAAACGAATATCCAGCTTGCCATTATCGTT-GGA   955175

Query   1028   CACGGTGACTTGTTCCGGGTTAAAGCGG-ACTCGCGT-AGCAGTT-AGATAGTCG--CTT   1082
               |||||||||||||||||||||||||||| ||||||||| | ||||| |||||||||   |||
Sbjct   955176 CACGGTGACTTGTTCCGGGTTAAAGCGGAACTCGCGTAACCAGTTAAGATAGTCGGCCTT   955235

Query   1083   AAAGAAACGCA-GGCAG-AAGCCACTGATATTCATCAT-CTGCA-GCGCAGGTGCTGCAT   1138
               |||||||| ||| ||||| |||||||||||||||||||| ||||| ||||||||||||||
Sbjct   955236 AAAGAAAGGCAGGGCAGAAAGCCACTGATATTCATCATCCTGCAGGCGCAGGTGCTGCAT   955295

Query   1139   CGCCCTGA--CTGATCACG-ATAGCATCGGCATAAATACCCAGCGAATCGTCACCTCG    1193
               || |||||   ||| ||||| ||||||||||||||||||||||    ||||||||||||
Sbjct   955296 CG-CCTGAACCTGTTCACGAATAGCATCGGCATAAATACCCAGCAGATCGTCACCTCG    955352
```

9.1.2 *fol* gene family sequencing data

folP and *folC* were synthesised by Epoch gene synthesis service rather than being cloned from *E. coli* directly.

9.1.2.1 *folE*

```
>110315_pET21dfolE21_T7
CCGTAGAAAGTACATTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGATGCCATCACTC
AGTAAAGAAGCGGCCCTGGTTCATGAAGCGTTAGTTGCGCGAGGACTGGAAACACCGCTGCGCCCGCCCGTGCAT
GAAATGGATAACGAAACGCGCAAAAGCCTTATTGCTGGTCATATGACCGAAATCATGCAGCTGCTGAATCTCGAC
CTGGCTGATGACAGTTTGATGGAAACGCCGCATCGCATCGCTAAAATGTATGTCGATGAAATTTTCTCCGGTCTG
GATTACGCCAATTTCCCGAAAATCACCCTCATTGAAAACAAAATGAAGGTCGATGAAATGGTCACCGTGCGCGAT
ATCACTCTGACCAGCACCTGTGAACACCATTTTGTTACCATCGATGGCAAAGCGACGGTGGCCTATATCCCGAAA
GATTCGGTGATCGGTCTGTCAAAAATTAACCGCATTGTGCAGTTCTTTGCCCAGCGTCCGCAGGTGCAGGAACGT
CTGACGCAGCAAATTCTTATTGCGCTACAAACGCTGCTGGGCACCAATAACGTGGCTGTCTCGATCGACGCAGTG
CATTACTGCGTGAAGGCGCGTGGCATCCGCGATGCAACCAGTGCCACGACAACGACCTCTCTTGGTGGATTGTTC
AAATCCAGTCAGAATACGCGCCACGAGTTTCTGCGCGCTGTGCGTCATCACAACTGATAAAAGCTTGCGGCCGCA
CTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCC
ACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGA
ACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCA
GCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCTTTCTCGCCACGTTCGCC
GGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGG
```

Escherichia coli genome assembly EcRV308Chr ,chromosome : 1
Sequence ID: emb|LM995446.1| Length: 4585620 Number of Matches: 1

Range 1: 2191036 to 2191709 GenBank Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1207 bits(1338)	0.0	672/674(99%)	0/674(0%)	Plus/Minus

Features: GTP cyclohydrolase I

```
Query  64       ATGCCATCACTCAGTAAAGAAGCGGCCCTGGTTCATGAAGCGTTAGTTGCGCGAGGACTG  123
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2191709  ATGCCATCACTCAGTAAAGAAGCGGCCCTGGTTCATGAAGCGTTAGTTGCGCGAGGACTG  2191650

Query  124      GAAACACCGCTGCGCCCGCCCGTGCATGAAATGGATAACGAAACGCGCAAAAGCCTTATT  183
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2191649  GAAACACCGCTGCGCCCGCCCGTGCATGAAATGGATAACGAAACGCGCAAAAGCCTTATT  2191590

Query  184      GCTGGTCATATGACCGAAATCATGCAGCTGCTGAATCTCGACCTGGCTGATGACAGTTTG  243
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2191589  GCTGGTCATATGACCGAAATCATGCAGCTGCTGAATCTCGACCTGGCTGATGACAGTTTG  2191530

Query  244      ATGGAAACGCCGCATCGCATCGCTAAAATGTATGTCGATGAAATTTTCTCCGGTCTGGAT  303
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2191529  ATGGAAACGCCGCATCGCATCGCTAAAATGTATGTCGATGAAATTTTCTCCGGTCTGGAT  2191470

Query  304      TACGCCAATTTCCCGAAAATCACCCTCATTGAAAACAAAATGAAGGTCGATGAAATGGTC  363
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2191469  TACGCCAATTTCCCGAAAATCACCCTCATTGAAAACAAAATGAAGGTCGATGAAATGGTC  2191410

Query  364      ACCGTGCGCGATATCACTCTGACCAGCACCTGTGAACACCATTTTGTTACCATCGATGGC  423
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2191409  ACCGTGCGCGATATCACTCTGACCAGCACCTGTGAACACCATTTTGTTACCATCGATGGC  2191350

Query  424      AAAGCGACGGTGGCCTATATCCCGAAAGATTCGGTGATCGGTCTGTCAAAAATTAACCGC  483
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2191349  AAAGCGACGGTGGCCTATATCCCGAAAGATTCGGTGATCGGTCTGTCAAAAATTAACCGC  2191290

Query  484      ATTGTGCAGTTCTTTGCCCAGCGTCCGCAGGTGCAGGAACGTCTGACGCAGCAAATTCTT  543
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2191289  ATTGTGCAGTTCTTTGCCCAGCGTCCGCAGGTGCAGGAACGTCTGACGCAGCAAATTCTT  2191230

Query  544      ATTGCGCTACAAACGCTGCTGGGCACCAATAACGTGGCTGTCTCGATCGACGCAGTGCAT  603
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  2191229  ATTGCGCTACAAACGCTGCTGGGCACCAATAACGTGGCTGTCTCGATCGACGCGGTGCAT  2191170

Query  604      TACTGCGTGAAGGCGCGTGGCATCCGCGATGCAACCAGTGCCACGACAACGACCTCTCTT  663
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2191169  TACTGCGTGAAGGCGCGTGGCATCCGCGATGCAACCAGTGCCACGACAACGACCTCTCTT  2191110

Query  664      GGTGGATTGTTCAAATCCAGTCAGAATACGCGCCACGAGTTTCTGCGCGCTGTGCGTCAT  723
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2191109  GGTGGATTGTTCAAATCCAGTCAGAATACGCGCCACGAGTTTCTGCGCGCTGTGCGTCAT  2191050

Query  724      CACAACTGATAAAA    737
                |||||||||| |||
Sbjct  2191049  CACAACTGATTAAA  2191036
```

9.1.2.2 *folB*

Final sequences after three rounds of SDM, as original BLAST search did not align with any known sequences in the database.

```
>115165_pET28afolBWT2_T7
TATGGAATTCCCGTACATTTCCCCTTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGC
AGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGATATTGTATTTATAGAGCAA
CTTTCGGTAATCACCACTATTGGTGTTTACGACTGGGAACAGACCATCGAACAGAAGTTAGTGTTCGATATCGAA
ATGGCGTGGGATAACCGTAAAGCGGCGAAAAGTGATGATGTGGCGGATTGCCTCAGTTACGCTGACATTGCAGAA
ACGGTGGTCAACCACGTCGAGGGGGCGCGTTTTGCGCTGGTGGAACGCGTGGCTGAAGAGGTGGCGGAGCTGCTG
TTAGCACGCTTCAACTCGCCGTGGGTGCGTATCAAACTCAGCAAGCCAGGCGCAGTGGCGCGGGCGGCGAATGTT
GGCGTAATCATTGAGCGTGGCAATAATCTGAAAGAAAATAATTAATAAGGATCCGAATTCGAGCTCCGTCGACAA
GCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGA
GTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTT
GCTGAAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTG
GTGGTTACGCGCAGCGT
```

MULTISPECIES: dihydroneopterin triphosphate 2'-epimerase [Enterobacteriaceae]
Sequence ID: ref|WP_001469253.1| Length: 122 Number of Matches: 1
▷ See 7 more title(s)

Range 1: 1 to 122 GenPept Graphics   ▽ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
200 bits(508)	4e-61	Compositional matrix adjust.	122/122(100%)	122/122(100%)	0/122(0%)	+1

```
Query  127  MDIVFIEQLSVITTIGVYDWEQTIEQKLVFDIEMAWDNRKAAKSDDVADCLSYADIAETV  306
            MDIVFIEQLSVITTIGVYDWEQTIEQKLVFDIEMAWDNRKAAKSDDVADCLSYADIAETV
Sbjct  1    MDIVFIEQLSVITTIGVYDWEQTIEQKLVFDIEMAWDNRKAAKSDDVADCLSYADIAETV  60

Query  307  VNHVEGarfalvervaeevaelllarfNSPWVRIKLSKPGAVARAANVGVIIERGNNLKE  486
            VNHVEGARFALVERVAEEVAELLLARFNSPWVRIKLSKPGAVARAANVGVIIERGNNLKE
Sbjct  61   VNHVEGARFALVERVAEEVAELLLARFNSPWVRIKLSKPGAVARAANVGVIIERGNNLKE  120

Query  487  NN  492
            NN
Sbjct  121  NN  122
```

9.1.2.3 *folK*

```
>110587_pET28afolK12_T7
GTATATCCGGAAACATTTCCCCTTCTTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCA
GCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGACTGGTGGACAGC
AAATGGGTCGCGGATCCTTATTACCATTTGTTTAATTTGTCAAATGCTCTTGTATGTAAGATTTGACGCAACATC
TCCCCATCAGGAAACACCAACTCCGGCGCGATTTCAAACAGCGGCCACAGCATAAATCCACGATTCTTCATATCG
TAGTGCGGAACGGTCAGGCGTTCAGTATTTATCACTTCATTACCAAACAGCATGATGTCGAGATCCAGCGTGCGT
GGTCCCCAGCGTTCAGCTTTGCGGACGCGACCTTGCTGCAATTCAATACGCTGTGTGTGATTGAGTAGCTCTTCA
GGTGCAAGAGAGGTTTCCAGCGCCACGGCTGCGTTTAAGTAATCGGGTTGATTTTGCGGCCCCAGCGGTGGGGTG
CGGTAAAACGAAGAAACGGTAAGAATGTGGCTTTCAGGGATATCGCCTAATGCTTTCAGGGCAGCATTGACCTGC
TCCAGCGGAGAGGCCAGATTGCTGCCTATGGCAATATACGCCACTGTCATATGGTACCCAAGGGCGAATTCCACA
TTGGTCGCTGCAGCCCGGGGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCCCCACC
ACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAAC
TAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGGTTTTTTGCTGAAAGGAGGAACTATATCCCGGATT
GGCGAATGGGACGCGCCCCTGTACCGCCCCTTAAGCGCGGCGGGAGTGGGTGGGTTACGCGCACCGTGACCGCT
ACACTTGCCAGCGCCCTACGCCCGCTTCCTTTTGCTTTTCTTTCCTTCCTTTCTCGGCCCCTTTCGCCCGGGCTT
TCCCCGTCAAGCTCTATAATCCGGG
```

2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase [Escherichia coli]
Sequence ID: ref|WP_000215142.1| Length: 159 Number of Matches: 1
▷ See 5 more title(s)

Range 1: 1 to 159 GenPept Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
328 bits(841)	3e-109	Compositional matrix adjust.	159/159(100%)	159/159(100%)	0/159(0%)	-3

```
Query  650  MTVAYIAIGSNLASPLEQVNAALKALGDIPESHILTVSSFYRTPPLGPQNQPDYLNAAVA  471
            MTVAYIAIGSNLASPLEQVNAALKALGDIPESHILTVSSFYRTPPLGPQNQPDYLNAAVA
Sbjct    1  MTVAYIAIGSNLASPLEQVNAALKALGDIPESHILTVSSFYRTPPLGPQNQPDYLNAAVA   60

Query  470  LETSLAPEELLNHTQRIELQQGRVRKAERWGPRTLDLDIMLFGNEVINTERLTVPHYDMK  291
            LETSLAPEELLNHTQRIELQQGRVRKAERWGPRTLDLDIMLFGNEVINTERLTVPHYDMK
Sbjct   61  LETSLAPEELLNHTQRIELQQGRVRKAERWGPRTLDLDIMLFGNEVINTERLTVPHYDMK  120

Query  290  NRGFMLWPLFEIAPELVFPDGEMLRQILHTRAFDKLNKW   174
            NRGFMLWPLFEIAPELVFPDGEMLRQILHTRAFDKLNKW
Sbjct  121  NRGFMLWPLFEIAPELVFPDGEMLRQILHTRAFDKLNKW   159
```

9.1.2.4 *folP*

```
>folP
CATATGAAACTGTTCGCACAAGGTACTTCTCTGGATCTGTCTCACCCACATGTGATGGGTATCCTGAATGTAACT
CCGGATAGCTTCAGCGATGGCGGCACCCACAACAGCCTGATCGACGCCGTTAAACATGCGAACCTGATGATTAAT
GCCGGCGCGACCATCATCGACGTTGGCGGCGAGTCCACCCGTCCGGGTGCCGCTGAAGTTAGCGTTGAAGAAGAA
CTGCAGCGTGTTATTCCGGTTGTAGAAGCAATCGCGCAGCGCTTTGAAGTTTGGATTTCCGTTGACACTTCTAAA
CCGGAAGTTATTCGTGAATCCGCTAAAGTTGGCGCTCACATTATTAACGATATCCGCTCCCTGAGCGAACCGGGT
GCCCTGGAAGCTGCAGCAGAAACCGGTCTGCCAGTGTGCCTGATGCACATGCAGGGTAACCCGAAAACTATGCAG
GAAGCCCCGAAATACGATGACGTGTTCGCCGAAGTAAACCGCTACTTTATTGAACAGATCGCTCGCTGCGAACAG
GCAGGTATCGCGAAAGAGAAACTGCTGCTGGACCCGGGCTTCGGTTTCGGCAAGAACCTGTCCCACAACTACAGC
CTGCTGGCCCGTCTGGCTGAATTCCACCACTTCAACCTGCCGCTGCTGGTGGGTATGAGCCGCAAATCCATGATT
GGCCAGCTGCTGAATGTGGGTCCGTCTGAACGTCTGAGCGGTAGCCTGGCTTGTGCTGTGATTGCAGCGATGCAG
GGCGCACATATCATCCGTGTACACGATGTTAAAGAAACCGTTGAAGCGATGCGCGTTGTGGAAGCTACCCTGTCC
GCAAAAGAGAACAAACGCTACGAATAAAAGCTT
```

dihydropteroate synthase [Escherichia coli]
Sequence ID: ref|WP_024242762.1| Length: 282 Number of Matches: 1

Range 1: 1 to 190 GenPept Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
394 bits(1012)	8e-136	Compositional matrix adjust.	190/190(100%)	190/190(100%)	0/190(0%)	+2

```
Query   62  MKLFAQGTSLDLSHPHVMGILNVTPDSFSDGGTHNSLIDAVKHANLMINAGATIIDVGGE  241
            MKLFAQGTSLDLSHPHVMGILNVTPDSFSDGGTHNSLIDAVKHANLMINAGATIIDVGGE
Sbjct    1  MKLFAQGTSLDLSHPHVMGILNVTPDSFSDGGTHNSLIDAVKHANLMINAGATIIDVGGE   60

Query  242  STRPGAAEVSVEEELQRVIPVVEAIAQRFEVWISVDTSKPEVIRESAKVGAHIINDIRSL  421
            STRPGAAEVSVEEELQRVIPVVEAIAQRFEVWISVDTSKPEVIRESAKVGAHIINDIRSL
Sbjct   61  STRPGAAEVSVEEELQRVIPVVEAIAQRFEVWISVDTSKPEVIRESAKVGAHIINDIRSL  120

Query  422  SEPGALEAAAETGLPVCLMHMQGNPKTMQEAPKYDDVFAEVNRYFIEQIARCEQAGIAKE  601
            SEPGALEAAAETGLPVCLMHMQGNPKTMQEAPKYDDVFAEVNRYFIEQIARCEQAGIAKE
Sbjct  121  SEPGALEAAAETGLPVCLMHMQGNPKTMQEAPKYDDVFAEVNRYFIEQIARCEQAGIAKE  180

Query  602  KLLLDPGFGF   631
            KLLLDPGFGF
Sbjct  181  KLLLDPGFGF   190
```

9.1.2.5 *folC*

```
>folC
CATATGATCATTAAACGTACTCCGCAAGCTGCATCTCCACTGGCATCTTGGCTGAGCTACCTGGAAAACCTGCAC
AGCAAAACTATCGACCTGGGCCTGGAACGTGTTTCCCTGGTAGCGGCTCGTCTGGGTGTTCTGAAACCGGCGCCA
TTCGTGTTCACCGTTGCAGGCACTAACGGCAAGGGTACCACTTGTCGCACCCTGGAGTCTATCCTGATGGCGGCA
GGCTACAAAGTAGGTGTTTACTCTTCTCCACACCTGGTTCGCTACACTGAACGTGTTCGTGTGCAGGGTCAGGAG
CTGCCGGAAAGCGCTCACACCGCTTCTTTTGCTGAAATCGAATCTGCGCGTGGTGATATCTCCCTGACTTACTTT
GAATACGGTACCCTGAGCGCACTGTGGCTGTTCAAACAGGCCCAACTGGATGTTGTGATCCTGGAAGTAGGCCTG
GGCGGTCGCCTGGACGCTACCAACATCGTGGATGCCGATGTGGCCGTAGTTACCTCTATCGCACTGGACCACACC
GACTGGCTGGGCCCGGACCGTGAATCTATTGGCCGCGAAAAGGCCGGCATCTTCCGCTCCGAAAAGCCGGCGATT
GTGGGCGAGCCGGAAATGCCGTCCACCATCGCGGACGTTGCGCAGGAGAAAGGTGCTCTGCTGCAGCGTCGCGGT
GTTGAGTGGAACTACTCCGTTACTGACCACGACTGGGCTTTCAGCGACGCTCACGGTACTCTGGAGAACCTGCCA
CTGCCACTGGTTCCGCAGCCGAACGCTGCAACCGCGCTGGCAGCACTGCGTGCATCTGGCCTGGAAGTATCTGAA
AACGCCATTCGCGATGGCATCGCCAGCGCCATTCTGCCGGGTCGTTTCCAGATCGTGTCCGAGAGCCCGCGTGTA
ATCTTCGATGTGGCACACAATCCGCACGCAGCGGAATACCTGACTGGTCGTATGAAAGCTCTGCCGAAGAATGGT
CGTGTTCTGGCTGTGATCGGCATGCTGCATGACAAGGATATCGCAGGTACCCTGGCGTGGCTGAAATCTGTGGTT
GACGATTGGTACTGTGCTCCACTGGAAGGCCCACGTGGTGCGACTGCGGAGCAGCTGCTGGAACACCTGGGTAAC
GGCAAATCTTTTGACTCCGTGGCTCAAGCATGGGACGCTGCTATGGCAGATGCTAAGGCGGAAGACACCGTTCTG
GTTTGCGGTTCTTTTCACACTGTTGCCCACGTTATGGAAGTAATCGATGCTCGTCGCAGCGGTGGTAAATAAAAG
CTT
```

bifunctional folylpolyglutamate synthase/ dihydrofolate synthase [Escherichia coli]
Sequence ID: ref|WP_032219939.1| Length: 422 Number of Matches: 1
▷ See 1 more title(s)

Range 1: 1 to 213 GenPept Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
431 bits(1109)	5e-148	Compositional matrix adjust.	213/213(100%)	213/213(100%)	0/213(0%)	+3

```
Query  66   MIIKRTPQAASPLASWLSYLENLHSKTIDLGLERVSLVAARLGVLKPAPFVFTVAGTNGK  245
            MIIKRTPQAASPLASWLSYLENLHSKTIDLGLERVSLVAARLGVLKPAPFVFTVAGTNGK
Sbjct  1    MIIKRTPQAASPLASWLSYLENLHSKTIDLGLERVSLVAARLGVLKPAPFVFTVAGTNGK  60

Query  246  GTTCRTLESILMAAGYKVGVYSSPHLVRYTERVRVQGQELPESAHTASFAEIESARGDIS  425
            GTTCRTLESILMAAGYKVGVYSSPHLVRYTERVRVQGQELPESAHTASFAEIESARGDIS
Sbjct  61   GTTCRTLESILMAAGYKVGVYSSPHLVRYTERVRVQGQELPESAHTASFAEIESARGDIS  120

Query  426  LTYFEYGTLSALWLFKQAQLDVVILEVGLGGRLDATNIVDADVAVVTSIALDHTDWLGPD  605
            LTYFEYGTLSALWLFKQAQLDVVILEVGLGGRLDATNIVDADVAVVTSIALDHTDWLGPD
Sbjct  121  LTYFEYGTLSALWLFKQAQLDVVILEVGLGGRLDATNIVDADVAVVTSIALDHTDWLGPD  180

Query  606  RESIGREKAGIFRSEKPAIVGEPEMPSTIADVA  704
            RESIGREKAGIFRSEKPAIVGEPEMPSTIADVA
Sbjct  181  RESIGREKAGIFRSEKPAIVGEPEMPSTIADVA  213
```

9.1.3 *gpt* sequencing data

```
>113666_pET28agpt2_T7
TTTTGAAAACCAAAACATTTTTCCCTTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGC
AGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGAGCGAAAAATACATCGTCACC
TGGGACATGTTGCAGATCCATGCACGTAAACTCGCAAGCCGACTGATGCCTTCTGAACAATGGAAAGGCATTATT
GCCGTAAGCCGTGGCGGTCTGGTACCGGGTGCGTTACTGGCGCGTGAACTGGGTATTCGTCATGTCGATACCGTT
TGTATTTCCAGCTACGATCACGACAACCAGCGCGAGCTTAAAGTGCTGAAACGCGCAGAAGGCGATGGCGAAGGC
TTCATCGTTATTGATGACCTGGTGGATACCGGTGGTACTGCGGTTGCGATTCGTGAAATGTATCCAAAAGCGCAC
TTTGTCACCATCTTCGCAAAACCGGCTGGTCGTCCGCTGGTTGATGACTATGTTGTTGATATCCCGCAAGATACC
TGGATCGAACAGCCGTGGGATATGGGCGTCGTATTCGTCCCGCCAATCTCCGGTCGCTAATAAGCTTGCGGCCGC
ACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGC
CACCGCTGAGCAATAAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAG
GAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCAGGTGTGGTGGTTACG
CGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACG
TTCGCCGGCTTTCCCCGTCAAGCCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACCGGCACC
TCGACCCCAAAAAACTTGATTTAGGGTGAATGGTTCACCGTAGTGGCCATCGCCCTGATTAAACGGTTTTCGCCC
```

TTGACGTTGGAGTCCACGTTCTTTATAGTGGACTTCTGTTCCATACTTGAACAACCCTCAAACCTATCTCGTCTA
TTCCTTTGAATTAATAACGAATTGGGCCGATTTCCGGCCTATTGTTTAAAAAATGGACCTGAATT

Escherichia coli strain ST540, complete genome
Sequence ID: gb|CP007391.1| Length: 4875682 Number of Matches: 1

Range 1: 3705276 to 3705738 GenBank Graphics   ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
830 bits(920)	0.0	462/463(99%)	0/463(0%)	Plus/Minus

Features: xanthine-guanine phosphoribosyltransferase

```
Query  124      CATATGAGCGAAAAATACATCGTCACCTGGGACATGTTGCAGATCCATGCACGTAAACTC  183
                || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3705738  CACATGAGCGAAAAATACATCGTCACCTGGGACATGTTGCAGATCCATGCACGTAAACTC  3705679

Query  184      GCAAGCCGACTGATGCCTTCTGAACAATGGAAAGGCATTATTGCCGTAAGCCGTGGCGGT  243
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3705678  GCAAGCCGACTGATGCCTTCTGAACAATGGAAAGGCATTATTGCCGTAAGCCGTGGCGGT  3705619

Query  244      CTGGTACCGGGTGCGTTACTGGCGCGTGAACTGGGTATTCGTCATGTCGATACCGTTTGT  303
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3705618  CTGGTACCGGGTGCGTTACTGGCGCGTGAACTGGGTATTCGTCATGTCGATACCGTTTGT  3705559

Query  304      ATTTCCAGCTACGATCACGACAACCAGCGCGAGCTTAAAGTGCTGAAACGCGCAGAAGGC  363
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3705558  ATTTCCAGCTACGATCACGACAACCAGCGCGAGCTTAAAGTGCTGAAACGCGCAGAAGGC  3705499

Query  364      GATGGCGAAGGCTTCATCGTTATTGATGACCTGGTGGATACCGGTGGTACTGCGGTTGCG  423
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3705498  GATGGCGAAGGCTTCATCGTTATTGATGACCTGGTGGATACCGGTGGTACTGCGGTTGCG  3705439

Query  424      ATTCGTGAAATGTATCCAAAAGCGCACTTTGTCACCATCTTCGCAAAACCGGCTGGTCGT  483
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3705438  ATTCGTGAAATGTATCCAAAAGCGCACTTTGTCACCATCTTCGCAAAACCGGCTGGTCGT  3705379

Query  484      CCGCTGGTTGATGACTATGTTGTTGATATCCCGCAAGATACCTGGATCGAACAGCCGTGG  543
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3705378  CCGCTGGTTGATGACTATGTTGTTGATATCCCGCAAGATACCTGGATCGAACAGCCGTGG  3705319

Query  544      GATATGGGCGTCGTATTCGTCCCGCCAATCTCCGGTCGCTAAT  586
                |||||||||||||||||||||||||||||||||||||||||||
Sbjct  3705318  GATATGGGCGTCGTATTCGTCCCGCCAATCTCCGGTCGCTAAT  3705276
```

9.2   Appendix 2 - SDS-PAGE analysis

The purified protein involved in NADP$^+$ and dyhdrofolate biosynthesis were analysed by SDS-PAGE to assess purity, and confirm the correct monomeric weight.
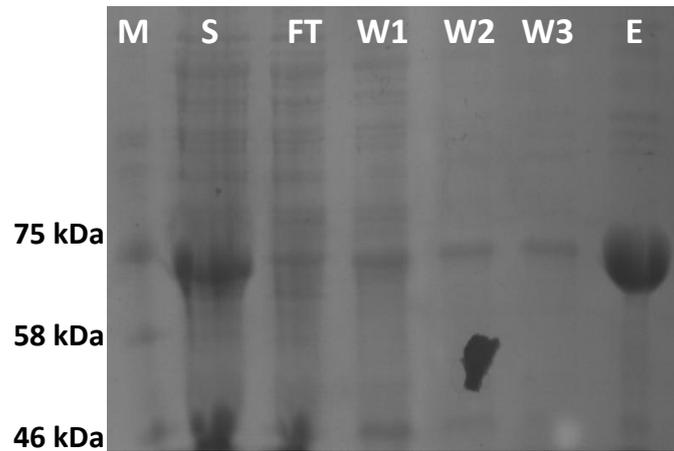
**Figure 9.1 SDS-PAGE analysis of aspartate oxidase following Ni2+-affinity purication. M = marker, S = supernatant, FT = flow through, W1-3 = washes and E = elution.**
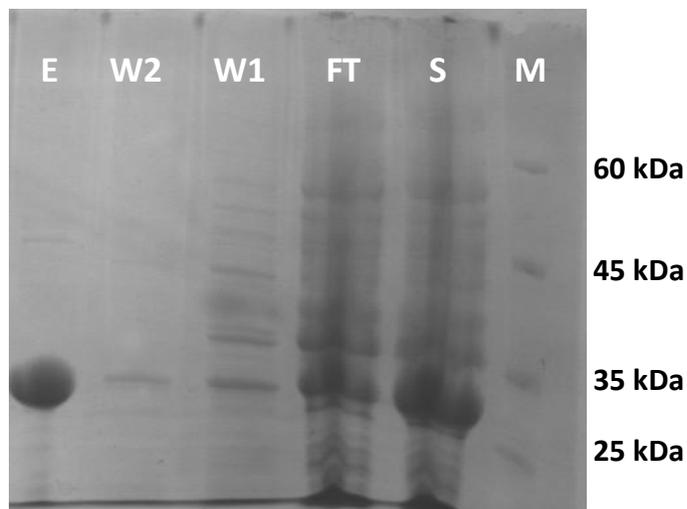

Figure 9.2 SDS-PAGE analysis of QAPRT following purification via Ni2+-affinity chromatography. E = elution, W2 - wash 2, W1 = wash 1, FT = flow through, S = supernatant and M = marker.
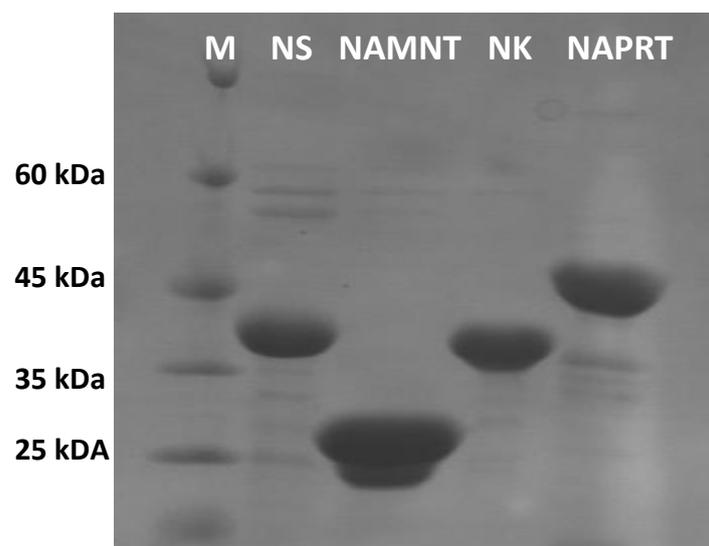
Figure 9.3 SDS-PAGE analysis of NADS, NAMNT, NADK and NAPRT following Ni2+-affinity purification. M = marker, NS = NADS and NK = NADK
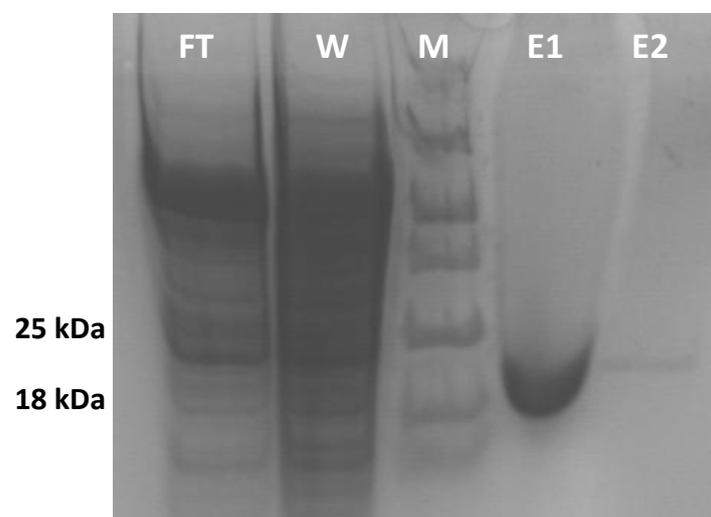

Figure 9.4 SDS-PAGE analysis of HPPK purifiaction via Ni^{2+}-affinity chromatography. FT = flow through, W = wash, M = marker, E1 = elution 1 and E2 = elution 2.
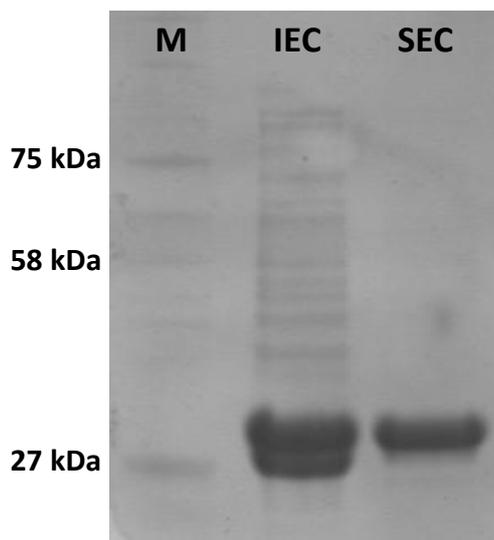
Figure 9.5 SDS-PAGE of DHPS fractions from ion-exchange and size-exclusion chromatography. M = marker, IEC = ion-exchange column, SEC = size-exclusion column.
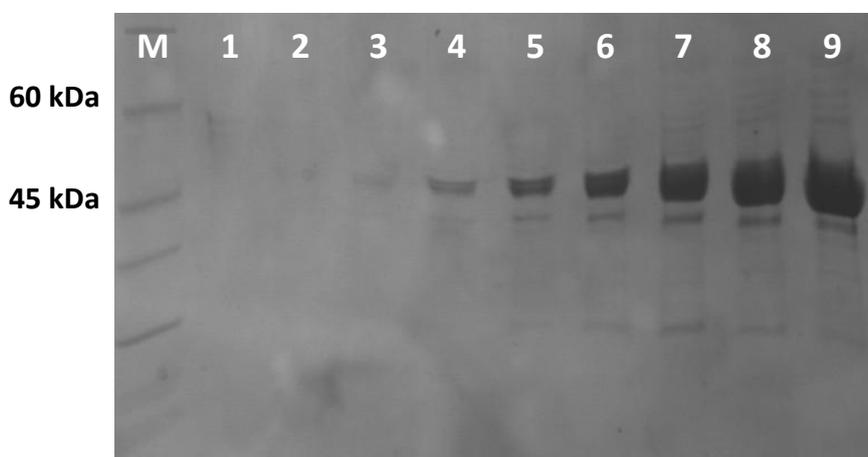

Figure 9.6 SDS-PAGE of DHFS containing fractions from ion-exchange chromatography. M = marker, 1-9 = fractions.

9.3 Appendix 3 - Stopped-flow spectroscopy data analysis

Double exponential fits were applied to the decrease in signal observed for each reaction trace, using the software provided on the stopped flow instrument (Applied Photophysics Workstation). This provided a rate constant for the reaction at each temperature. Each rate constant was an average of at least 3 individual traces, and this was repeated at least 3 independent times for each temperature. Thus, reported rate constants are an average of at least 9 individual traces, with standard deviation of the 3 independent runs used to calculate the error for each reported rate constant. Examples of individual traces (red) and the corresponding double exponential fits (black) for light and heavy BsDHFR and ^{13}C-labelled NADPH and NADPD are shown below.
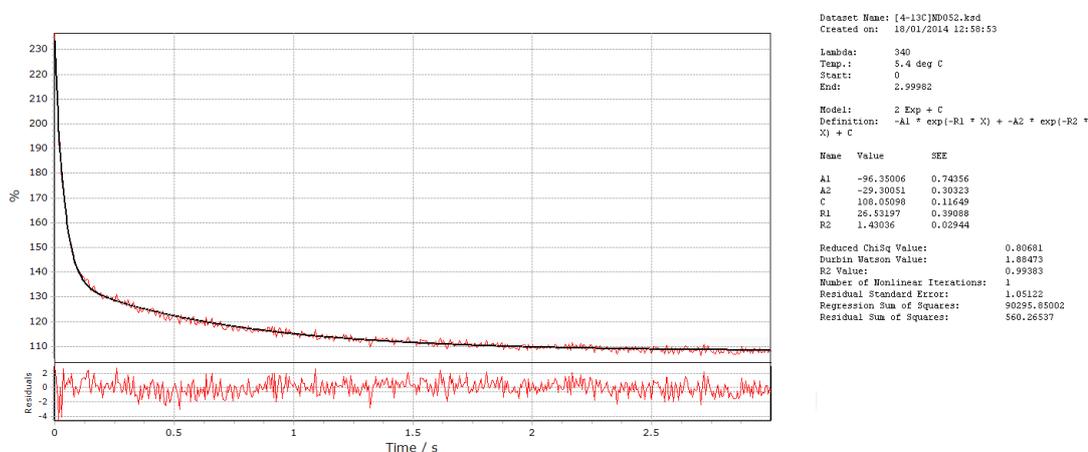

Figure 9.7 Reaction trace from the pre-steady state reaction of EcDHFR with [4-^{13}C]-NADPD at 5 °C.
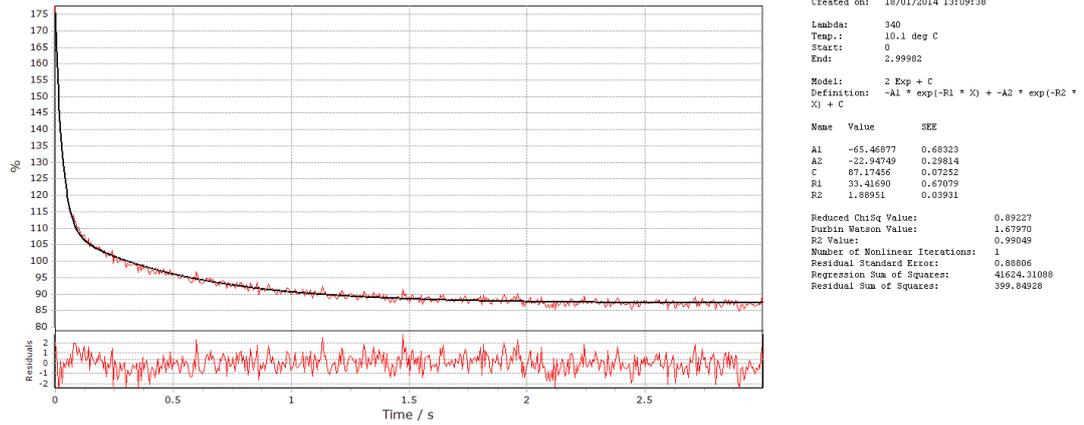
Figure 9.8 Reaction trace from the pre-steady state reaction of EcDHFR with [4-13C]-NADPD at 10 °C.
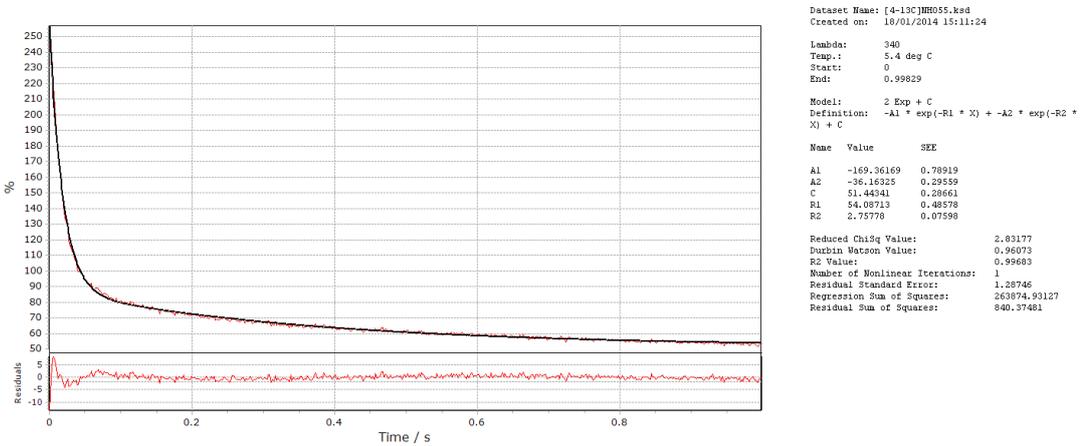

Figure 9.9 Reaction trace from the pre-steady state reaction of EcDHFR with [4-^{13}C]-NADPH at 5 °C.
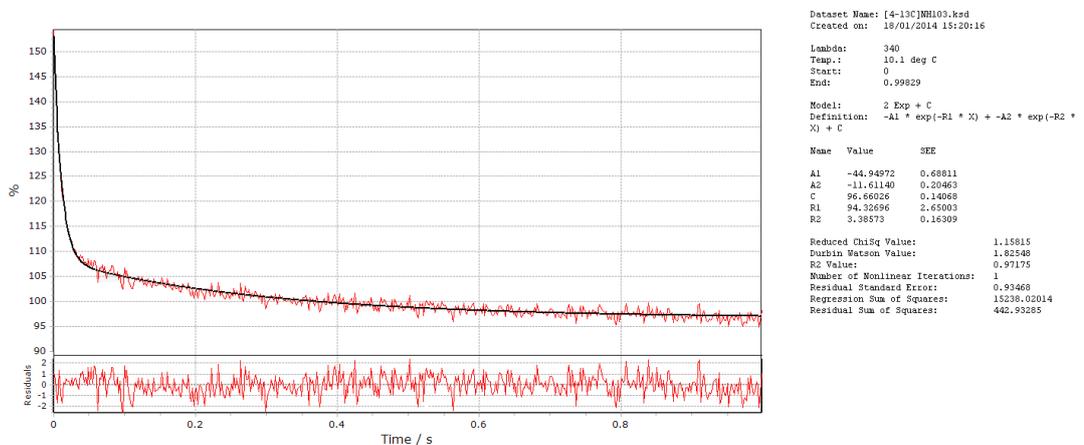
Figure 9.10 Reaction trace from the pre-steady state reaction of EcDHFR with [4-^{13}C]-NADPH at 5 °C.

Dataset Name: 13c15nNH254_avg.ksd
Created on: 16/08/2014 14:54:49

Lambda: 340
Temp.:   24.7 deg C
Start: 0
End: 0.79961

Model: 2 Exp + C
Definition: -A1 * exp(-R1 * X) + -A2 * exp(-R2 * X) + C

Name   Value   SEE

A1   -0.03524   0.00022
A2   -0.02095   0.00010
C   0.39430   0.00003
R1   117.65909   1.41940
R2   6.72625   0.05118

Reduced ChiSq Value: 1.00352
Durbin Watson Value: 0.31837
R2 Value: 0.99818
Number of Nonlinear Iterations: 1
Residual Standard Error: 0.00029
Regression Sum of Squares: 0.02398
Residual Sum of Squares: 0.00004

Figure 9.11 Reaction trace from the pre-steady state reaction of $^{13}C^{15}N^{2}H$-labelled BsDHFR with NADPH at 25 °C.

Dataset Name: NH251.ksd
Created on: 16/08/2014 12:19:37

Lambda: 340
Temp.:   24.7 deg C
Start: 0.00156
End: 0.79961

Model: 2 Exp + C
Definition: -A1 * exp(-R1 * X) + -A2 * exp(-R2 * X) + C

Name   Value   SEE

A1   -0.02912   0.00046
A2   -0.03195   0.00009
C   0.32124   0.00005
R1   141.40765   3.33740
R2   4.89497   0.03236

Reduced ChiSq Value: 3.45129
Durbin Watson Value: 0.41521
R2 Value: 0.99816
Number of Nonlinear Iterations: 3
Residual Standard Error: 0.00038
Regression Sum of Squares: 0.04050
Residual Sum of Squares: 0.00007

Figure 9.12 Reaction trace from the pre-steady state reaction of BsDHFR with NADPH at 25 °C.