
Structure and Function of Enteric Pathogen Glyceraldehyde-3-Phosphate Dehydrogenases

Thesis submitted for the degree of
Doctor of Philosophy
The University of Leicester

by

Paul Ronald Elliott BSc (Nottingham)
Department of Biochemistry
University of Leicester

September 2008

**Structure and Function of Enteric pathogen Glyceraldehyde-3-Phosphate
Dehydrogenases**
Paul Ronald Elliott

Abstract

The availability of published genomes from all domains of life has provided insight into biochemical processes for many organisms. Frequently the mapping of classical pathways onto genome-derived data is used to deduce metabolic pathways in an otherwise uncharacterised system. Whilst this method may be sufficient as a prelude to further biochemical analysis, the function of genes may be assigned by extrapolation from homologs, and this may not be correct. This study highlights the dangers of such a process, focusing on the glycolytic/gluconeogenic enzymes glyceraldehyde 3-phosphate dehydrogenase (GAPDH) from the human pathogenic species *Helicobacter pylori* and *Campylobacter jejuni*. *H. pylori* has two genes encoding GAPDH (*gapA* and *gapB*). These are both annotated as NAD⁺-dependent glyceraldehyde 3-phosphate dehydrogenases. This study has demonstrated enzymatically and structurally that *gapA* encodes a NADP-dependent GAPDH, whilst *gapB* encodes a NAD⁺-dependent GAPDH, furthermore GAPDHB is a better phosphorylating erythrose-4-phosphate dehydrogenase. Structural analysis of GAPDHA and GAPDHB showed key residues providing specificity for the coenzyme NADP⁺ over NAD⁺ and this finding was used to search for other putative NADP⁺-dependent GAPDHs within the *Campylobacteriales* order. Other NADP⁺-dependent GAPDHs were identified; including that of *C. jejuni*, which has only one annotated GAPDH-encoding gene. Structural and enzymatic analysis confirmed *C. jejuni's* GAPDH is NADP⁺ dependent, although dual specificity is observed. This further shows the importance of experimental data to describe a system. Finally, a mutagenic approach was undertaken to determine the mechanism underlying the differing substrate specificities between GAPDHA and GAPDHB. Whilst the structural analysis was unable to provide a determinant of substrate specificity, these structures provided clear evidence for a reaction mechanism used by all phosphorylating GAPDHs. The significance of the findings is discussed in the context of the metabolism of these pathogens. This work demonstrates the importance of the synergy between structural and genomic analysis.

Acknowledgements

As always in science, one's research is not carried out in isolation, through either collaboration or financial and moral support. Therefore, I would like to thank the following people for their varying help and support throughout this study.

Several third year project students who worked in the laboratory under my supervision provided various levels of input. Work on GAPDHA was assisted by Shabaz Mohammad who assisted in His₆-GAPDHA cloning, expression, purification and crystallisation. Helen Melrose worked on generating the mutants for *H. pylori* and some preliminary crystallisation experiments. Finally, Louise Edgell assisted in the cloning, expression, enzymatic analysis and crystallisation of cGAPDH. Louise additionally assisted in the refinement of holo-cGAPDH.

Mrs Jacquie Greenwood provided excellent technical assistance in the laboratory and was involved with Daniel Evans in the purification and crystallisation of apo-GAPDH.

I would like to thank various other people within the Biochemistry Department for their useful contribution and assistance with other projects which I undertook during my thesis. In particular, Dr Alex Gingras, Dr Ben Goult, Professor David Critchley, Dr Martin Dickens and Dr Igor Barsukov for their assistance and advice.

I greatly acknowledge Dr Peter Moody for his continued support throughout my thesis and his commitment to ensuring my career development. More importantly his willingness to allow me to pursue my research goals, unhindered and un-phased.

I would like to thank my parents Ronald and Christine, my brothers Philip and Christopher and my sister Jennifer for their continued support, understanding and patience.

I am indebted to the generosity of Dominique, Michael, Emily and Tim Shuttlewood for their hospitably and kindness throughout 2008 and 2009 and making me feel part of the family.

I would like to thank Chloe for her understanding and support during the thesis.

Finally, my gratitude to the Medical Research Council for providing the funding for this work.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
Abbreviations.....	viii
Chapter 1. Introduction The <i>Campylobacterales</i> order.....	1
1.1 Introduction.....	2
1.2 The Epsilon proteobacteria class.....	2
1.3 The Campylobacterales order.....	4
1.3.1 <i>Wolinella succinogenes</i>	5
1.3.2 <i>Campylobacter jejuni</i>	5
1.4 Genetic variability within the Campylobacterales.....	6
1.5 <i>Helicobacter pylori</i>	7
1.5.1 Morphology.....	7
1.5.2 Growth Requirements.....	8
1.5.3 Taxonomy.....	9
1.5.4 Epidemiology.....	10
1.5.5 Transmission.....	10
1.5.6 Disease Spectrum.....	10
1.5.6.1 Gastritis.....	11
1.5.6.2 Gastroduodenal Ulceration.....	12
1.5.6.3 Gastric Cancer.....	12
1.5.6.4 MALT Lymphoma.....	12
1.5.7 Diagnosis.....	13
1.5.7.1 Invasive.....	13
1.5.7.1.1 Endoscopy.....	13
1.5.7.1.2 Histology.....	13
1.5.7.1.3 Culture biopsy.....	14
1.5.7.1.4 Rapid Urease test.....	14
1.5.7.1.5 Molecular tests.....	14
1.5.7.2 Non-Invasive.....	15
1.5.7.2.1 Serology.....	15
1.5.7.2.2 Urea breath test.....	15
1.5.8 Treatment.....	15
1.5.9 Virulence factors.....	16
1.5.9.1 Vacuolating Cytotoxin.....	18
1.5.9.2 Urease.....	18
1.5.9.3 Cytotoxin-Associated Gene.....	19
1.5.9.4 Neutrophil Activating Protein.....	20
1.5.9.5 Lipopolysaccharide.....	20
1.6 Protection against Oxidative Stress in <i>H. pylori</i>	21
1.7 Metabolism within the Campylobacterales.....	21
1.7.1 Carbohydrate Metabolism.....	22
1.8 Aldehyde Dehydrogenase Super family.....	26
1.8.1 Glyceraldehyde-3-phosphate Dehydrogenase.....	27
1.8.1.1 Glycolytic GAPDH.....	28
1.8.1.1.1 Structure of GAPDH.....	28
1.8.1.1.2 Reaction Mechanism.....	31
1.8.1.2 Photosynthetic GAPDH.....	34
1.8.1.3 Non-Phosphorylating GAPDH.....	36
1.8.2 Pyruvate metabolism.....	36
1.8.3 The Krebs cycle.....	37
1.8.4 Respiratory chains.....	39
1.8.5 Amino acid metabolism.....	40

1.9	Focus of current study	41
Chapter 2.	Materials and Methods	42
2.1.	Materials	43
2.1.1.	Chemicals and Reagents.....	43
2.1.2.	Chromatographic media and Membranes.....	43
2.1.3.	Crystallisation accessories	43
2.2.	Molecular Biology Methods	44
2.2.1.	Primer design.....	44
2.2.2.	GAPDH ORF amplification by PCR	45
2.2.3.	TOPO isomerase cloning.....	46
2.2.4.	Transformation of <i>E. coli</i> strains	47
2.2.5.	Colony PCR screening.....	47
2.2.6.	Site-directed mutagenesis	48
2.3.	Protein Methods	50
2.3.1.	SDS Polyacrylamide Gel Electrophoresis	50
2.4.	Crystallographic methods	50
2.4.1.	Protein Crystallisation.....	51
2.4.1.1.	Robotics in Protein crystallography.....	53
2.4.2.	Description of X-rays	53
2.4.3.	The Fourier Transform	54
2.4.4.	Cryo cooling	55
2.4.5.	Diffraction data collection and processing	55
2.4.6.	Scattering of X-rays by an atom	56
2.4.7.	Scattering of X-rays by a crystal.....	58
2.4.8.	The electron density equation	59
2.4.9.	The Phase problem.....	59
2.4.10.	Molecular Replacement	60
2.4.11.	Model building.....	61
2.4.12.	Refinement.....	61
2.4.13.	Structure validation	62
2.4.13.1.	Omit maps	62
2.5.	Enzymatic analysis	63
2.6.	Methods for the analysis of GAPDHA.....	64
2.6.1.	Cloning and over expression.....	64
2.6.2.	GAPDHA purification	65
2.7.	Methods for the analysis of GAPDHB.....	66
2.7.1.	Cloning and protein over expression	66
2.7.2.	GAPDHB purification	66
2.8.	Methods for analysing <i>H. pylori</i> mutants	67
2.8.1.	Mutagenesis.....	67
2.8.2.	Protein over expression and purification	67
2.9.	Methods for analysis of <i>C. jejuni</i> GAPDH	68
2.9.1.	Cloning and protein over expression	68
2.9.2.	Protein purification	69
Chapter 3	70
Structures of Glyceraldehyde-3-phosphate Dehydrogenase A	70
3.1	Introduction.....	71
3.2	GAPDHA cloning and expression.....	72
3.3	Enzymatic analysis	73
3.4	Crystallisation of NADP ⁺ -bound GAPDHA.....	76
3.5	Data collection of His ₆ -GAPDHA	78
3.6	Data processing.....	79
3.7	Structure determination of GAPDHA	80

3.7.1	Apo-GAPDHA	80
3.7.2	Holo-GAPDHA	82
3.8	Analysis of apo- and holo-GAPDHA	84
3.8.1	Apo-GAPDHA	84
3.8.2	Holo-GAPDHA	86
3.9	Structural analysis	88
3.9.1	Holo-GAPDHA	88
3.9.2	Crystal Packing	89
3.9.3	NADP ⁺ binding	91
3.10	Determinants of NADP ⁺ -binding	94
3.10.1	Comparison with other NADP-dependent GAPDHs	94
3.10.2	Comparison with NADP ⁺ binding in other Rossmann folds	99
3.10.3	Effects of NADP ⁺ binding	100
3.11	Implications for <i>H. pylori</i> carbohydrate metabolism	102
3.12	NADP ⁺ -dependent GAPDH's within the <i>Campylobacterales</i> order	103
3.13	Conclusion	105
Chapter 4.		106
Glyceraldehyde-3-phosphate Dehydrogenase from <i>Campylobacter jejuni</i>		106
4.1.	Introduction	107
4.2.	Cloning, over expression and purification of cGAPDH	108
4.3.	Enzymatic analysis of cGAPDH	110
4.4.	Crystallisation of cGAPDH-NADP ⁺	113
4.5.	Data collection and analysis of holo-cGAPDH	115
4.6.	Structure determination of cGAPDH	117
4.7.	Stereochemical analysis of cGAPDH model	119
4.8.	Structural analysis of cGAPDH	120
4.8.1.	Overall Structure	120
4.8.2.	Active site	122
4.8.3.	NADP ⁺ Binding site	124
4.9.	Analysis of holo-cGAPDH and GAPDHA	125
4.10.	Conclusion	127
Chapter 5. Structure of Glyceraldehyde-3-Phosphate Dehydrogenase B		128
5.1.	Introduction	129
5.2.	GAPDHB cloning and expression	130
5.3.	Crystallisation of GAPDHB	131
5.4.	Data collection and processing	132
5.5.	Structure determination of GAPDHB	133
5.6.	Structure analysis of holo-GAPDHB	134
5.6.1.	Stereochemical analysis of GAPDHB	134
5.7.	Overall Structure	136
5.8.	N-terminal hexa-histidine linker	137
5.9.	Unit cell analysis	138
5.10.	NAD ⁺ binding to GAPDHB	140
5.11.	Enzymatic analysis of GAPDHB	143
5.12.	Conclusion	146
Chapter 6		147
Ternary Complexes of <i>H. pylori</i> GAPDHA and GAPDHB mutants		147
6.1	Introduction	148
6.2	Site-directed Mutagenesis	150
6.3	Protein expression and purification	150
6.4	Crystallisation of GAPDHA and GAPDHB mutants	151
6.4.1	GAPDHA c149s	151
6.4.2	GAPDHB c151a and c151s	152

6.5	Data collection for GAPDH mutants.....	154
6.5.1	GAPDHA c149s.....	154
6.5.2	GAPDHB c151s.....	157
6.5.3	GAPDHB c151a.....	158
6.6	Data processing.....	159
6.6.1	GAPDHA c149s.....	159
6.6.2	GAPDHB mutants.....	160
6.7	Structure Determination.....	162
6.7.1	GAPDHA c149s.....	162
6.7.2	GAPDHB mutants.....	163
6.8	Stereochemical analysis.....	165
6.9	Structural analysis of mutant complexes.....	167
6.9.1	GAPDHA c149s + e4p.....	167
6.9.2	GAPDHB c151s + g3p.....	169
6.9.3	GAPDHB c151s + e4p.....	173
6.9.4	GAPDHB c151a + g3p.....	177
6.9.5	GAPDHB c151a + e4p.....	179
6.10	Conclusions.....	182
6.10.1	GAPDHB substrate specificity.....	182
6.10.2	Analysis with other GAPDH complexes.....	183
6.10.3	<i>H. pylori</i> GAPDH and the Flip-Flop reaction mechanism.....	186
Chapter 7	Discussion.....	187
7.1	Introduction.....	188
7.2	NADP ⁺ -dependent GAPDHs.....	188
7.2.1	Cooperativity.....	188
7.2.2	Determinants of coenzyme specificity.....	189
7.3	Substrate specificity.....	190
7.4	Analysis of GAPDH reaction mechanism.....	192
7.5	Metabolic inferences within the <i>Campylobacterales</i> order.....	193
7.6	Design of drug targets for GAPDHs.....	194
	Publications during thesis.....	195
	Bibliography.....	203

Abbreviations

Å	Angstrom (1 Å = 0.1 nm)
α, β, γ	crystallographic cell angles
a, b, c	crystallographic cell vectors
DTT	dithiothreitol
°	degree
e4p	erythrose-4-phosphate
EDTA	ethylenediamine-tetraacetic acid
$\rho(x, y, z)$	electron density
EtOH	ethanol
f	atomic scattering factor
$F(S)$	atomic structure factor
g3p	glyceraldehydes-3-phosphate
GAPDH	Glyceraldehyde-3-phosphate Dehydrogenase
IPTG	isopropyl β -D-1-thiogalactopyranoside
K	kelvin
K_M	Michaelis constant
k_{cat}	turn over number (overall rate constant)
k	wave vector
LB	Luria-Bertani
MeOH	methanol
MES	2-(N-Morpholino)ethanesulphonic acid
M	molarity
m	meter
m, μ , n	milli, micro, nano
MPD	2-methyl-2,4-pentanediol
Mwt	molecular weight
NAD(P) ⁺	β -nicotinamide adenine dinucleotide (phosphate)
NAD(P)H	β -nicotinamide adenine dinucleotide (phosphate) reduced form
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
ϕ, ψ	phi, psi (peptide torsional angles)
R	crystallographic R-factor
R_{free}	crystallographic R-factor calculated with 5 % of observations used
R_{merg}	merging R-factor
SDS	sodium dodecylsulphate
TAE	Tris acetate EDTA
Tris	Tris (hydroxymethyl)aminomethane
TEMED	N, N, N', N',-tetramethylethylenediamine

Chapter 1 Introduction
The Campylobacterales order of the epsilon proteobacteria class of bacteria

1.1 Introduction

Metabolic pathways have been well characterised for the "classical" systems, for example, mammals and *E. coli*. Often, these processes are mapped onto other species, particularly when genomic information is available. The effect of this is incorrectly annotated databases, which tend to serve as starting points in the design of new experiments. This study highlights one such case with the enteric pathogens *Helicobacter pylori* and *Campylobacter jejuni*. Both species are found within the epsilon class of proteobacteria phylum, a poorly characterised class, composed of species that are pathogenic to humans and hence of significant importance.

1.2 The Epsilon proteobacteria class

The proteobacteria phylum is composed of five classes, one of these classes is the epsilon proteobacteria. The epsilon proteobacteria colonise a diverse spectrum of ecological niches, ranging from oceanic hydrothermal vents (Corre *et al.*, 2001), oil-fields (Kodama and Watanabe, 2004) and sulphidic cave springs (Engel *et al.*, 2003), to the gastrointestinal tracts of animals (Wolin *et al.*, 1961, Marshall and Warren, 1983). Within this class are two main orders; the *Campylobacterales* order, which in general, comprises of families whose species are host-associated and the *Nautiliales* order, which consists of free-living organisms found in the diverse ecological habitats. Figure 1.1 shows a schematic representation of the relationship between the families that comprise the two orders of the epsilon proteobacteria class. For simplicity, only the species of bacteria that were involved in this study are shown.

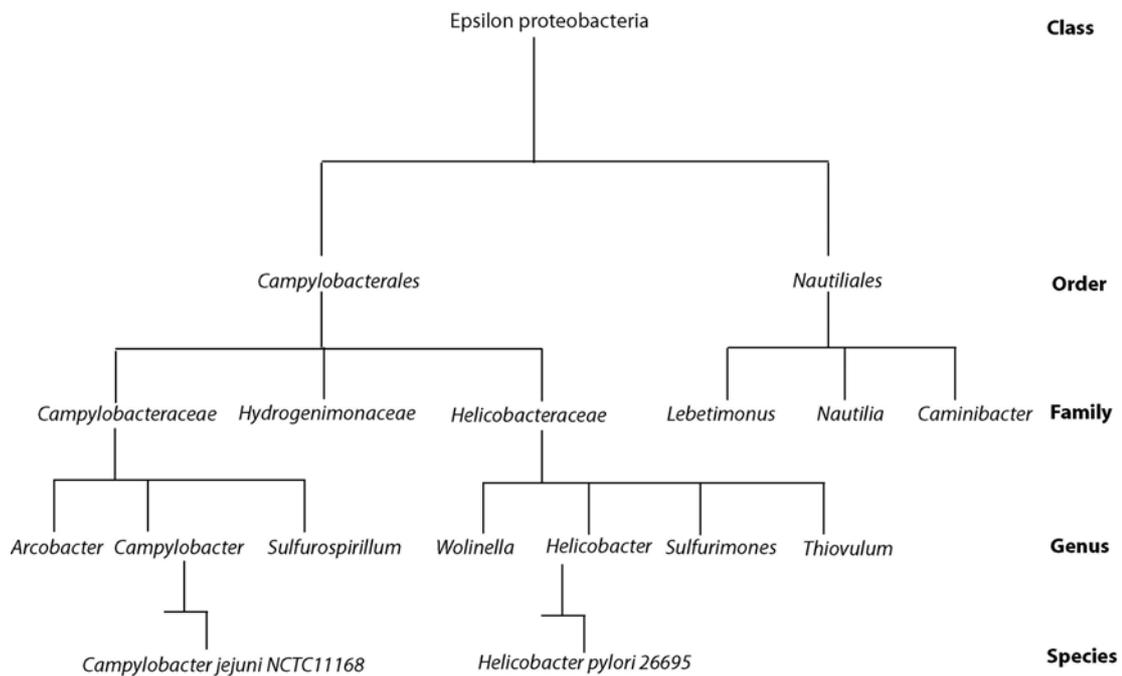


Figure 1. 1 A schematic representation of the relationship between the two orders of the epsilon proteobacteria. For clarity only the genera of the *Campylobacteraceae* and *Helicobacteraceae* families are shown, as are the two species *Campylobacter jejuni* and *Helicobacter pylori* used in this study.

All species belonging to the *Campylobacterales* and *Nautiliales* orders are difficult to cultivate and study *in-vivo*. However, there is a greater interest with the *Campylobacterales* order due to their dependence on host-interactions and the diseases associated through these interactions (reviewed in Chapter 1.5.6). This bias towards research within the *Campylobacterales* is reflected by the available completed genome sequences of the epsilon proteobacteria, with all seven completed genome sequences coming from the *Campylobacterales* order. The completely sequenced genomes include: *Campylobacter jejuni* NCTC11168 (Parkhill *et al.*, 2000) and *Helicobacter pylori* 26695 (Tomb *et al.*, 1997) (both in this study), *H. pylori* J99 (Alm *et al.*, 1999), *H. hepaticus* ATCC51449 (Suerbaum *et al.*, 2003), *C. jejuni* RM1221 (Fouts *et al.*, 2005), *Wolinella succinogenes* DSM1740 (Barr *et al.*, 2003) and *Thiomicrospira denitrificans* ATCC33889 (Copeland *et al.*, 2006).

In general, the *Nautiliales* order are environmental bacteria which play important roles in the carbon, nitrogen and sulphur cycles, reviewed in (Campbell *et al.*, 2006). The *Nautiliales* order is made up of species that are chemolithotrophic, in that they are capable of utilising basic carbon containing molecules such as carbon monoxide and carbonates as the main carbon source for biosynthesis. In addition, they also utilise the oxidation of inorganic and organic compounds to derive the energy for these biosynthetic pathways. Other organisms within the order are chemoorganotrophic in that they are able to utilise organic compounds to generate the biosynthetic energy for carbon fixation. The metabolism of these bacteria is significantly different that that of the *Campylobacterales* order and is not discussed further within this study.

1.3 The Campylobacterales order

The *Campylobacterales* order has been intensely investigated due to three of the genera notably; *Campylobacter*, *Helicobacter* and *Wolinella* containing species that may be regarded as living within a hostile environment, but rather co-inhabiting within the host organism. In the case of *Campylobacter* and *Wolinella* genera, a commensal interaction occurs, where the avian and cattle, respectively do not develop symptoms of infection. Although *Campylobacter jejuni* is one of the main causes of bacterial food-borne illness in humans, causing acute gastritis, which can progress to Guillain-Barré syndrome (Nachamkin *et al.*, 1998) (Chapter 1.3.2). *Helicobacter pylori* however, can persist in the human gastrointestinal tract asymptotically but can also bring about gastritis and has been identified as a key factor for gastric cancer (Cover and Blaser, 1996) (Chapter 1.5.6.3).

The focus of this study was on the glycolytic/gluconeogenic genes encoding for Glyceraldehyde-3-phosphate Dehydrogenase (GAPDH) from *Helicobacter pylori* 26695 and *Campylobacter jejuni* NCTC1168. The vast majority of this work focused particularly on the two GAPDH genes from *H. pylori*, termed *gapA* (gi: 6626253) (NCBI, NIH) and *gapB* (gi: 2494645) (NCBI, NIH). For completeness the *Wolinella* and the *Campylobacter* genera will briefly be discussed to highlight their similarities (and differences) to the *Helicobacter* genus.

1.3.1 *Wolinella succinogenes*

Wolinella succinogenes was first isolated from the rumen of cattle (Wolin *et al.*, 1961) and has been shown to be zoonotic, in that its infection does not harm the host and is not believed to be involved in infection of humans or other animals, reviewed in (Eppinger *et al.*, 2004). Its physiology is similar to *Helicobacter* and *Campylobacter* species concerning its growth in non-strictly anaerobic conditions, i.e. the presence of 2% oxygen. However, *W. succinogenes* has a greater number of complete metabolic pathways compared to *Helicobacter* and *Campylobacter* species. Genomic analysis of *W. succinogenes* reveals glucose and hexose transporters are absent, as well as key enzymes from the Entner-Doudoroff pathway and the oxidative pentose-phosphate pathway (Baar *et al.*, 2003a). This genomic analysis suggests a mainly gluconeogenic role of the organism for generation of key five carbon sugars required for nucleotide biosynthesis (Baar *et al.*, 2003a).

1.3.2 *Campylobacter jejuni*

Campylobacter jejuni is a commensal organism of chickens and other avian species, where the bird does not develop any symptoms or ill effects due to infection. However, once in humans, probably as a result of ingested, contaminated chicken (although transmission to humans can occur through unpasteurised milk or water referred in (Allos, 2001). The bacteria infect the gastrointestinal tract. Infection in humans results in diarrhoea, commonly accompanied by abdominal cramps and fever. In addition to the chronic gastroenteritis, *C. jejuni* infection can also result in Guillain-Barré Syndrome (GBS), an immune-mediated disorder affecting the peripheral nervous system, reviewed in (Yuki, 2001). GBS is characterised by loss of tendon reflexes, weakness within the limbs and autonomic dysfunctions and is self-limiting, with partial or full-recovery occurring within weeks to months (Yuki, 2001). It is believed that molecular mimicry associated with *C. jejuni* results in GBS, whereby antibodies against a lipopolysaccharide on the cell wall of *C. jejuni* recognise the GM1 ganglioside of the peripheral nerve axons (Yuki *et al.*, 1990, Walsh *et al.*, 1991). With an estimated two and a half million cases of *C. jejuni* annually within the United States, *C. jejuni* is regarded as a serious health and economic problem (Samuel *et al.*, 2004).

1.4 Genetic variability within the Campylobacterales

The availability of the complete genome sequences from three genera of the *Campylobacterales* order, has provided evidence for features specific to the epsilon proteobacteria, reviewed in (Gupta, 2006). Of interest it should be noted that for pathogenic organisms which are adapted to living within the host, their genomes undergo a process of reductive evolution (Andersson and Andersson, 1999, Oshima, Oshima et al., 2004). The process of reductive evolution results in a reduction of the genome size with cellular processes that are no longer required becoming redundant. Within the *Campylobacterales*, *W. succinogenes* has the largest genome of 2.11 Mbp. This and the 16S RNA phylogenic relationship of *W. succinogenes* to the other genera within the order suggests that it is a common ancestor and therefore has retained the gene pool, which the other genera have lost (Eppinger *et al.*, 2004). The genome of *H. pylori* appears to be still going through a process of reductive evolution; therefore, it must contain genes, which are no longer required because of duplication and evolution (Eppinger *et al.*, 2004). It has been shown that strain variation between patients infected with *H. pylori* is highly polymorphic with frequent horizontal gene transfer between strains (Suerbaum *et al.*, 1998). Furthermore, strains undergo genetic variation during infection, hence are continuously evolving within the gut to adjust to the environment as a life-long pathogen (Israel *et al.*, 2001).

1.5 *Helicobacter pylori*

The presence of bacteria in the human stomach had been known before the successful isolation of a spiral bacterial species from the human gut, later known as *Helicobacter pylori*, by Barry Marshall and Robin Warren (Marshall and Warren, 1983). Self-ingestion experiments by Marshall (Marshall *et al.*, 1985) and Morris (Morris and Nicholson, 1987) (for which they were awarded the Nobel Prize in Physiology and Medicine in 2005) demonstrated that *H. pylori* can colonise the human stomach and cause inflammation of the gut epithelium. In the case of Marshall a persistent case of gastritis developed which was only treated after sequential treatment with first doxycycline and then bismuth subsalicylate (Kusters *et al.*, 2006).

The following sections provide an overview of the suggested mechanisms underlying *H. pylori's* contribution to disease and an overview of its metabolism.

1.5.1 Morphology

Individual gram-negative bacteria measure between 2 to 4 μm in length and 0.5 to 1 μm in width. The bacteria are sometimes rod-shaped (Figure 1.2A), coccoid shaped bacteria have been observed, as in Figure 1.2B, though these are thought to be dead cells as a result from antibiotic treatment or prolonged culturing (Kusters *et al.*, 1997). However, the common morphology is spiral-shaped (Figure 1.2C). Two to six flagella of 3 μm in length confer motility and have been shown to be required for infection (Eaton *et al.*, 1992).

1.5.2 Growth Requirements

H. pylori is a microaerophile, it requires 2 to 5% oxygen as *H. pylori* uses oxygen as a terminal electron acceptor (Chapter 1.7.4). Although *H. pylori*'s natural habitat is the gastric mucosa, which is acidic, it is actually a neutrophile, with optimum growth occurring at neutral pHs (Scott *et al.*, 2002, Stingl *et al.*, 2002, Bauerfeind *et al.*, 1997). Several studies have shown that a complex medium is required to cultivate *H. pylori in vitro*, the most commonly used is Stuart's Transport medium (Ashdown, 1978) this contains a majority of amino acids required to biosynthetic processes discussed in Chapters 1.6 and 1.7.2.

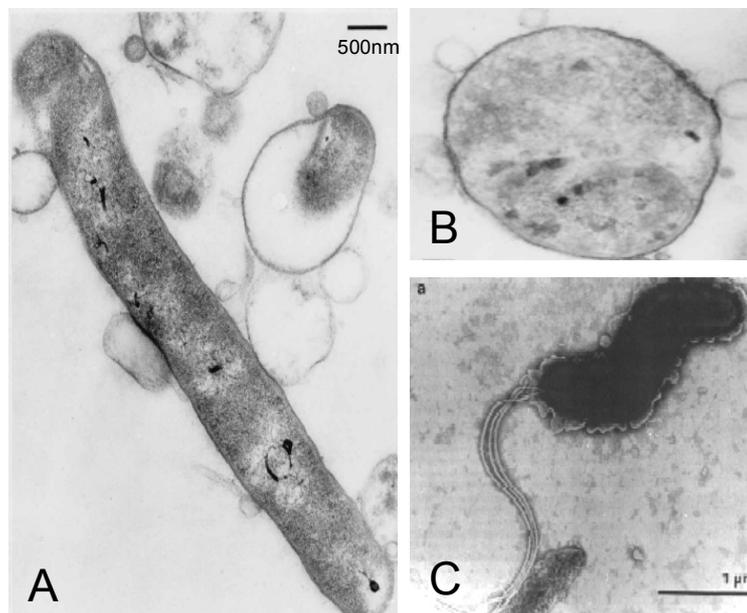


Figure 1. 2 Several morphologies of *H. pylori* have been observed; A. rod-shaped bacterium, B. coccoidal after prolonged culturing or antibiotic treatment and C. the common spiral-shaped bacterium showing several flagella. Panels A and B are transmission electron micrographs from (Kusters *et al.*, 1997) and panel C an electron micrograph from (Geis *et al.*, 1989)

1.5.3 Taxonomy

Intensive research into the *Helicobacter* genus has resulted in over 100 different species of *Helicobacter* being identified (National Center for Biotechnology Information). The number of species, which comprise the genus *Helicobacter*, is nearly one-third the total amount of the entire epsilon class, as shown in Figure 1.3. Species of *Helicobacter* are responsible for causing infection of the gastric mucosa of different animals. For example, *H. felis* is found in cats and dogs and the *H. heilmannii* species, which has a wide host range, is found within wild animals and *H. mustelae*, which is found in ferrets. Other *Helicobacter* species colonise the lower gastrointestinal tract and are known as enterohepatic. Enterohepatic *Helicobacter* species colonise the colon and the ileum with the rodent-infecting *H. hepaticus* being well characterised (Rogers and Fox, 2004).

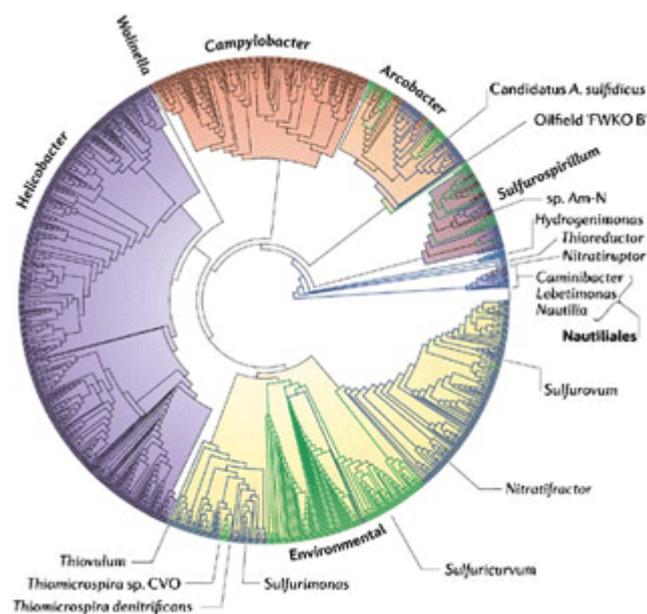


Figure 1. 3 Phylogenetic tree of the epsilon subdivision of the *Proteobacterial* class highlighting the different genera and the number of species making up each genus, taken from (Campbell *et al.*, 2006).

1.5.4 Epidemiology

Within developing countries 80% of the population are found to be infected with *H. pylori* (Perez-Perez *et al.*, 2004), whereas in industrialised countries only 40% of the population test positive for infection and this is mostly reserved to adults and the elderly (Pounder and Ng, 1995). The prevalence of *H. pylori* infection is decreasing in developed countries, whilst it remains constant within developing countries (Magalhaes Queiroz and Lizza, 2006). It is believed that a reduction in childhood infection, as a result of, improved hygiene has led to the decline of infection within developed countries. The notion that infection occurs early on in life and lasts as a life-long infection is supported through the observation that, in developing countries *H. pylori* infection rises rapidly within the first five years of life and remain constant thereafter (Fiedorek *et al.*, 1991). There are higher levels of infection amongst the elderly in developed countries, compared to lower levels of childhood infections. This is believed to be related to higher infection rates several decades ago where socioeconomic status during childhood would have led to increased chances of infection (Parsonnet, 1995).

1.5.5 Transmission

The mechanisms of *H. pylori* transmission are unknown, since *H. pylori* is sensitive to atmospheric oxygen, temperature and requires a complex set of nutrients (Ashdown, 1978). It is likely that transmission occurs from person-to-person via oral-oral or fecal-oral routes, in particular within childhood through contact with close family members (Konno *et al.*, 2005).

1.5.6 Disease Spectrum

Symptoms of *H. pylori* infection range from gastritis to gastroduodenal ulceration and gastric cancer. All patients infected with *H. pylori* have histological gastritis and have a 20% lifetime risk of developing gastric ulcer disease and a 2% risk of developing gastric cancer (Ernst, 2000), (Kuipers, 1999). The progress of disease is summarised in Figure 1.4, taken from (Suerbaum and Josenhans, 2007). The risk of developing these disorders depends on a variety of host, bacterial and environmental factors discussed below. In 1994 a convention organised by the National Institute of Health concluded that *H. pylori* was the major cause of gastric ulcers and recommended that infected individuals be treated to eradicate the organism (Zoorob, 1994).

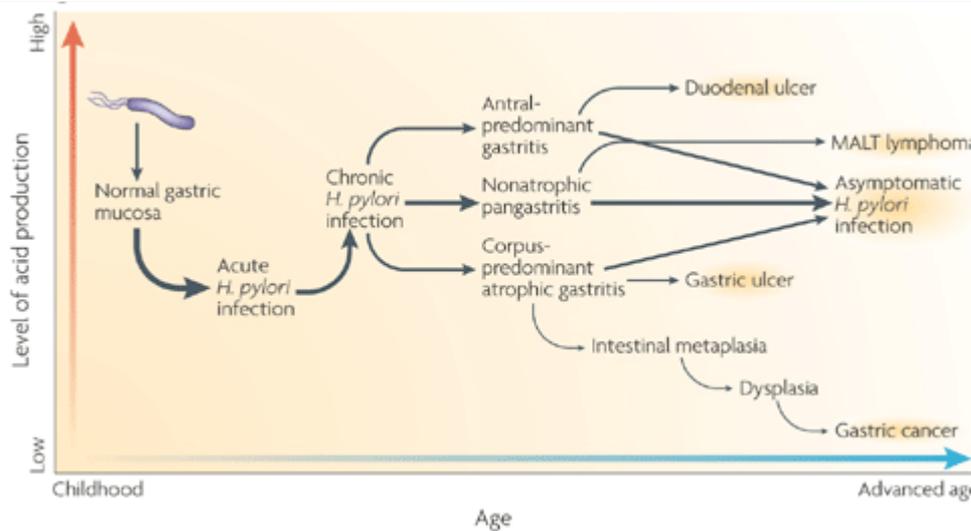


Figure 1. 4 Progression of *H. pylori* infection. Infection typically begins in childhood and is a life-long infection. Acute gastritis tends to develop into chronic gastritis whereby it becomes a life-long infection. Most patients infected with *H. pylori* will remain at the chronic stage of infection. However, 20% of patients will develop gastric or duodenal ulcer disease and a further 2% develop gastric cancer. The disease outcome is strongly dependent upon the level of acid secretion and is discussed in the subsequent sections. Figure taken from (Suerbaum and Josenhans, 2007)

1.5.6.1 Gastritis

H. pylori colonisation of the gastric mucosa leads to gastritis in virtually all infected individuals. There is a clear link between gastritis and *H. pylori*, although Peterson had suggested that *H. pylori* was simply colonising inflamed tissue rather than causing inflammation (Peterson, 1991). Studies however, on children with secondary causes of gastritis i.e. those suffering from Crohn's disease or eosinophilic gastroenteritis (Drumm *et al.*, 1990) have shown that the bacterium is not present in the majority of these cases. This is further supported through studies showing eradication of *H. pylori* results in the relief of symptoms of gastritis in children (Yeung *et al.*, 1990). Other risk factors leading to gastritis include autoimmune diseases such as Crohn's disease and pernicious anemia, as well as chemical damage due to alcohol abuse or nonsteroidal anti-inflammatory drugs reviewed in (Kusters *et al.*, 2006). The level of acid secretion and gastritis are related through the counteractive effects of acid on bacterial growth, versus bacterial growth and associated mucosal inflammation on acid secretion (Kusters *et al.*, 2006). These counteractive effects have an effect on the disease outcome resulting in either duodenal ulceration to gastric cancer and MALT lymphoma. In patients with a functional acid secretion system, *H. pylori* colonises the gastric antrum, resulting in antrum-predominant gastritis. Whereas patients with impaired

acid secretion (by proton-pump inhibitors) have a more even distribution of *H. pylori*, therefore, resulting in corpus-predominant atrophic gastritis.

1.5.6.2 Gastroduodenal Ulceration

It is accepted that between 90-95% of duodenal ulcers and 75% of gastric (peptic) ulcers are as a result of *H. pylori* infection (Covacci *et al.*, 1999). It is more common for children to have duodenal ulcers rather than gastric ulcers, with gastric ulcers being observed in people over the age of 40 (Brown, 2000). Ulceration mostly occurs in regions of high inflammation, (hence the relation between acid secretion and the type of ulceration) (Veldhuyzen van Zanten *et al.*, 1999) as shown in Figure 1.4.

1.5.6.3 Gastric Cancer

H. pylori was designated as a Class I carcinogen by the World Health Organisation in 1994 (International Agency for Research on Cancer, 1994) in light of the geographical associations between the distribution of *H. pylori* infection and incidence of gastric cancer (Forman *et al.*, 1990). The risk of developing gastric cancer is associated with a variety of environmental factors, which affect levels of acid secretion and also strain-specific factors. For example it has been shown that *cagA*-positive strains (*cagA* is described in Chapter 1.5.9.3) increases the risk of gastric cancer (Parsonnet *et al.*, 1997).

1.5.6.4 MALT Lymphoma

All *H. pylori* infected gastric tissue contains B-cell lymphomas, which are not found in un-infected gastric tissue. Eradication of *H. pylori* in infected patients tends to remove B-cell lymphomas (Bayerdorffer *et al.*, 1995). In rare cases the B-cell lymphomas proliferate to form a MALT lymphoma and eradication of *H. pylori* infection after a MALT lymphoma has developed does not cause remission in approximately one third of patients (Terai *et al.*, 2008).

1.5.7 Diagnosis

There are two different ways of testing for *H. pylori* infection, each with their own advantages and disadvantages. Invasive tests result in the taking of gastric specimens for culturing or histology, whilst non-invasive tests utilise peripheral samples such as urine, blood or breath samples. Whilst non-invasive tests have the advantage of reduced risk and inconvenience to the patient, some of these tests only detect exposure to *H. pylori* and do not determine whether an active infection is ongoing.

1.5.7.1 Invasive

1.5.7.1.1 Endoscopy

Endoscopy provides a definitive method for identifying inflammation of the gut and diagnosis of the disease type. Recent advances in endoscopic methods have provided the ability to visualise *H. pylori* within the gut (Kiesslich *et al.*, 2005b, Uedo *et al.*, 2006). Confocal laser endoscopy developed by Kiesslich and co-workers (Kiesslich *et al.*, 2005b) utilises intravenously administered fluorescence stains, fluorescein and acriflavin which when used in combination allow the visualisation of *H. pylori* (Kiesslich *et al.*, 2005a). Other endoscopic techniques such as narrow band endoscopy also allows the visualisation of *H. pylori* (Uedo *et al.*, 2006). Endoscopy also provides biopsy specimens that can be used for culture, histology, rapid urease test and PCR analysis, thus providing further tools to examine the infection and provide a systematic route for treatment.

1.5.7.1.2 Histology

The use of a variety of stains that are specific for inflammatory cells and *H. pylori* allow for (in 90% of cases) a diagnosis to be carried out within two to three days (Cutler *et al.*, 1995). The use of stains such as haematoxylin and eosin allow for staining of inflammatory cells to ascertain the type of disease, with giemsa and genta stains detecting *H. pylori*. The advantage of using the genta stain is its high sensitivity and low costs (Rotimi *et al.*, 2000). An updated Sydney system (Dixon *et al.*, 1994) provides a visual analogue scale to score *H. pylori* infection, though this is dependent on where within the gut the biopsy was taken.

1.5.7.1.3 Culture biopsy

H. pylori is gram-negative, urease positive, oxidase and catalase positive, meaning culturing of the bacteria is highly specific (Ricci *et al.*, 2007). However, culturing of *H. pylori* is difficult due to its growth only in a microaerophilic environment. After a sample has been taken from a patient, it is transferred to Stuart's transport medium for no longer than 24 hr (Kjøller *et al.*, 1991). An advantage of culturing *H. pylori* is its antibiotic resistance can be determined through the E-test (Mishra *et al.*, 2006). This can then save time and money in treatment of the pathogen. In addition, the culturing of antibiotic-resistant *H. pylori* allows molecular tests (Chapter 1.5.7.1.5) and biochemical work to investigate the molecular systems governing antibiotic resistance, with the aims of devising new treatments to those currently available to be investigated (Chapter 1.5.8).

1.5.7.1.4 Rapid Urease test

The rapid urease test utilises *H. pylori* urease activity. Upon the addition of urea, urease catalyses its hydrolysis to ammonia and hydrogen carbonate. The pH increase associated with ammonia production can be detected by phenol red. There are several commercial-based tests, which utilise, gel, paper or liquid-based techniques reviewed by Ricci (Ricci *et al.*, 2007). Approximately 10^4 organisms are required for positive detection of *H. pylori*.

Although it should be noted that patients suffering from achlorhydria or those on proton-pump inhibitors will give false negatives due to the action of *H. pylori*'s own urease giving a local high pH destroying the pathogen, discussed in (Basset *et al.*, 2004).

1.5.7.1.5 Molecular tests

Molecular tests are highly sensitive techniques that utilise samples taken from invasive gastric biopsies and from non-invasive stool samples (although to a much lesser extent). Molecular tests do not require living *H. pylori* and so it is possible to investigate antibiotic resistance without the prior need for culturing. However, no information concerning whether an infection is ongoing or has occurred can be obtained since the dead pathogens DNA will still be detected.

PCR-based methods can be used to detect clarithromycin resistance through the analysis of the 23S rRNA gene in *H. pylori*. Whereby it has been shown that single point mutations exist within the AT rich region of the 23S rRNA gene (Fontana *et al.*, 2002). The

disadvantage of PCR-based methods is contaminations resulting in false positives (Lisby, 1999).

1.5.7.2 Non-Invasive

1.5.7.2.1 Serology

This is the most widely used method for detecting exposure to *H. pylori*, although is not able to show whether an infection is ongoing, as it relies upon the detection of anti-*H. pylori* IgG antibodies. From the blood test, the antibodies can be detected by a variety of different methods including: enzyme-linked immunosorbent assay (ELISA), Latex agglutination and Western Blotting, reviewed in (Ricci *et al.*, 2007).

1.5.7.2.2 Urea breath test

The urea breath test is considered one of the most accurate non-invasive methods for the detection of *H. pylori*. ^{13}C urea is ingested whereby in the gastric mucosa it is hydrolysed to ^{13}C -labelled CO_2 , which diffuses into the epithelial layer and is expelled on the breath after 20 minutes. The amount of ^{13}C present is detected with mass spectrometry. ^{14}C urea has been used reviewed in (Ricci *et al.*, 2007) and is detected with a scintillation counter. Recent developments with an improvement in infra-red spectroscopy has made the method cheaper by not using mass spectrometry (Taniguchi *et al.*, 1996). Patients who have had gastric surgery or are on proton-pump inhibitors or ranitidine do not give a reliable result for the same reason as described in Chapter 1.5.7.1.4.

1.5.8 Treatment

Single therapy treatment against *H. pylori* is not efficient with only 40% eradication upon treatment with clarithromycin for 14 days (Peterson *et al.*, 1993). Other antibiotics commonly used against *H. pylori* include, azithromycin, amoxicillin and metronidazole. Recently rifabutin and furazolidone have been used (Perri *et al.*, 1998, Fakheri *et al.*, 2004), although these tend to be used in second-line rescue therapy in the instances of metronidazole resistance reviewed in (Kusters *et al.*, 2006). Triple therapy tends to be used and consists of two antibiotics and either a PPI or bismuth compound. This method of treatment is used in most countries leading to higher eradication rates of *H. pylori* infection (Kusters *et al.*, 2006). The action of bismuth compounds against *H. pylori* has been investigated but its mechanism of action is unknown (Armstrong *et al.*, 1987). The use of various combinations of antibiotics

has been investigated but eradication rates have never been able to surpass 95% references within (Kusters *et al.*, 2006).

Antibiotic resistance is the main cause for failure of *H. pylori* eradication, which tends to arise from improper patient adherence to the treatment, some of this is related to the side-effects of some of the antibiotics, such as nausea and abdominal pain (Bell *et al.*, 1992). The antibiotic furazolidone tends to cause particular problems amongst patients (Bell *et al.*, 1992). However, smoking has been shown to increase the chances of eradication failure (Suzuki *et al.*, 2006). To limit the antibiotic resistance it has been suggested that in areas with a resistance rate of 20%, culturing and antibiotic susceptibility testing should be carried out prior to treatment (Cavallaro *et al.*, 2006).

Vaccination against *H. pylori* would eliminate the issue of antibiotic resistance and a single vaccine would be more efficient than relying on patients to take a course of antibiotics over several days, whilst sometimes experiencing the side effects from their treatment. Several lines of research on developing a vaccination against *H. pylori* are focusing on inactivated whole cells, outer membrane vesicles with most research focusing on developing vaccines based on *H. pylori* antigens, reviewed in (Giudice *et al.*, 2001).

1.5.9 Virulence factors

Infection with *H. pylori* always results in chronic gastritis, though for some patients no further complications develop, this has led to the notion that some strains are more virulent than others (Blaser and Atherton, 2004). Investigations showed that strains that were more virulent were able to induce morphological changes of *in vitro* cultured cells (Leunk *et al.*, 1988). The most virulent strains were found to express the vaculating toxin, VacA and also cytotoxin-associated gene (*cagA*). The most virulent strains were referred to as Type I strains, whilst the less virulent strains as Type II strains (Xiang *et al.*, 1995). Only *cagA* gene is expressed in Type I strains whilst *vacA* is expressed in both types. Figure 1.5 summarises the suspected virulence factors and implies as to how they are thought to contribute to the associated diseases.

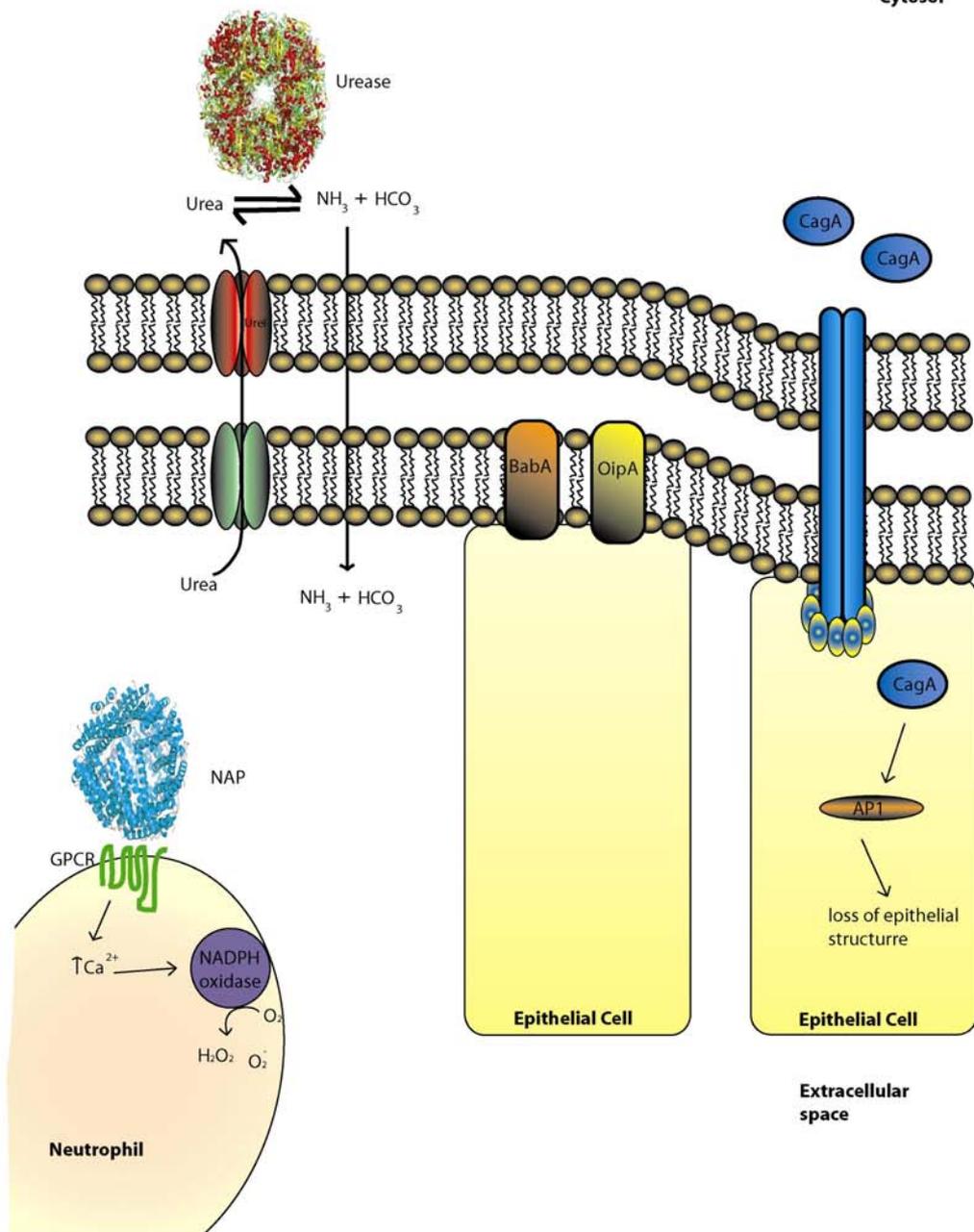


Figure 1. 5 Schematic representation of the suspected virulence factors of *H. pylori* and their possible mode of action within the gut epithelium resulting in many of the disease outcomes. A pH-activated urea transporter is thought to transport urea into the cytoplasm of *H. pylori* whereby it is acted upon by urease. Neutrophil Activating Protein (NAP) is an extracellular protein which results in Neutrophils activation leading to the inflammation of the gut. Various adhesions including, OipA and BabA attach *H. pylori* to the epithelial cell layer. The cytotoxin-associated gene (cagA) is thought to act on internal signalling pathways within epithelial cells, possibly providing life-long

1.5.9.1 Vacuolating Cytotoxin

The *vacA* gene encodes a pro-toxin, which undergoes proteolysis to yield an 88kDa mature toxin, which is then secreted by the bacterium. Further cleavage occurs within the extracellular space yielding a 34 kDa N-terminal and a 54 kDa C-terminal fragment, which appear associated as large macromolecular complexes of six to seven VacA monomers (Lupetti *et al.*, 1996). Acidic pH causes disassembly of the oligomeric complex into membrane-inserting monomers which form anion-selective channels releasing urea and bicarbonate that can be utilised by *H. pylori* (Molinari *et al.*, 1998a). The VacA channel is eventually internalised and changes the ion permeability of the endosomes leading to water influx and vesicle swelling (vacuolisation). Vacuolisation causes alteration in cell trafficking and endosomal protein degradation. This maybe important for the prolonged survival of *H. pylori* as impairment of antigen processing in the antigen-processing compartment reduces the number of surface-presented peptides to bind to MHC class II antigens (Molinari *et al.*, 1998b), hence invasion from the immune system.

1.5.9.2 Urease

Urease production makes up 6% of total protein production in *H. pylori* (Labigne and de Reuse, 1996, Eaton *et al.*, 1991) and catalyses the hydrolysis of urea to carbon dioxide and ammonia. Gnotobiotic pig models have shown that urease-deficient strains of *H. pylori* are unable to colonise the gastrointestinal tract (Eaton *et al.*, 1991), highlighting the importance of urease as a colonisation factor

Ureases are found in the cytoplasm of all bacteria but urease of *H. pylori* has also been reported associated with the surface of the bacterium, with up to 30% total urease found on the surface *in vivo* (Dunn *et al.*, 1997). This surface-associated urease has been thought to provide a “cloud” of ammonia to neutralise the gastric acid and allow *H. pylori* to survive until it was able to colonise the pH neutral gastric epithelial cells, which are protected by the mucosal layer (Labigne and de Reuse, 1996) and reviewed in (Moran, 1996). The crystal structure of urease shows it to be a nickel-containing dodecamer consisting of 62 kDa UreA subunit and 27kDa UreB subunit in an 4($\alpha\beta$)₃ arrangement (Ha *et al.*, 2001), unlike most other bacterial ureases which contain three subunits ($\alpha\beta\gamma$) (Marais *et al.*, 1999). Previous studies had shown that isolated urease is inactivated at pH <5, Ha *et al* showed that conditions that mimicked physiological conditions did not inactivate the enzyme at pH 3 and proposed (based on the crystal structure) that the 12 active sites were protected by additional residues.

More recent evidence has cast doubt over extracellular urease providing acid resistance (Sachs *et al.*, 2005) and (Marcus and Scott, 2001). The monitoring of $^{14}\text{CO}_2$ emissions from intact and lysed bacteria at varying pH has showed that at lower pHs there is an increase in $^{14}\text{CO}_2$ production from intact bacterium implying the existence of a pH-activated urea transporter which was shown to be UreI (Scott *et al.*, 2002). Deletion of the integral membrane *ureI* gene prevented the pH-dependent increase in urease activity (Scott *et al.*, 2002).

1.5.9.3 Cytotoxin-Associated Gene

The *cagA* locus is found only in Type I strains and occupies a 40 kbp region of the *H. pylori* genome (Censini *et al.*, 1996). This region is known as the pathogenicity island (PAI) which is composed of 31 genes, some of which encode for type IV secretion machinery allowing for transfer of molecules into eukaryotic cells. The secretion machinery allows the transfer of the 140 kDa CagA, peptidoglycan and other factors into the host cell (Christie and Vogel, 2000, Fischer *et al.*, 2001). Once inside the host cell CagA interacts with a variety of proteins through its initial interaction with the tyrosine phosphatase SHP-2 (Yamazaki *et al.*, 2003), leading to morphological changes of the epithelial cells (Moese *et al.*, 2004) (Naumann *et al.*, 1999). CagA is also phosphorylated by src kinase and then binds to the kinase via its SH2 domain leading to inactivation of the kinase and hence no further phosphorylation of CagA (Selbach *et al.*, 2003, Tsutsumi *et al.*, 2003). This negative feedback regulating CagA activity may be responsible for allowing lifelong colonisation of the host by ensuring survival of the epithelial cell (Kusters *et al.*, 2006).

1.5.9.4 Neutrophil Activating Protein

Neutrophil Activating Protein (NAP) is a 150 kDa dodecamer that is produced within the cytosol of *H. pylori*, the crystal structure shows each subunit is comprised of a four-helix bundle (Dundon *et al.*, 2001, Tonello *et al.*, 1999)). Purified NAP is able to induce chemotaxis of monocytes and neutrophils (Satin *et al.*, 2000). It is believed that upon bacterial lysis, NAP is released and binds to and activates neutrophils, possibly through a G-protein-coupled receptor (GPCR) (Satin *et al.*, 2000). NAP binding to the GPCR is thought to activate an internal signalling cascade that results in release of intracellular calcium leading to the assembly of a functional NADPH oxidase (Evans *et al.*, 1995) (Satin *et al.*, 2000). Activation of NADPH oxidase catalyses the superoxide anion and hydrogen peroxide formation from molecular oxygen, which have been shown to contribute to tissue damage (Rautelin *et al.*, 1994). NAP does not act as an adhesion (Blom *et al.*, 2001) and other factors are required to stimulate a burst in polymorphonuclear neutrophils through adhesions of the bacterium (Allen, 2001). Neutrophils are found in far greater excess than macrophages in *H. pylori*-induced infected gastric mucosa (Ernst, 2000, Allen, 2001) and it is speculated that the activation of neutrophils resulting in inflammation of the mucosa causes the release of nutrients to support the growth of *H. pylori* (Dundon *et al.*, 2001).

1.5.9.5 Lipopolysaccharide

LPS are fucosylated oligosaccharide antigens that are structurally similar to human blood group antigens (Kusters *et al.*, 2006). In other gram-negative bacteria Lewis antigens activate the innate immune response, whilst in *H. pylori* they are a weak activator of the innate immune system (Muotiala *et al.*, 1992, Bliss *et al.*, 1998). One possibility for bacteria possessing LPS that are similar to human epitopes is to avoid detection by the immune system and hence allow it to persist as a life-long infection.

1.6 Protection against Oxidative Stress in *H. pylori*

H. pylori uses oxygen as a terminal electron acceptor and is unable to use nitrite or formate, as is the case for anaerobic organisms, thus *H. pylori* requires 2% oxygen for survival (Mendz *et al.*, 1997). Several genes are found within *H. pylori*'s genome to overcome oxidative damage resulting from molecular oxygen as a terminal electron acceptor and oxidative stress generated by the hosts immune system. *H. pylori* express an Fe-dependent superoxide dismutase (Fe-SOD), which catalyses the reduction of superoxide to hydrogen peroxide (Spiegelhalder *et al.*, 1993) and a mono-functional catalase that lacks peroxidase activity, catalysing the breakdown of hydrogen peroxide to water and oxygen (Hazell *et al.*, 1991). Alkylhydroperoxide reductase also is found, which catalyses the NADH-dependent reduction of alkylhydroperoxides to the corresponding alkyl alcohol (Pesci and Pickett, 1994, Alm *et al.*, 1999).

1.7 Metabolism within the Campylobacterales

The determination of the first genomic sequence of *H. pylori* (Tomb *et al.*, 1997), *C. jejuni* (Parkhill *et al.*, 2000) and *W. succinogenes* (Baar *et al.*, 2003b) has led to the identification of genes involved in virulence, replication and metabolism. This analysis in conjunction with more classical biochemical approaches has provided insight into the *Campylobacterale*'s ability to colonise the gastrointestinal system of different species and in the case of *H. pylori* and *C. jejuni*, contribute towards the different disease associated with them. Like any analysis, parallels can only be made to existing systems where experimental data has confirmed biological processes. Where, for example, in *H. pylori* knowledge of the genomic sequence allows inference on the metabolic processes and only the combination of experimental knowledge from the organism and parallels drawn with similar species supports and builds upon the known genomic model. In general, the mapping of metabolism from other organisms to a particular organism based upon its genomic sequence does not give a full understanding of that organism's metabolism. As is always the case, experimental observations provide the greatest amount of insight into an organism's metabolism, where hypothesis based upon genomic data can be tested. This is no exception to *H. pylori* and the other genera within the *Campylobacterales*. In the case of *H. pylori*, genomic data has provided hypothesis to test metabolic studies, only experimental work confirms or disproves these hypotheses, and this is summarised in the following sections. The remainder of the Chapters within this thesis are a continuation of this argument and demonstrate the importance of experimental work in providing questions about the metabolism of *H. pylori*

and *C. jejuni* based on the work of three genes within metabolism of these species. The following sections will concentrate on metabolism within *H. pylori*, which is considered similar to that of *C. jejuni* (Eppinger *et al.*, 2004).

1.7.1 Carbohydrate Metabolism

Glucose is the only carbohydrate source that can be metabolised by *H. pylori* (Mendz *et al.*, 1993) with genomic analysis suggesting the lack of other carbohydrate permeases and only the presence of the GluP glucose/galactose transporter (Tomb *et al.*, 1997). Once within the cell phosphorylation of glucose is performed by a glucokinase rather than a hexokinase (Mendz and Hazell, 1993), again confirming the lack of utilisation of other carbohydrates. It is believed that glucose is not the main source of energy but is utilised when other sources have been exhausted (Mendz *et al.*, 1994a). Experimental work has confirmed enzymatic activities of the oxidative and nonoxidative parts of the pentose-phosphate pathway (Mendz *et al.*, 1995a), which leads to the generation of NADPH, which is utilised in various other anabolic processes and the generation of five carbon sugars for nucleotide biosynthesis. Several groups have been unable to detect activities of some enzymes within glycolysis (Mendz *et al.*, 1994a, Chalk *et al.*, 1994) with genomic analysis confirming the absence of phosphofructokinase and pyruvate kinase (see reaction scheme in Figure 1.6). Though such data should be viewed in context as the gene encoding for a putative phosphoglycerate mutase has been identified but no activity has been reported (Marais *et al.*, 1999). An alternative pathway to glycolysis, the Entner-Doudoroff pathway is able to catalyse the formation of pyruvate and glyceraldehyde-3-phosphate from 2-keto-3-deoxy-gluconate-6-phosphate. The energy yield from the Entner-Doudoroff pathway is less than that from glycolysis but it provides an alternative route and a mode of metabolising aldonic acids. The Entner-Doudoroff pathway is present in *E. coli* but is inducible, whereas in *H. pylori* it is constitutive, suggesting its role to substitute for an apparently incomplete glycolysis. Enzymatic activity has also been reported for most of the key, irreversible reactions of gluconeogenesis (Mendz and Hazell, 1993) with the corresponding genes involved identified however, no gene has been identified for the last irreversible step of gluconeogenesis, the dephosphorylation of glucose-6-phosphate. Although it could be argued that, this step of gluconeogenesis would seem irrelevant, as there is no storage of glucose within *H. pylori*.

Figure 1.6 shows the classical carbohydrate metabolic pathways; glycolysis, gluconeogenesis, pentose-phosphate shunt and the Entner-Doudoroff pathway. All pathways

can feed into each other, therefore no pathway functions in complete isolation. This is not a complete picture of all the metabolic processes known within *H. pylori* (or *E. coli* in general) as this is out of the scope of this work. However, on a simplified scale, one of the common links between all the pathways shown is through the three-carbon sugar glyceraldehyde-3-phosphate.

Glyceraldehyde-3-phosphate is the first three-carbon sugar of glycolysis (last of gluconeogenesis) and is the entry point from the Entner-Doudoroff pathway into the glycolytic pathway for the metabolism of glucose, through the absence of the first irreversible step of glycolysis. Glyceraldehyde-3-phosphate is also utilised by the pentose-phosphate pathway. In glycolysis glyceraldehyde-3-phosphate is acted upon by glyceraldehyde-3-phosphate dehydrogenase (GAPDH) by an oxidative phosphorylation reaction (discussed in greater detail in Chapter 1.8.1) to yield 1,3-bisphosphoglycerate. GAPDH is also able to catalyse the reversible "gluconeogenic" reaction generating glyceraldehyde-3-phosphate and inorganic phosphate from 1,3-bisphosphoglycerate and NADH. In *H. pylori*, two genes have been identified for GAPDH; *gapA* (gi:6626253) and *gapB* (gi:2494645).

Inspection of the genome of *W. succinogenes* within the *Helicobacteraceae* family has shown that key enzymes of the oxidative pentose phosphate pathway and the Entner-Doudoroff pathway are absent, whereas glycolysis and gluconeogenesis appear to be intact (Baar *et al.*, 2003a). Genes encoding for enzymes for glucose transport and other carbohydrate transporters also appear absent in *W. succinogenes*'s genome and probably glucose metabolism does not occur. Therefore, a gluconeogenic pathway exists and enzymes for gluconeogenesis have been identified in the *W. succinogenes* genome. This is in contrast to *H. pylori* where enzymatic activities for glycolysis have not been detected whereas, the oxidative pentose phosphate pathway and the Entner-Doudoroff pathway have and this is supported by genomic analysis. This highlights how each family is diverse due to their complex relationship with their host organism

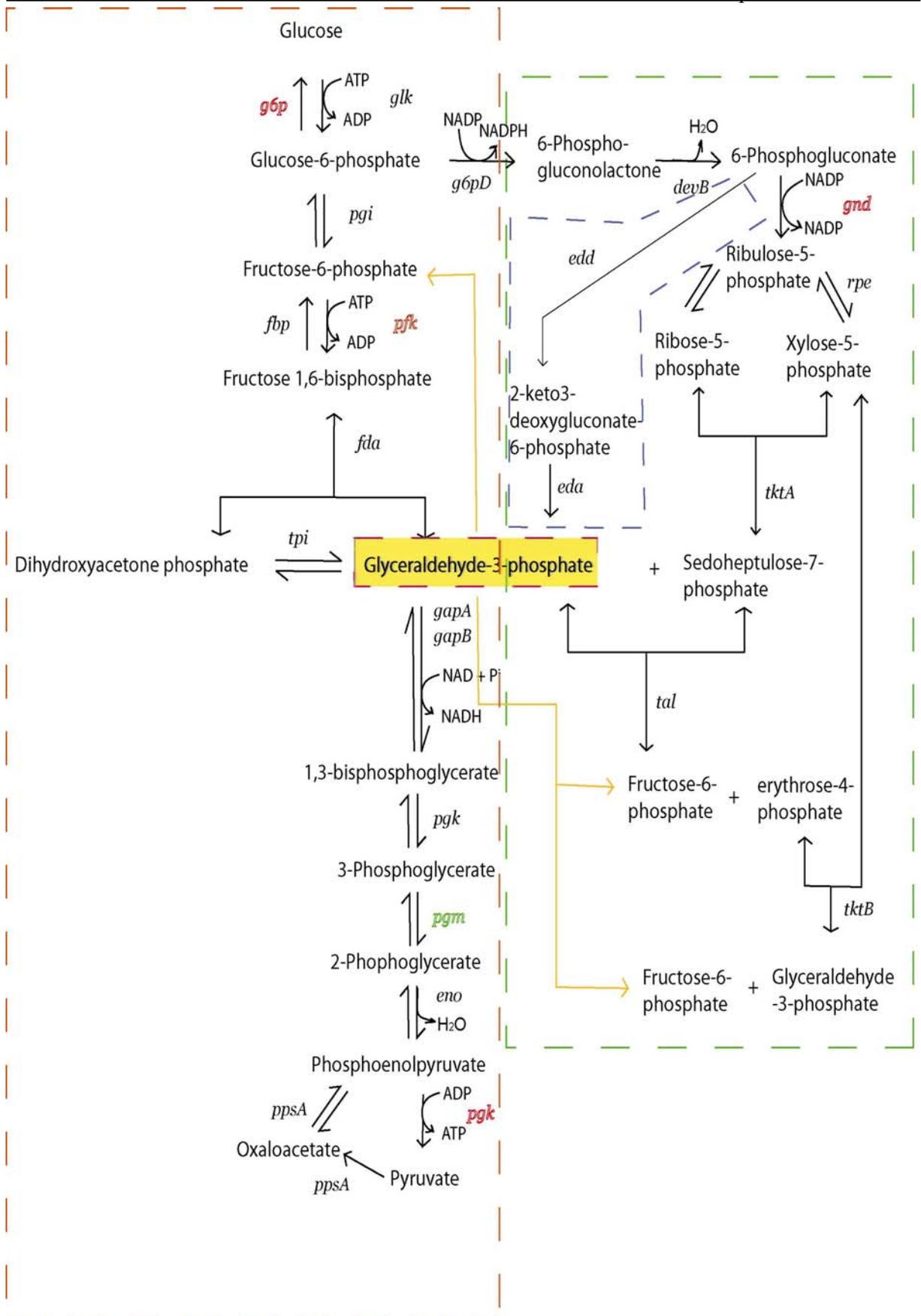


Figure 1. 6 The main carbohydrate metabolic pathways in *H. pylori*. The area enclosed by the blue box is the Entner-Doudoroff pathway, green box is the pentose-phosphate pathway, the red box is the glycolytic and gluconeogenic pathways. Enzyme genes are shown in italics with those in red not being identified in the published genome. The gene *pgm* is shown in green as the gene has been identified but no enzyme activity has been detected *in vitro*. The genes encoding the enzymes of the pathway are: *g6p*, glucose-6-phosphatase; *glk*, hexokinase; *pgi*, phosphoglucose isomerase; *pfk*, phosphofructo kinase; *fbp*, fructose 1,6-bis phosphatase; *fda*, aldolase; *tpi*, triose phosphate isomerase; *gapA/B*, GAPDH; *pgk*, phosphoglycerate kinase; *pgm*, phosphoglycerate mutase; *eno*, enolase; *pgk*, pyruvate kinase; *pps*, phosphoenolpyruvate carboxykinase; *g6pD*, glucose-6-phosphate dehydrogenase; *devB*, 6-phosphogluconolactonase, *gnd*, 6-phosphogluconate dehydrogenase; *rpe*, phosphopentose isomerase; *tktA*, transketolase; *tal*, transaldolase; *tktB*, transketolase; *edd*, 6-phosphogluconate hydolyase; *eda*, 2-keto-3-deoxygluconate 6-phosphate aldolase.

Until now, this study has focused on the *Campylobacterales* order, focusing on the *Helicobacter pylori* species. The morphology of the organism has been examined and its involvement in disease has been highlighted. The main virulence factors thought to bring about disease have also been highlighted. Current antibiotic treatments have also been described, where it is clear that due to the onset of antibiotic resistance new treatments are required. The previous and next sections highlight the little knowledge at the molecular level of how the bacteria's metabolism is fully understood. Analysis of what is known about the bacteria's metabolism shows that one divergence point of several pathways is glyceraldehyde-3-phosphate. Therefore, this study sought to undertake a structural approach to investigate how two enzymes, of annotated identical function, act upon glyceraldehyde-3-phosphate, with the view that they could potentially become drug targets and promote further research into the metabolism within the *Campylobacterales*, in particular *H. pylori*. The next few sections will investigate in greater detail the structure and mechanism in context of GAPDHs from all three domains of life.

1.8 Aldehyde Dehydrogenase Super family

The aldehyde dehydrogenase (ALDH) super family of enzymes catalyse the NAD- and NADP-dependent oxidation of aldehydes to the corresponding carboxylic acids. These enzymes are found throughout all domains of life with several ALDHs found in each organism. For example, *E. coli* has 17 identified ALDHs (Sophos and Vasiliou, 2003). Based upon sequence analysis, the ALDH superfamily has been categorised into several groups: semialdehyde dehydrogenases, non-specific ALDHs, betaine dehydrogenases, non-phosphorylating glyceraldehyde-3-phosphate dehydrogenases, phenylacetaldehyde dehydrogenases, lactaldehyde dehydrogenases and ALDH-like proteins (Sophos and Vasiliou, 2003). To date a variety of structures of aldehyde dehydrogenases have been solved all displaying a wide range of substrate and coenzyme specificity. All of the structures solved support the mechanism for NAD(P)⁺-dependent aldehyde oxidation which consists of the following steps; initially nucleophilic attack by an active site cysteine at the carbonyl group, followed by collapse of the tetrahedral intermediate with hydride transfer to NAD(P)⁺ and subsequent hydrolysis of the thioester intermediate yielding the free carbonic acid (Racher, 1955). A typical monomeric fold of an aldehyde dehydrogenase is shown below in Figure 1.7 Briefly the enzyme consists of an nucleotide binding domain which is a typical Rossmann fold (Buehner *et al.*, 1974), a catalytic domain and a hydrophobic oligomerisation domain.

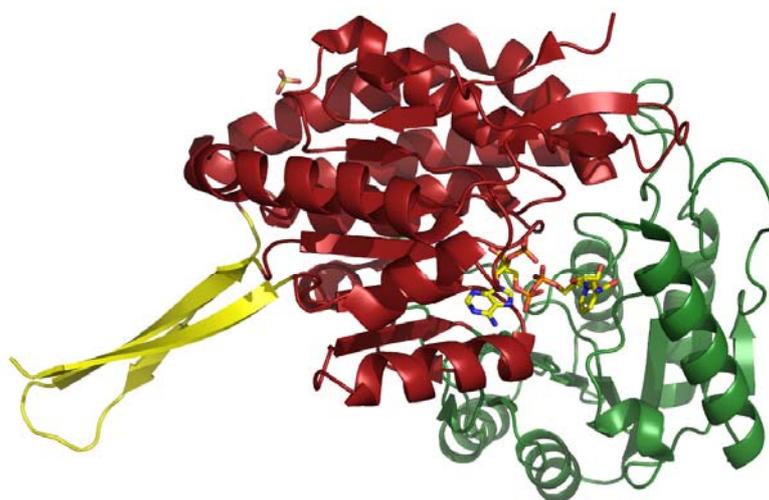
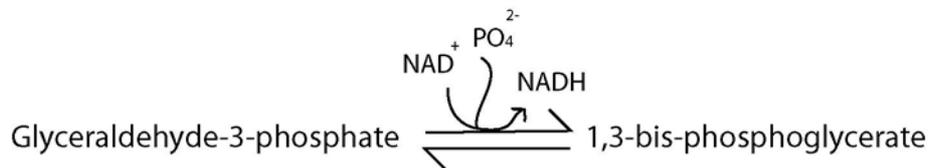


Figure 1. 7 Monomeric structure of lactaldehyde dehydrogenase (2ILU).
The nucleotide binding domain is shown in red, with the catalytic domain in green and the oligomerisation domain in yellow.

1.8.1 Glyceraldehyde-3-phosphate Dehydrogenase

Glyceraldehyde-3-phosphate Dehydrogenase catalyses the reversible oxidative phosphorylation of glyceraldehyde-3-phosphate to 1,3-bisphosphoglycerate resulting in the reduction of one molecule of NAD^+ to NADH . After hydride transfer, phosphate attack causes the collapse of the hemi-thioacyl intermediate, generating 1,3-bisphosphoglycerate. This is in contrast to aldehyde dehydrogenases where water attack results in the acid being formed. The function of the inorganic phosphate attack by GAPDH is to yield the highly unstable 1,3-bisphosphoglycerate. The high energy acyl-phosphate bond within 1,3-bisphosphoglycerate allows for the transfer of the phosphate to ADP by phosphoglycerate kinase, yielding the high energy store ATP.



There are three main types of GAPDH with different subcellular localisations, performing different roles. Firstly, the classical glycolytic NAD^+ -dependent enzyme found in all organisms (EC 1.2.1.12). Secondly, the NADP^+ -dependent GAPDH (EC 1.2.1.13), a key component of the reductive pentose phosphate pathway in the chloroplast stroma and the cytoplasm of cyanobacteria. Finally, the cytosolic non-phosphorylating GAPDH (EC 1.2.1.9), which is actually an ALDH rather than a GAPDH (Habenicht *et al.*, 1994), which catalyses the irreversible NADP^+ -dependent oxidation of glyceraldehyde-3-phosphate to 3-phosphoglycerate. In *Archaea* there is also a non-phosphorylating GAPDH that does not utilise NAD^+ or NADP^+ but rather a ferredoxin for oxidation of glyceraldehyde-3-phosphate, this enzyme termed GAPOR will not be discussed further as it has only been identified in a few *Archaea* and proceeds through a different mechanism irrelevant to this study (Van der Oost *et al.*, 1998).

1.8.1.1 Glycolytic GAPDH

The classical glycolytic GAPDH has been extremely well characterised and catalyses the reversible NAD⁺-dependent oxidative phosphorylation of glyceraldehyde-3-phosphate to 1,3-bisphosphoglycerate as previously mentioned. Its structure and mechanism is discussed below.

1.8.1.1.1 Structure of GAPDH

The crystal structures of GAPDHs have been determined from across the three domains of life, from *Eukarya*; human muscle (Mercer *et al.*, 1976), lobster (Murthy *et al.*, 1980, Buehner *et al.*, 1974, Shen *et al.*, 2000), yeast (Ferreira-da-Silva *et al.*, 2006) and plants (Fermani *et al.*, 2001) to *Bacteria*; (Skarzynski *et al.*, 1987, Duee *et al.*, 1996, Yun *et al.*, 2000), including thermophiles (Skarzynski *et al.*, 1987, Tanner *et al.*, 1996, Korndorfer *et al.*, 1995) and finally *Archaea*; (Isupov *et al.*, 1999, Charron *et al.*, 2000). In all structures with the exception of non-photosynthetic organisms (discussed below) GAPDH exists as a 222 symmetry-related homo-tetramer. A typical monomeric structure of GAPDH from *Bacillus stearothermophilus* is shown in Figure 1.8. For all GAPDHs each monomer is composed of two domains. A NAD(P)⁺ binding domain (residues 1-147) which has the topology of a Rossmann fold (Buehner *et al.*, 1974) consisting of a 3-layer $\alpha\beta\alpha$ sandwich architecture. The intermolecular interactions that occur between neighbouring subunits within the tetramer have been well characterised, with most interactions occurring across the P-axis and least across the Q-axis (Skarzynski *et al.*, 1987), as is clear from Figure 1.9B.

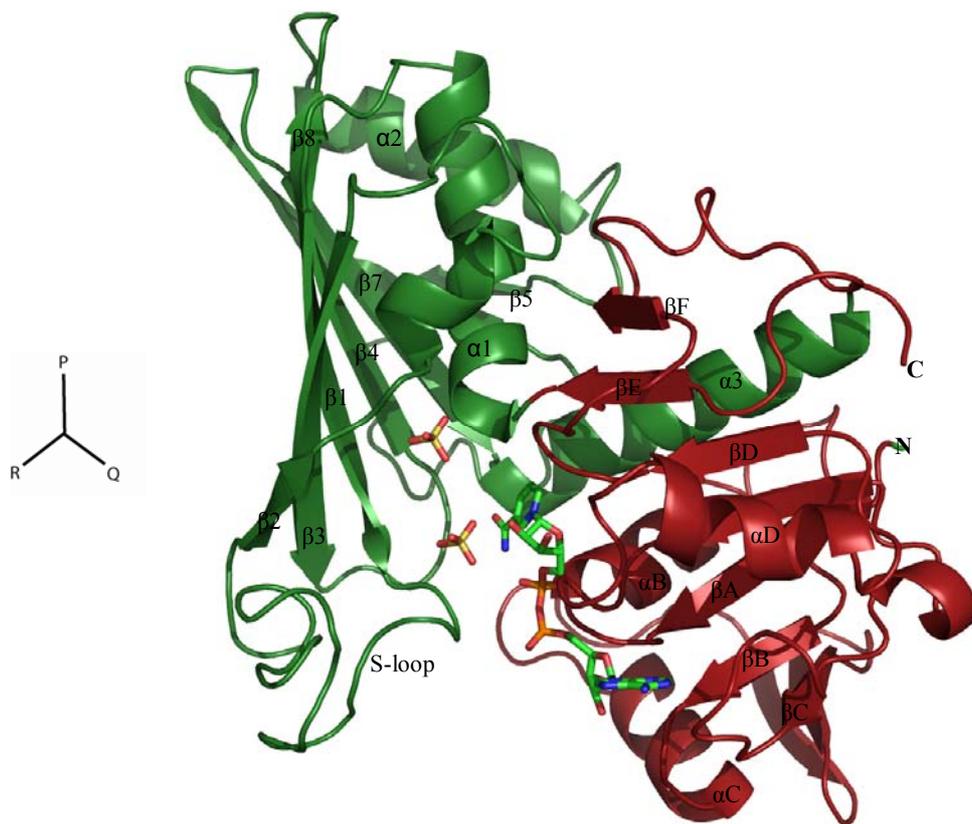


Figure 1. 8 Typical structure of a subunit of GAPDH from *Bacillus stearothermophilus* (1GD1). Residues 1-145 and 330-335 form the nucleotide binding domain displaying a typical Rossmann fold (Moras *et al.*, 1975) and is shown in red. The catalytic domain consists of residues 146-329 and is shown in green. NAD⁺ and two sulphate molecules are depicted as sticks. For clarity the coordinate axis are shown as depicted for the orientation for the tetramer as in Figure 1.9.

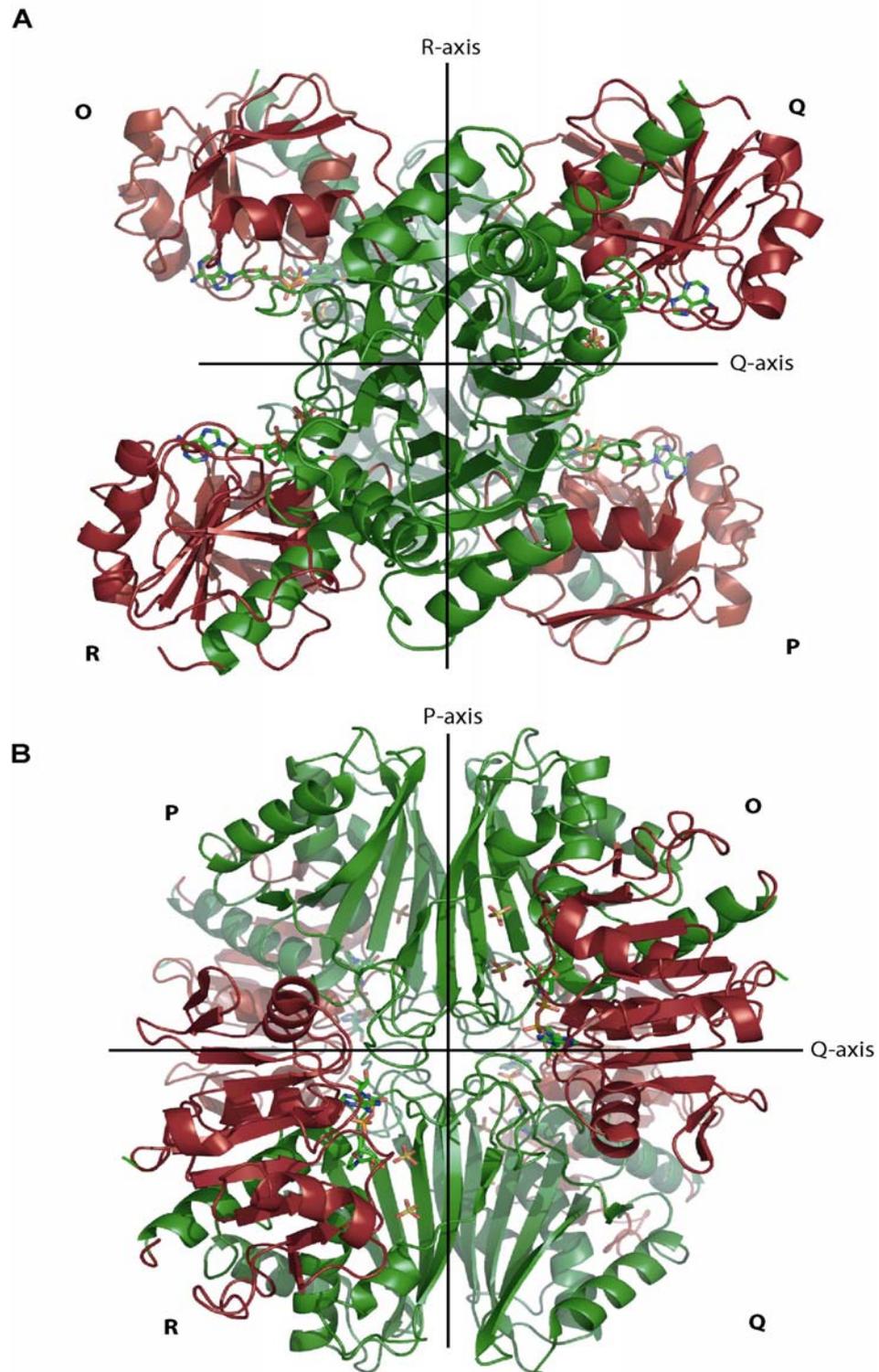


Figure 1. 9 Tetrameric arrangement of *Bacillus stearothermophilus* GAPDH, showing the arrangement of the monomer within the tetramer using the same colouring scheme as in Figure 1.8 for the catalytic and nucleotide binding domains. Individual subunits are denoted O-R in accordance with the original nomenclature. Panel A is the view along the P-axis whilst panel B depicts the view along the R-axis

1.8.1.1.2 Reaction Mechanism

The reaction mechanism of phosphorylating GAPDH has been intensely investigated for over half of a century. Steady state kinetics first suggested a ping-pong reaction (Krimsky and Racker, 1955), (Velick and Hayes, 1953), (Trentham, 1971) (Hilvers *et al.*, 1964)) in which a stable enzyme-NAD⁺ complex is formed. The enzyme-NAD⁺ complex is observed through spectroscopy as a broad extensive band extending from 300 to 360nm in the UV spectrum, known as the Racker band (Racker and Krimsky, 1952).

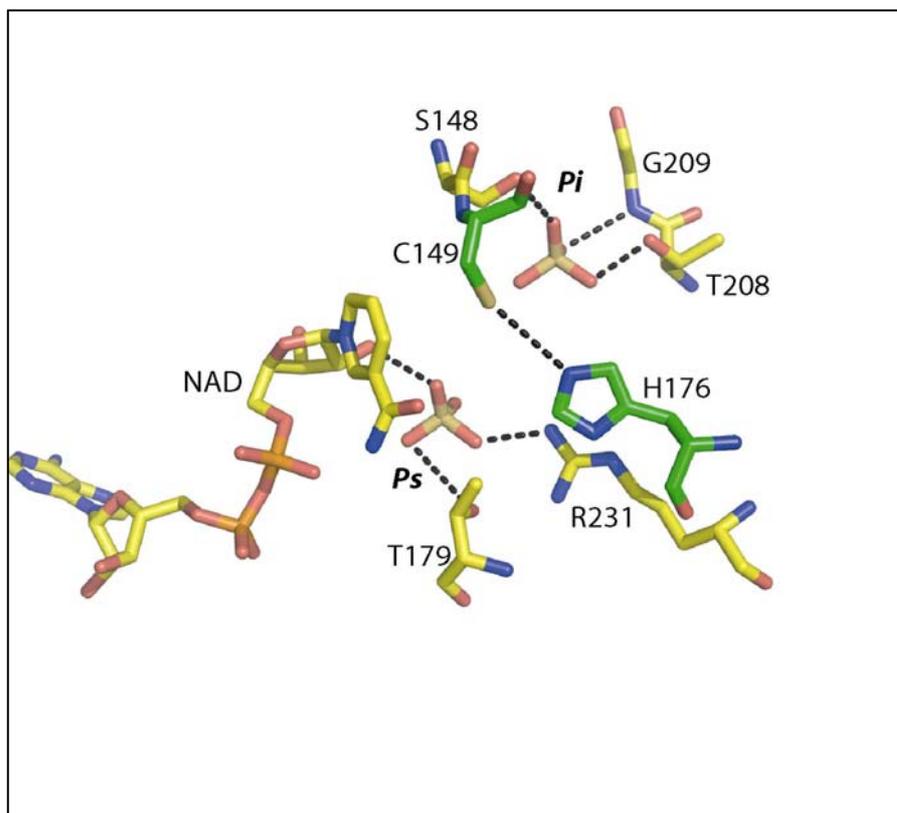


Figure 1. 10 Catalytic site of holo-GAPDH from *Bacillus stearothermophilus*. The location of the substrate phosphate-binding site, Ps, formed by residues T179 and R231. The nicotinamide ribose 2' hydroxyl group also contributes to the Ps site. The Pi site, where the attacking inorganic phosphate binds is formed from residues S148, T208 and G209. The catalytic cysteine C149 and histidine H176 which stabilises the C149 thiolate and also stabilises the Michaelis complex and acts as an acid and base during the reaction.

The Racker band is believed to be a charge-transfer band formed from the close contact between the active site cysteine (Cys149) and the nicotinamide moiety of NAD⁺, which was inferred from the crystal structure (Moras *et al.*, 1975). During the reaction process, the Racker band disappears upon binding of substrate. Site directed mutagenesis demonstrated that Cys149 is the active site cysteine (Mougin *et al.*, 1988), (the addition of

iodoacetate had shown that a cystine was the residue responsible for attacking the aldehyde (Racker and Krimsky, 1952)). Binding of substrate occurs simultaneously at all four active sites within the tetramer (Krimsky and Racker, 1955, Leslie and Wonacott, 1984). Interestingly, inhibitors or analogues of either NAD⁺ or g3p bind to GAPDH as half-of-sites reactivity (Ehrenfeld *et al.*, 1981, Levitzki, 1974) where the enzyme functions not as a tetramer but as a dimer of dimers (Levitzki, 1974). The rate determining step is the exchange between NADH and NAD⁺ and this requires a high pH (Trentham, 1971). Therefore, the assay is performed at pH 8.4. Once NAD⁺ has replaced NADH then phosphate attack occurs releasing 1,3-bisphosphoglycerate, summarised in the reaction scheme in Figure 1.11.

Coenzyme binding is strongly cooperative, with negative cooperativity being observed for the rabbit muscle and bacterial GAPDH's (Peczon and Spivey, 1972, Biesecker and Wonacott, 1977). Asymmetry has been reported for *Bacillus stearothermophilus* GAPDH when only one NAD⁺ molecule occupies the tetramer (Leslie and Wonacott, 1984). Upon binding of coenzyme to the remainder of the tetramer, a conformational change between the apo and holo structures is observed (Leslie and Wonacott, 1984). Whereby, Leslie and Wonacott showed a rotation of approximately 4 ° of the nicotinamide binding domain with respect to the catalytic domain. A similar effect on binding coenzyme is noted between the crystal structures of apo and holo GAPDHA from *H. pylori*, discussed in Chapter 3.

Crystallographic analysis of the holo-GAPDH from *Bacillus stearothermophilus* provided insight into the reaction mechanism of GAPDH (Skarzynski *et al.*, 1987). Skarzynski and co-workers observed within the active site the location of two sulphate molecules, which were within two distinct sites, believed to be phosphate-binding sites, previously modelled by (Moras *et al.*, 1975) as shown in Figure 1.10.

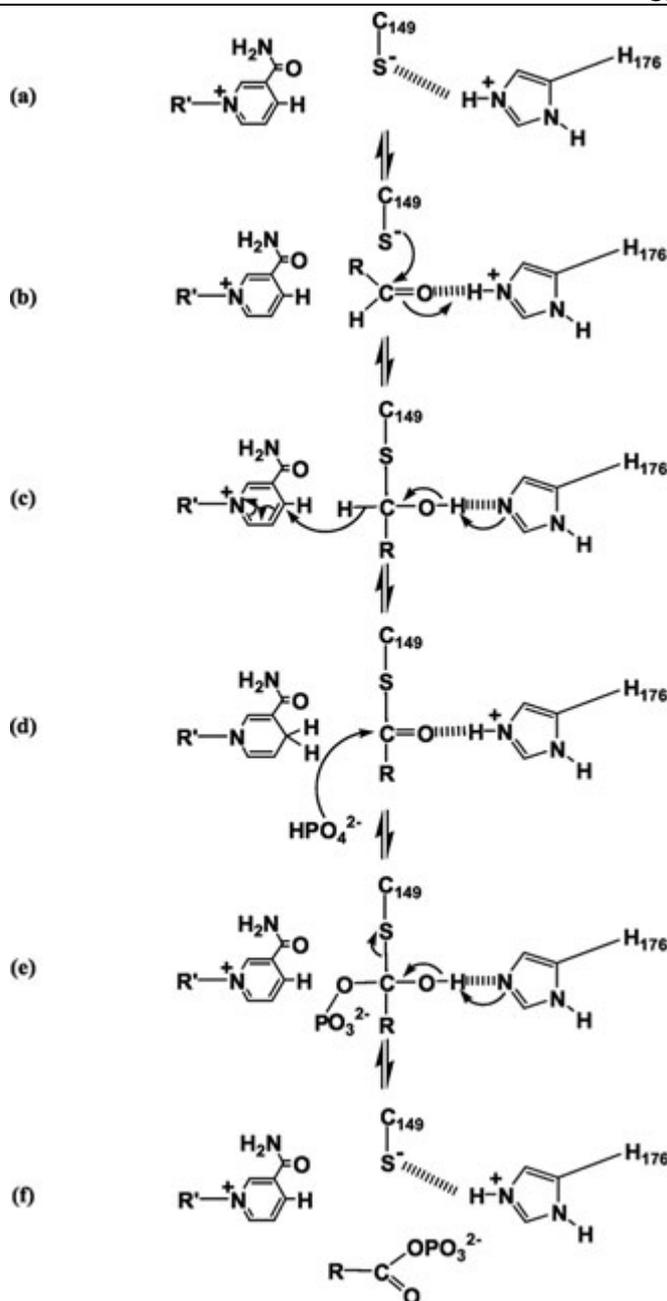


Figure 1. 11 Proposed reaction mechanism for GAPDH, taken from (Didierjean *et al.*, 2003). Briefly, the base-stabilised thiolate attacks the carbonyl of glyceraldehyde-3-phosphate yielding the thioacyl intermediate stabilised by the catalytic histidine. Hydride transfer is permitted from the carbonyl to NAD⁺ yielding NADH. A new molecule of NAD⁺ binds causing the phosphate of the acyl intermediate to move out of the Pi site it is currently occupying into the Ps site formed by the binding of a new molecule of NAD⁺. This allows phosphate to bind to the Pi site and attack the thio-acyl intermediate, leading to collapse of the intermediate and release of the substrate.

These phosphate-binding sites are known as the Pi and Ps site and correspond to the inorganic phosphate and substrate phosphate sites respectively. Further work involving the non-reversible inhibitor, glycidol phosphate that mimicked the thioacylenzyme intermediate but could not act as a leaving group due the absence of a hydroxyl group showed that the substrate phosphate occupied the Pi site and could not be modelled into the Ps site (Skarzynski *et al.*, 1987). Although the attacking phosphate has to attack the thio-acylenzyme intermediate from the Pi site and during this stage of reaction the substrate phosphate has to be in the Ps site. This lead to the hypothesis of a flip-flop mechanism, shown in Figure 1.11, whereby the substrate phosphate binds to the Pi site in the formation of the thio-acylenzyme intermediate and hydride transfer occurs reducing the NAD^+ to NADH . Next the substrate phosphate moves out of the Pi site and occupies the Ps site formed by the binding of a new NAD^+ molecule, allowing inorganic phosphate to bind to the Pi site and attack the thio-acyl intermediate causing its collapse and release of 1,3-bisphosphoglycerate. Recent crystallographic analysis has alluded to the flip-flop mechanism. Binding of g3p to *E. coli* GAPDH in the absence of coenzyme demonstrated that the substrate phosphate group occupied both the Pi and Ps sites (Yun *et al.*, 2000). Recently the ternary complex between the enzyme acyl-intermediate and the coenzyme has been reported where the substrate phosphate occupies the Pi site (Moniot *et al.*, 2008). All these structures have the substrate phosphate occupying either the Pi or the Ps site when coenzyme is bound. Chapter 6 presents the ternary complexes where both the Pi and Ps sites are occupied with coenzyme bound.

1.8.1.2 Photosynthetic GAPDH

Photosynthetic GAPDHs of cyanobacteria, algae and higher plants utilise NAD^+ and NADP^+ as coenzymes with a preference towards NADP^+ . These GAPDHs are located within the chloroplasts of plants and are composed of either one type of subunit (A) forming a homotetramer (A_4) or two types of subunits resulting in a hetrotetramer (A_2B_2). The homotetrameric form of GAPDH which is found in lower and higher plants is specific for NADP^+ (Cerff and Chambers, 1979, Ferri *et al.*, 1990). The crystal structure of the GAPDH A_4 isoform from spinach chloroplasts has been solved and demonstrates the atomic preference for NADP^+ over NAD^+ (Fermani *et al.*, 2001, Falini *et al.*, 2003). The specificity for NADP^+ over NAD^+ in the structure from spinach chloroplasts is determined through a hydroxylated side chain close to the 2'-phosphate group that in this case is threonine. A positively charged arginine interacts with the 2'-phosphate group and further stabilises the adenine ring through a stacking effect. There is also an interaction between a lysine residue from the s-loop from the R-axis related subunit which interacts with the 2'-phosphate (Fermani *et al.*, 2001). These

interactions are discussed in detail in Chapters 3 and 4 within the context of NADP binding for *H. pylori* and *C. jejuni* respectively.

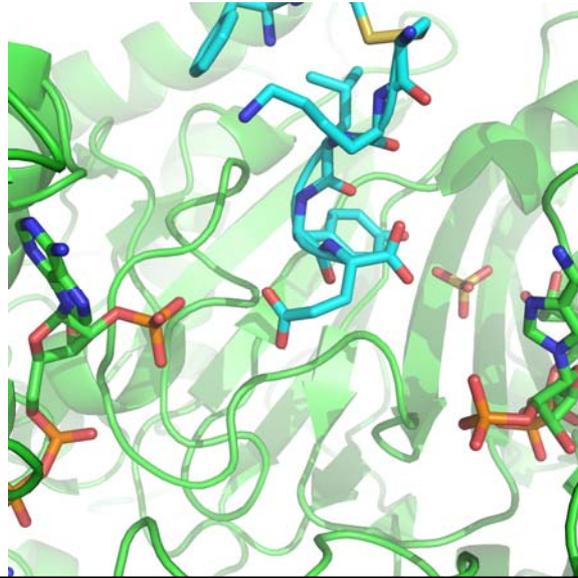


Figure 1. 12 Interaction of the 2' phosphate group of NADP through the CTE, shown with blue carbons. Once the enzyme is oxidised through the action of thioredoxin the enzyme is inhibited, and is only active once the disulphide bridge in the CTE is reduced releasing NADP⁺ and freeing the Ps site.

For the A₂B₂ GAPDH in higher plants the B subunit contains a C-terminal extension (CTE) which regulates the activity of the enzyme in light and dark reactions, through the redox state of two cysteines in the CTE, which can form a disulphide bond through action of thioredoxin. The crystal structure of the A₂B₂ in the oxidised form has been recently reported (Fermani *et al.*, 2007), shown in Figure 1.12. Where it was shown that oxidation of the CTE resulted in no coordination to the 2'-phosphate of NADP⁺ due to the interacting arginine forming a salt-bridge with a glutamate within the CTE (Fermani *et al.*, 2007). This level of control is in stark contrast to the regulation of GAPDHs in bacteria and mammalian cells. Within these organisms the metabolic flux of the pathway is controlled at other key enzymes other than GAPDH, for example 6-phosphofructo kinase. The regulation of these enzymes at key points in the pathway allow for enzymes like GAPDH to be controlled by the equilibrium of the system.

1.8.1.3 Non-Phosphorylating GAPDH

Non-phosphorylating GAPDH (GAPN) appears limited to within *Archaea*, which utilise a semi-phosphorylating branch of the Entner-Doudoroff pathway for sugar metabolism (Siebers and Schonheit, 2005). GAPNs are NADP⁺-dependent and have closer sequence homology to aldehyde dehydrogenases, rather than GAPDHs, thus they constitute a subdivision of the aldehyde dehydrogenase super family (Ettema *et al.*, 2008). The product from the oxidative hydrolysis of GAPN, 3-phosphoglycerate, can be utilised further in the glycolytic pathway, the result of which, is no ATP generation due to the reaction by phosphoglycerate kinase not occurring. The crystal structure of GAPN from *Streptococcus mutans* revealed an invariant threonine was responsible for the acylation step due to its interaction with the nicotinamide ring, positioning the ring for hydride transfer (Pailot *et al.*, 2006). Due to all the GAPDHs within this study being phosphate-dependent, GAPN is not discussed any further.

1.8.2 Pyruvate metabolism

Pyruvate has been shown to be metabolised anaerobically by *H. pylori* to yield lactate, ethanol and acetate whereas aerobically acetate is the main product from pyruvate metabolism (Chalk *et al.*, 1994). The true physiological conditions for the growth of *H. pylori* are microaerobic (5% oxygen). Under these conditions lactate, acetate, formate, succinate and alanine were produced from pyruvate metabolism (Mendz *et al.*, 1994b). The generation of succinate suggests an active Krebs cycle. The enzyme D-Lactate dehydrogenase which converts pyruvate to lactate has been identified within the genome of *H. pylori* (Tomb *et al.*, 1997). For the conversion of pyruvate into acetyl-CoA for entry into the Krebs cycle an aerobic pyruvate dehydrogenase or anaerobic pyruvate-formate lyase is absent. Instead a flavodoxin-dependent oxidoreductase found in obligate anaerobes catalyses the reversible reaction (shown in Figure 1.13). The pyruvate:acceptor oxidoreductase has been analysed from *H. pylori* and shown to be oxygen sensitive (Hughes *et al.*, 1998) with a flavodoxin:quinone reductase regenerating oxidised flavodoxin by reducing NADP (Maurice *et al.*, 2007).

1.8.3 The Krebs cycle

The Krebs cycle in *H. pylori* has been a classic example of how extrapolation from genomic analysis and limited biochemical knowledge can lead to miss-interpretation of a metabolic pathway. It was originally hypothesised and in some cases shown that *H. pylori* contained a non-cyclic pathway consisting of a dicarboxylic moiety: oxaloacetate being reduced to malate, fumarate and finally succinate and the tricarboxylic acid pathway. The tricarboxylic acid pathway proceeds oxidatively through oxaloacetate, citrate, isocitrate and finally α -ketoglutarate (Piston *et al.*, 1999). Apparent conflict with regards to the absence of α -ketoglutarate dehydrogenase but the presence of an α -ketoglutarate oxidase occurred (Piston *et al.*, 1999). The lack of other enzymes such as malate dehydrogenase from the genome but an apparent generation of oxaloacetate were suggestive of other "non-classical" enzymes fulfilling this role. The α -ketoglutarate:ferredoxin oxidoreductase complex of four heterogenous subunits was purified by Hughes *et al* (Hughes *et al.*, 1998) and shown to be oxygen sensitive. A succinyl-CoA:acetoacetyl-CoA transferase has been shown to convert succinyl-CoA to succinate (Corthesy-Theulaz *et al.*, 1997). With a malate:quinone oxidoreductase performing the role of malate dehydrogenase converting malate to oxaloacetate (Kather *et al.*, 2000). The microaerophilic fate of pyruvate and the complete Krebs cycle is shown in Figure 1.13 adapted from (Kather *et al.*, 2000). It is of interest to note that isocitrate dehydrogenase is an NADP⁺ dependent enzyme rather than NAD⁺ (Piston *et al.*, 1999) again showing the possibility that *H. pylori* uses NADPH as the main electron pool rather than NADH and this is discussed further in Chapter 7

It has been suggested that fumarate may act as a terminal electron acceptor in anaerobic respiration and the enzyme responsible for catalysing the reduction of fumarate to succinate (described in Chapter 1.8.4), fumarate reductase has been shown to be a therapeutic target against *H. pylori* (Mendz *et al.*, 1995b). Although kinetic measurements of the malate:quinone oxidoreductase suggesting that the enzyme does not reduce oxaloacetate to malate although another enzyme may carry out the reverse reaction (Kather *et al.*, 2000).

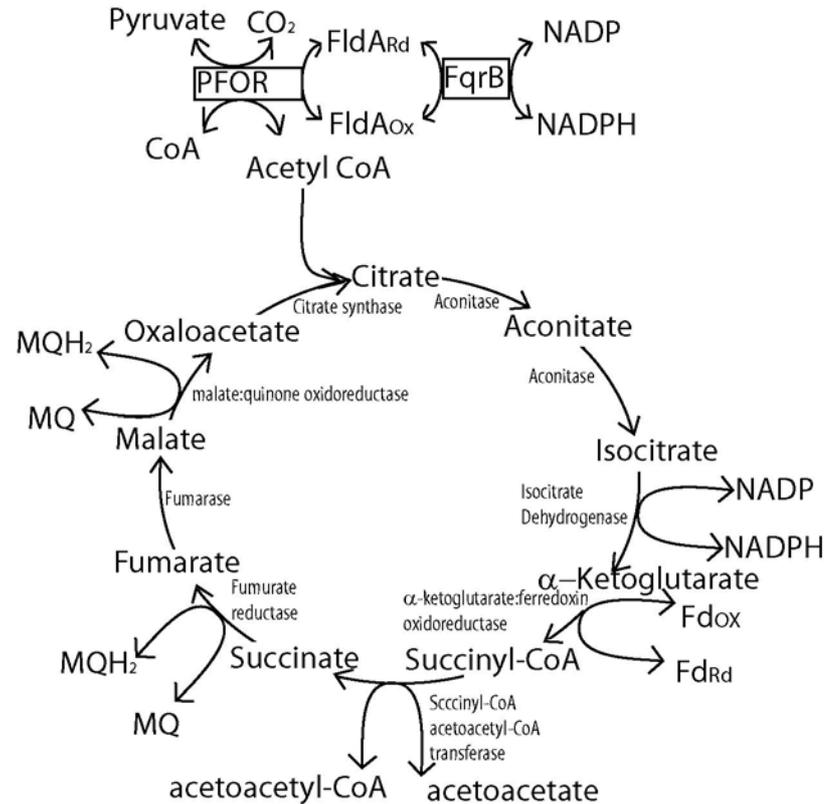


Figure 1. 13 The fate of pyruvate and the Krebs cycle in *H. pylori*. In microaerobic conditions pyruvate is oxidatively decarboxylated with CoA to give acetyl-CoA by the flavodoxin-containing pyruvate:acceptor oxidoreductase (PFOR) complex with the flavodoxin being regenerated by the flavodoxin:quinone reductase (FqrB) reducing NADP⁺. Acetyl-CoA enters the Krebs cycle and is condensed with oxaloacetate releasing CoA. Citrate is isomerised to isocitrate by aconitase in dehydration and subsequent hydration steps. The NADP⁺-dependent isocitrate dehydrogenase catalyses oxidative decarboxylation of isocitrate to α -ketoglutarate. Instead of a further oxidative decarboxylation of α -ketoglutarate by α -ketoglutarate dehydrogenase, α -ketoglutarate:ferredoxin oxidoreductase oxidatively decarboxylates α -ketoglutarate to succinyl-CoA yielding reduced ferredoxin probably re-oxidised by a NADP-dependent flavodoxin reductase. Succinyl-CoA acetoacetyl transferase is able to transfer the CoA group to acetoacetate, found in the gastric mucosa. Regeneration of acetoacetate from acetoacetyl-CoA proceeds by an unknown mechanism. Fumarate reductase oxidises succinate to fumarate and the final stage is the oxidation of malate to oxaloacetate by malate:quinone oxidoreductase. Figure adapted from (Maurice *et al.*, 2007) and (Kather *et al.*, 2000).

1.8.4 Respiratory chains

For the generation of ATP, membrane associated complexes are responsible for generating a proton gradient. Genomic analysis has indicated the presence of a NADH-quinone oxidoreductase 1 (NDH1) complex (Complex I) found in the respiratory chains of mitochondria and bacteria. However, in *H. pylori* there are no genes encoding the NADH dehydrogenase subunits NDQ1 and NDQ2 of NDH1 (Tomb *et al.*, 1997) although there is an FeS cluster in the NQO3 subunit of NDH1 which may serve as a docking site for another protein to feed the electrons into the complex (Finel, 1998). The NDH2 enzyme which is responsible for transfer of electrons from NADH into the respiratory chain is absent in *H. pylori* (Tomb *et al.*, 1997) suggesting that NADH is not the electron donating source in *H. pylori*, this is in contrast to other more classical pathways in bacteria and eukaryotic organisms. Within bacteria the respiratory chain is modular being comprised of dehydrogenase complexes, quinone pools and terminal oxidoreductases with oxygen as the terminal electron acceptor for aerobic respiration and other substrates such as dimethyl sulphoxide, sulphate, nitrate or trimethylamine N-oxide for anaerobic respiration (Marais *et al.*, 1999). The utilisation of other terminal electron acceptors in *H. pylori* is unlikely due to key genes being absent (Tomb *et al.*, 1997). Although this does not rule out anaerobic respiration in *H. pylori* since the presence of fumarate reductase activity (described below) and activity for an Ni-Fe hydrogenase which is activated in anaerobic conditions (Maier *et al.*, 1996) and bears close sequence similarity to the Ni-Fe hydrogenase identified in *Wolinella succinogenes*.

NADPH oxidation has been observed in membrane extracts of *H. pylori* with NADPH oxidation being six times more rapid than NADH oxidation (Chang *et al.*, 1995) with increased rates of NADPH instead of NADH oxidation being reported by other groups (Chen *et al.*, 1999, Hughes *et al.*, 1998). Although the exact mechanism of how electrons are delivered from NADPH into the respiratory chain and Complex I are unknown (Finel, 1998). Classical inhibitor analysis on cell extracts from *H. pylori* have shown the activity of Complex II (succinate dehydrogenase) with succinate cytochrome *c* reductase which links Complex II and III. Interestingly fumarate reductase activity has been shown to be more reducing than oxidising within *H. pylori*, which would lead to reduction of fumarate to succinate rather than its oxidation (Chen *et al.*, 1999).

1.8.5 Amino acid metabolism

Amino acids are key requirements for the cultivation of *H. pylori in vitro* with arginine, alanine, histidine, isoleucine, leucine, methionine, phenylalanine, serine and valine all required for growth (Reynolds and Penn, 1994). The dependence on these amino acids can be attributed to the lack of key anabolic enzymes in the genome (Alm *et al.*, 1999, Tomb *et al.*, 1997). Due to carbohydrates not being the main source for energy production and the synthesis of key metabolic intermediates, amino acid metabolism can provide a key alternative through feeding into the Krebs cycle with key metabolites produced from amino metabolism being identified as acetate, formate, lactate and succinate (Mendz *et al.*, 1993). It has been shown that growth medium comprising of no carbohydrates but amino acids is sufficient for growth of *H. pylori* (Mendz and Hazell, 1995) and when glucose is present, amino acids are metabolised first with glucose metabolism only occurring once the amino acids have been exhausted (Mendz *et al.*, 1993). The catabolism of arginine to urea and ornithine by an arginase has been investigated *in situ* (Mendz *et al.*, 1998). The amino acids serine, proline and alanine, are found in abundance within the human gut and *H. pylori* cells have been shown to metabolise these amino acids to pyruvate (Nagata *et al.*, 2003). In *C. jejuni* the oxygen-sensitive L-serine dehydratase has been shown to be a key enzyme in growth of the bacteria within poultry (Velayudhan *et al.*, 2004). Unpublished work undertaken during this study resulted in the cloning, expression and purification of L-serine dehydratase from *H. pylori*. Preliminary analysis showed L-serine dehydratase to be an oxygen-sensitive Fe-S containing enzyme presumably containing a 4Fe-4S cluster as in *C. jejuni*, which could be reconstituted under anaerobic conditions. Further work is continuing to elude the crystal structure of the 4Fe-4S bound enzyme (Elliott, Unpublished).

An interesting feature of *H. pylori* is the utilisation of D-amino acids, in particular D-alanine, which typically is used for peptidoglycan synthesis by bacteria. The abundance of D-alanine is particularly high compared to other components for peptidoglycan synthesis and D-alanine has been shown to be metabolised to components for the Krebs cycle (Nagata *et al.*, 2003). The racemisation of L-alanine to D-alanine occurs by an alanine racemase which has recently been characterised in *H. pylori*, favouring the racemisation of L- to D-alanine (Saito *et al.*, 2007).

1.9 Focus of current study

The preceding sections have outlined the bacteria comprising the *Campylobacterales* order, with particular emphasis on *H. pylori*. An overview has been presented as to how infection occurs and the disease outcomes associated with infection, through the action of virulence factors. The methods of detecting and treating the bacteria have been summarised as has the problems associated with current treatment, in particular, antibiotic resistance. The remainder of the chapter addressed the metabolism of *H. pylori* with emphasis on data derived from genomic and experimental analysis. Within carbohydrate metabolism, a key point for the convergence of several pathways was at glyceraldehyde-3-phosphate. Therefore, the focus has been the investigation of the oxidative phosphorylation of glyceraldehyde-3-phosphate by GAPDH. Within *H. pylori* there are two *gap* genes annotated as NAD⁺-dependent, termed *gapA* and *gapB*. The focus of this study has been the investigation of these two genes from *H. pylori* with the aims of using structural analysis to determine the mechanisms underlying substrate and coenzyme specificity. The ultimate aim of which is to identify whether there could be significant features of either enzyme, which could make them drug targets.

Characterisation of GAPDHA (Chapter 3) shows it to be NADP⁺-dependent, the coenzyme specificity underlying this is investigated in Chapter 3. The outcome of the investigation of holo GAPDHA leads to the identification of other putative NADP⁺-dependent GAPDHs within the *Campylobacterales* order. Of significance and interest is the GAPDH from *C. jejuni*, where Chapter 4 investigates the apparent dual coenzyme specificity. Chapter 5 investigates the second GAPDH in *H. pylori*, GAPDHB and addresses different substrate specificity, notably between glyceraldehyde-3-phosphate and erythrose-4-phosphate. Chapter 6 attempts to resolve the difference between substrate specificity observed in Chapter 5 and whilst the structures are unsuccessful in discriminating between the two substrates, further understanding of the enzyme mechanism is obtained.

Chapter 2. Materials and Methods

2.1. Materials

2.1.1. Chemicals and Reagents

All chemicals were of analytical grade where possible, except where stated and were purchased from Sigma™. Erythrose-4-phosphate was stipulated as being 60-75% pure and contained approximately 30% glyceraldehyde-3-phosphate contamination as confirmed by Sigma.

Molecular biology kits for extraction and purification of plasmid DNA and for gel-based DNA extraction and/or PCR purification were obtained from Qiagen™ as Mini-prep® and Gel extraction® respectively.

2.1.2. Chromatographic media and Membranes

Hi-Trap™ metal chelating columns and Hi-Trap™ SP Sepharose™ columns were purchased from GE Healthcare. The two main columns used for desalting were Hi-Prep™ 26/10 desalting column and PD-10 disposable chromatography column, both from GE Healthcare. Viva-spin™ columns for 20 ml, 6 ml and 500 µl sample volumes were supplied by Sartorius and had a molecular weight cut-off of 10 kDa. Dialysis tubing of molecular weight cut-off of 12-14 kDa was from Spectrum® Laboratories and was pre-treated by boiling in 2 mM EDTA then dH₂O prior to use.

2.1.3. Crystallisation accessories

Crystallisation plates used for sitting drop vapour diffusion were MRC plates purchased from Innovadyne Technologies, Inc, tape for sealing the crystallisation trays were obtained from Hampton Research Ltd, California, USA. Sparse matrix crystallisation screens (Jancarik and Kim, 1991) used were; Wizard™ I and II and Cryo I and II from Emerald BioSystems, Washington, USA. For cryocrystallography crystals were mounted in Litho™ loops from Molecular Dimensions Ltd. Prior to data collection at the ESRF, Grenoble, France, frozen crystals were transferred into pucks for the sample chamber at the ESRF, with all magnetic caps and vials being SPINE standard and purchased from Hampton Research Ltd.

2.2. Molecular Biology Methods

2.2.1. Primer design

As all ORFs (open reading frames) were cloned into the TOPO[®] pET151/D vector for protein over expression, the primers were designed in accordance to the manufacture's instructions. All forward primers had to start with the **cacc** sequence at the 5' end of the primer, this was to ensure the topoisomerase reaction worked. The remainder of the primer was designed to provide a suitable annealing length complementary to the two strands of the target DNA, with a melting temperature (T_m) between 50-58 °C and a T_m difference between the forward and reverse primers no more than 2 °C to allow efficient PCR. The design of primers for mutagenesis is discussed in Chapter 2.2.5.

Construct	Forward Primer	Reverse Primer
GAPDHA	CACCATGCCAATTAGAAT	TATATAGCACAAAATTAA
GAPDHB	CACCATGAAAATTTTATCATTGGATTG	TTAATAATGATACATAACTGG
cGAPDH	CACCATGGCTGTAAAAGTTGCTATAAATGG	TTAAGCCTTATTTGCAATATATACTGC
GAPDHA C149A	GCCATTGTTTCTAACGCCTCTTCCACGA	GATAGCGTTAGTCGTGGCAGAGGCGTT
GAPDHA C149S	GCCATTGTTTCTAACGCCTCTGCCACGA	GATAGCGTTAGTCGTGGAAGAGGCGTT
GAPDHB C151A	CTAACGCCTCTTCTACGACTAACGCTATCG	GTCGAAGAGGCGTTAGAAACAATGGCT
GAPDHB C151S	CTAACGCCTCTGCGACGACTAACGCTATCG	GTCGCAGAGGCGTTAGAAACAATGGCT

Table 2. 1 Forward and reverse primer sequences for the generation of all the constructs used within this study.

2.2.2. GAPDH ORF amplification by PCR

The Polymerase Chain Reaction (PCR) was used to amplify the open reading frame (ORF) of the corresponding GAPDH genes. The PCR was carried out using Kod DNA Polymerase™ from Novagen, using a modification of the manufacture's protocol. Primers were purchased from Invitrogen as lyophilised and desalted. Prior to PCR each primer was diluted to 90 μM in TE (40 mM Tris-EDTA pH 8.0) and then diluted to a 30 μM working stock in miliQ water (Millipore). Briefly, the components were set up in order on ice with the amounts shown in Table 2.2

	Control	Sample
Q (μl)	33.1	32.6
10 x Buffer I/II (μl)	5	5
2 mM each dNTPs (μl)	5	5
25 mM MgCl_2 (μl)	2.5	2.5
30 μM Forward Primer (μl)	2	2
30 μM Reverse primer (μl)	2	2
DNA template (μl)	-	0.5
Kod DNA Polymerase (μl)	0.4	0.4

Table 2. 2 Components for amplifying the target ORF. The above solutions were added in order on ice in thin-walled 200 μl PCR tubes. Amplification was carried out in a Mastercycler® personal thermocycler, Eppendorf in accordance with the manufacture's conditions. For all reactions, the primer annealing temperature was 50 °C, with an extension time of 40 seconds.

After the PCR reaction, the samples were returned to ice with 5x loading buffer added (final concentration of 25 mM Tris pH 7.5, 20 mM NaCl, 30% glycerol, 0.1% Orange G). DNA was separated based upon size by agrose gel electrophoresis. Typically, a 1% agrose gel was poured containing 8 μgml^{-1} ethidium bromide and transferred to a gel tank containing 1x Tris acetate EDTA (TAE) solution (40 mM Tris-acetate, 1 mM EDTA). Samples were run against a ladder of specific, varying, DNA sizes in the 100 bp or 1 kbp range (New England Biolabs). DNA bands were visualised by UV-light on a transilluminator. DNA bands of the appropriate size were cut out from the gel with the DNA extracted using a Gel extraction® kit (Qiagen) and eluted and stored in 5 mM Tris pH 8.0 prior to use.

2.2.3. TOPO isomerase cloning

The method of choice for cloning the GAPDH ORFs into a suitable expression vector was using the Champion™ TOPO® cloning kit from Invitrogen. The vector of choice for expression of the target ORF was pET151/D (Invitrogen) containing the T7 bacteriophage promoter for high, inducible expression and an N-terminal hexa-histidine affinity tag and TEV (tobacco etch virus) cleavage site, as shown in Figure 2.1.

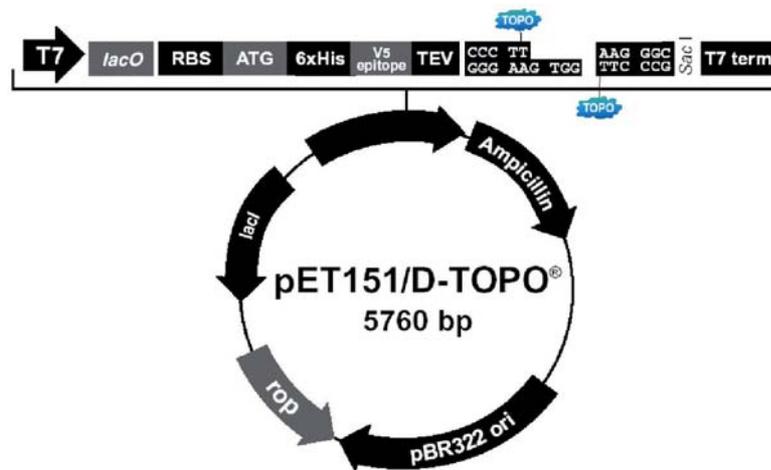


Figure 2. 1 Vector map of the pET151/D TOPO® vector for cloning of GAPDH ORFs prior to protein expression.

The TOPO® cloning strategy utilises pre-linearised vector containing topoisomerase at the 5' and 3' ends of the vector. The enzyme recognises the cacc sequence of the amplified PCR ORF and performs blunt end independent ligation similar to the method by Shuman, (Shuman, 1994). A modification of this method allows greater than 90% efficiency of cloning (Invitrogen).

The reaction between the purified PCR product and 10 ng of vector is followed according to the manufacture's instructions, with half of the reaction mixture transformed into chemically competent DH5a *E. coli* and selected against the appropriate antibiotic.

2.2.4. Transformation of *E. coli* strains

Chemically competent *E. coli* (previously prepared by the rubidium chloride method) were removed from storage at -80 °C and thawed on ice. Either 4 µl of cloning mix from the TOPO[®] method (Chapter 2.2.3) or 1 µl of purified plasmid DNA was added to 50-100 µl of competent cells and incubated on ice for 30 minutes. The samples were placed in a 42 °C water bath for exactly 30 sec to heat shock the cells and then placed immediately on ice for 2 minutes. To each sample 450 µl S.O.C. medium was added (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose per litre of water) and incubated with shaking at 37 °C for one hour. Selection for successful transformants is permitted by plating the cells onto LB-agar (5 g tryptone, 2.5 g yeast extract, 5 g NaCl, 7.5 g bacteriological agar per half litre) containing the appropriate antibiotic (50 µgml⁻¹ ampicillin/carbenicillin and/or 34 µgml⁻¹ chloramphenicol) and incubated overnight at 37 °C. Plates containing colonies were stored at 4 °C, typically no longer than two weeks, with all protein expression experiments undergoing a new transformation prior to expression.

2.2.5. Colony PCR screening

Thin walled PCR tubes were set up on ice with all the components from Table 2.1 added at half the volumes shown, with the exception of the miliQ, which was omitted from all tubes. Alongside this were duplicate tubes containing 50 µl of miliQ. For each PCR reaction a single colony was picked from the transformed topoisomerase reaction described above with a sterile 200 µl tip and resuspended in the miliQ water. Next 16.6 µl of the miliQ is taken (with the same tip) and dispensed into the thin walled PCR tubes and the tip is stabbed into a fresh LB-agar plate containing 50 µgml⁻¹ ampicillin before being discarded. During the PCR reaction 500 µl S.O.C. medium was added to the eppendorfs containing the remainder of the miliQ and the tubes are returned to 37 °C for the duration of the PCR reaction. After the amplification had finished 5 x loading buffer was added and successfulness of the reaction analysed by agarose gel electrophoresis. Colonies which contained the pET151/D-ORF would produce a positive PCR reaction of the appropriate size of the ORF. The corresponding eppendorf containing the S.O.C. medium can then be used as a starter culture for plasmid preparation the following day.

The above method of diluting the colony into miliQ prior to the PCR reaction has the advantage of allowing for a starter culture to be set-up for plasmid preparation the following day, speeding up the whole process by a day. In addition, it increases the successfulness of the PCR reaction owing to agar being diluted, which can inhibit the PCR reaction (Personal communication with the Oxford Protein Production Facility).

2.2.6. Site-directed mutagenesis

The standard approach of amplifying the entire vector by PCR and then digesting the parental vector with DpnI before transformation is not used as the method of first choice for mutagenesis. Instead primers were designed as for any mutagenic reaction, i.e. a 6-15 bp overlap between the 3' end of the forward primer and the 3' end of the reverse primer. Next, a PCR reaction is performed using the ORF forward and reverse cloning primers, with the forward ORF primer being used with the mutant reverse primer and vice versa in two separate PCR reactions ran in parallel to each other. The PCR reaction is analysed by gel electrophoresis and this confirms if the mutant primers are successful. Next, the PCR DNA is extracted and used to set up another PCR reaction containing forward and reverse ORF primers. During this reaction, the two PCR products anneal due to their 6-15 bp homology region and an overhanging PCR is performed, yielding one PCR product, the sum of the two individual PCR reactions. This method involves another cloning step, which is avoided by the DpnI approach. However, this method is directional because the first PCR reactions would fail if there was a problem with the mutant primer, therefore the mutation is incorporated within the first PCR and is therefore directional. A schematic of this method is shown in Figure 2.2.

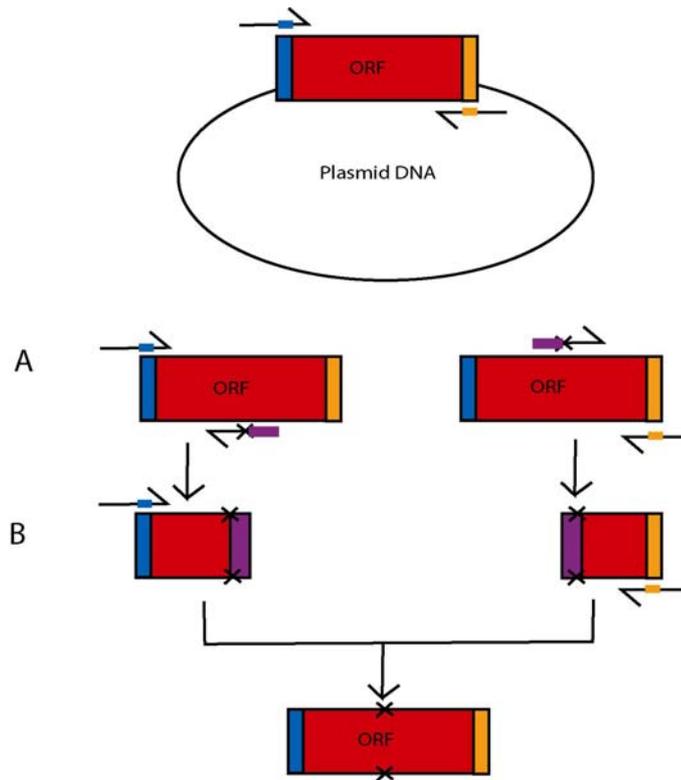


Figure 2. 2 Mutagenic PCR-based approach using overhanging PCR. In the first reaction two PCR reactions are ran in parallel, with amplification of the ORF using the ORF forward primer and mutant reverse primer and vice versa (reaction A). The PCR products are analysed by gel electrophoresis to confirm the reaction has been successful and the PCR products isolated and mixed together in another PCR reaction using ORF forward and reverse primers (reaction B), yielding a mutant ORF PCR-product which can be re-cloned into the expression vector.

2.3. Protein Methods

2.3.1. SDS Polyacrylamide Gel Electrophoresis

SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE) was carried out using gels poured for use in the Bio-Rad Mini-PROTEAN™ III system. Typically 12% acrylamide gels were poured and consisted of resolving gel: 375 mM Tris pH 8.8, 0.1%w/v SDS, 12 %v/v acrylamide, 0.1 %w/v ammonium persulphate, 0.0004%v/v TEMED and stacking gel: 125 mM Tris pH 6.8, 0.1%w/v SDS, 4%v/v acrylamide, 0.02%w/v ammonium persulphate, 0.0002%v/v TEMED. Samples were prepared prior to electrophoresis by adding sample buffer (2x buffer consists of: 200 mM Tris pH 6.8, 20%v/v glycerol, 4%w/v SDS, 0.05%w/v bromophenol blue and 5%v/v β -mercaptoethanol) and heating to 95 °C for 10 minutes before loading. Samples were loaded onto the gel against a protein ladder consisting of proteins of known molecular weight (NEB) and ran at 190 V for approximately 40 min in SDS-PAGE running buffer (10x SDS-running buffer: 3%w/v Tris base, 14%w/v glycine, 1%w/v SDS). Gels were stained using 0.4%w/v Coomassie Brilliant Blue R250, 50%v/v methanol, 10%v/v acetic acid for 5 min after warming in the microwave. Gels were subsequently destained in 10%v/v acetic acid for 30 min and transferred to dH₂O.

2.4. Crystallographic methods

The following section will only deal with the crystallographic methods and principles applied during this study. Several textbooks provide an excellent introduction to the theory of protein crystallography and a complete overview of the methods required to solve the phase problem, other than molecular replacement which is covered in this section. The reader is referred to the books by Blundell and Johnson (Blundell and Johnson, 1976) and Sherwood (Sherwood, 1976) for a thorough introduction to the theory and mathematics of protein crystallography.

2.4.1. Protein Crystallisation

For the determination of the three-dimensional structure of proteins (and small molecules) through diffraction analysis, crystals have to be obtained. Scattering of X-rays occurs at an atomic level through the interaction of the incident X-ray beam with the electrons orbiting an atom. This scattering is weak with the majority of the incident X-rays passing through the atom. If the scattering centres are arranged in an ordered, repeating manner then the cumulative effect of scattering will be strong enough to be measured, this is achieved through the crystalline state. The crystal as a whole diffracts the X-rays.

The process of crystallising proteins for structure determination has been intensely studied over the last 60 years. The process of protein crystallisation is an immense field in its own right and will only be covered briefly, the reader is referred to (McPherson, 1998) for a thorough description of the subject.

The unique chemical nature of proteins makes them labile with complex interactions existing between other protein molecules and the protein with its solvent environment. These complex interactions are as a result of the proteins three dimensional structure and can not be predicted. Therefore, the challenge to the crystallographer is how to preterb the equilibrium of the protein within its soluble environment, to produce an ordered crystalline state, without causing the protein to enter a state of amorphous precipitation. Conceptually, an infinite number of interrelated factors may affect the thermodynamic and kinetic nature of crystallisation. Several main variables are often explored in crystallisation; temperature, protein concentration, pH of precipitating solution and precipitant. Different types of precipitants have different contributions to affecting the hydration of the protein in solution, promoting nucleation and crystal growth. For example, salts either compete for solvent on the surface of the protein "salting-out" or protect specific charged groups "salting-in". Other precipitants reduce the dielectric constant of the solution, (such as polyethylene glycols (PEGs)) increasing the distance over which protein-protein interactions occur. Organic solvents bind to water molecules also effectively reducing the dielectric constant.

The purpose of the crystallisation experiment is to slowly reduce the solubility of the protein whilst retaining its structural integrity. The process of crystallisation can be described with a phase diagram, as shown in Figure 2.3. The phase diagram in Figure 2.3 is that for a vapour diffusion experiment. where the protein exists either in solution (A); a metastable zone (C), which is supersaturated but no nucleation occurs however, crystal growth is supported; a

nucleation point where protein molecules start to aggregate in a crystallised form (B); and a precipitation zone where protein separates from solution as an amorphous solid. There are a variety of techniques, which favour a particular transition through the phase diagram, such as batch, dialysis, counter diffusion and vapour diffusion. All these methods have their own advantages and disadvantages (again for a thorough description the reader is referred to (McPherson, 1998) and references therein).

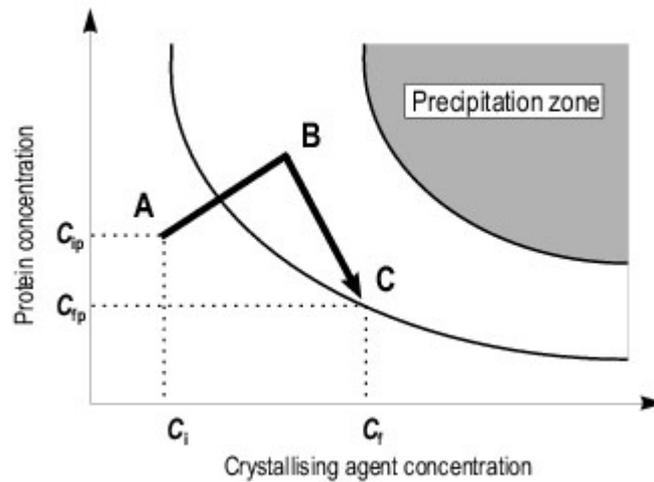


Figure 2. 3 Phase diagram for the vapour-diffusion method of crystallisation. At the start of the experiment the protein exists in the soluble state (position A). As the concentration of protein and precipitant increases through vapour diffusion the protein enters the supersaturation zone. Here nucleation can occur (B) and crystal growth may be permitted through the metastable zone (C) where the protein concentration of the solution will decrease. This figure is taken from www-cryst.bioc.cam.ac.uk/.../node3.html.

In all situations when one is provided with a new protein to crystallise the conditions, which favour crystallisation are not known. An approach of systematically altering each possible variable e.g. pH, precipitant concentration, protein concentration, temperature etc for a range of different precipitants would lead to a combinatorial explosion. Therefore, rather than sample all the matrix of crystallisation space, random conditions are sampled which can lead to initial "hits" (i.e. where potential for growing single crystals is seen). These conditions can then be refined systematically. The initial search is known as incomplete factorial screening (Carter and Carter, 1979). The method of incomplete factorial screening has evolved into sparse matrix screening (Jancarik and Kim, 1991). Sparse matrix screening applies some "prior knowledge" by using conditions that have given crystals for other proteins, therefore

removing conditions, which are less likely to give initial crystallisation hits. Although one can never predict how a protein will behave on attempting to crystallise it, careful examination of its behaviour before and during the crystallisation process can maximise the chances of success.

2.4.1.1. Robotics in Protein crystallography

The use of robotics has significantly improved the process of finding and refining crystallisation conditions, the reasons for this are two-fold. Firstly, the volumes required are much reduced and secondly the reproducibility is improved. Furthermore, the combination of these with elimination of operator fatigue (and error) allows many more conditions to be investigated reliably. In the case of proteins expressed in mammalian or insect cells, or isolated from source organisms, the amount of protein available may be very limited. The use of nano-drop dispensing robotics allows for sitting and hanging drop crystallisation experiments to be prepared in a 96-well format within a matter of minutes. There are a variety of nano-drop dispensing crystallisation robots on the market, the one used in this study was the Genomics Solutions Cartesian Honeybee (8 + 1) (Harvard Bioscience) incorporating a humidified chamber. Plate storage with integrated imaging systems allows the visualisation and record keeping of each drop, providing one with the ability to monitor remotely the progress of the crystallisation experiment. The hotel-imaging system used in this study was the CrystalProHT imaging system (TriTek). Details of the protocols used to grow X-ray quality crystals are described in the subsequent chapters.

2.4.2. Description of X-rays

X-rays are electromagnetic waves with a wavelength (λ) between 0.01 and 1 nm. The interaction of a particle with X-rays will depend upon the size of the particle with respect to the wavelength. A particle, which is larger than the wavelength, will experience a negligible effect from the incident wave, whilst a particle, which is smaller than the wavelength, will experience an acceleration through its interaction with the electronic component of the wave. The effect of this force is oscillation of the particle causing a source of a new electromagnetic wave (assumed of exact frequency to the incident wave). The wavelength of X-rays from $\text{Cu}\alpha$ is 0.15 nm (1.5 Å) whilst the size of a water molecule is ~ 4 Å. Therefore, the contribution of scattering is from protons and electrons (neutrons have no charge and so cannot interact with the X-rays). The intensity of scattering is proportional to the square of the charge to mass ratio. The mass of a proton is 1837 times greater than an electron;

therefore, the intensity ratio between electron and proton is 1837^2 . Hence, an electron scatters x-rays some 1837^2 times more efficient than protons. Thus, X-ray diffraction is the measurement of scattering from electrons comprising molecules. The greater the number of electrons then the greater scattering.

A three-dimensional wave propagating from a single source can be described by the following equation;

$$\Psi(\mathbf{r}, t) = \Psi_0 \cos(\mathbf{k} \cdot \mathbf{r} - \omega t) \quad (\text{equation 2.1})$$

Where, ψ is the amplitude of the wave, ω is the frequency and \mathbf{k} is the wave vector for positions x , y and z (\mathbf{r}) in space.

2.4.3. The Fourier Transform

The Fourier transform $F(\mathbf{k})$ of any function of $f(\mathbf{x})$ can be calculated. A significant feature of the Fourier transform is its relationship to the Inverse Fourier transform, for example;

$$F(\mathbf{k}) = \int f(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} d\mathbf{r} \quad (\text{equation 2.2})$$

and the inverse transform

$$f(\mathbf{r}) = \int F(\mathbf{k}) e^{-i\mathbf{k} \cdot \mathbf{r}} d\mathbf{k} \quad (\text{equation 2.3})$$

The Fourier transform in equation 2.2 represents the three dimensional wave shown in equation 2.1. Therefore, the Fourier transform is a convenient method to deconvolute any number of different waves added together. The function $f(\mathbf{r})$ is known as the amplitude function and will contain information about the scattering centre. The Fast Fourier Transform (FFT) is used to speed up the process of calculating electron density from structure factors (and vice versa) (Ten Eyck, 1985).

2.4.4. Cryo cooling

Prior to data collection crystals are mounted in the appropriately sized loop before being cryo cooled. The process of cryo cooling, as reviewed in (Garman and Schneider, 1997) into a stream of boiled off liquid nitrogen at 100 K or submerging the mounted crystal directly into liquid nitrogen has two main advantages. Firstly, a reduction in radiation damage caused primarily by radicals generated from water molecules within the crystal, which then attack the protein. Other effects, such as local heating and direct radiation damage through the emission of electrons from molecules initiating chemical reactions (for example deamination of side chains) also occur. The second advantage of cryo-cooling is the reduction in internal movement, which can lead to an improvement in resolution. Different cryoprotectants have to be found which firstly permit the vitrification (formation of a non-ordered glass-like phase) of water and secondly does not damage the crystal as a result of osmotic shock. This process is normally carried out by screening a range of potential solutions (as described for holo GAPDHA crystals in Chapter 3.5) although some systematic methods may be applied, reviewed in (Garman and Schneider, 1997).

2.4.5. Diffraction data collection and processing

Once a series of images for different oscillation angles have been collected (typically at two different values of φ) then an auto-indexing routine can be applied to give an estimation of the cell parameters and the crystal symmetry. This process was carried out using the routines in *MOSFLM* (Powell, 1999), which uses one dimensional FFT. The auto-indexing is transformed onto each of the characteristic 44 lattices with a penalty applied for each transformation. From the Laue symmetry the amount of data required to fully sample the reciprocal space will be known. *MOSFLM* (Leslie, 1992) uses the program *STRATEGY* where from the crystal orientation an optimum data set can be collected. Data is collected using a small oscillation range (typically 0.2-1 °) which avoids overlaps. Once a complete data set has been collected then the cell parameters initially provided from the auto-indexing routine and the beam position and divergence are refined in *MOSFLM* (Leslie, 1992). Finally, once the cell parameters and the beam position are refined the intensities are added together and subtracted from the background. Using small oscillation ranges results in partial reflections being recorded. *MOSFLM* (Leslie, 1992) relies on two-dimensional integration whereby partial reflections are measured over several images (2-3) resulting in fully recorded spots

upon scaling and merging. Other programs such as *XDS* (Kabsch, 1988) rely on three-dimensional integration which includes information as to how the spot profile changes. In all cases a mask is applied to define the spot area and profile fitting (typically over 10-20 images) is used to integrate all the pixels and write out a single file containing the intensities as an mtz file (McClauglan Terry Zelinka).

The unmerged intensities written out after integration in *MOSFLM* (Leslie, 1992) will have intensities on different scales as a result of variation of beam intensity during data collection or the volume of the crystal being hit by the X-ray beam. Other factors such as radiation damage and absorption of the X-rays also will affect the intensities. Therefore, the reflections from different images will need to be scaled and this is achieved through the program *SCALA* (Evans, 2006). Comparison of symmetry-related reflections allows the determination of the scales to be applied. The agreement between scaled symmetry related reflections is tabulated and thus *SCALA* (Evans, 2006) can give information as to whether the correct Laue group is being used. Next, the intensities are converted to structure factor amplitudes, F . For a perfect system the structure factor amplitude is related to the square root of the intensities, $|F| = \sqrt{I}$. However, for non-perfect data this is not the case so the program *TRUNCATE* (French and Wilson, 1978) estimates the closest approximation of F based upon the probability distribution of intensities.

2.4.6. Scattering of X-rays by an atom

Two types of scattering can occur, coherent (Thomson) scattering or incoherent (Compton) scattering. During Compton scattering, the X-ray interacts with a loosely bound electron and is deflected after transferring some energy to the electron. The difference in energy of the incident and scattered X-ray means that the scattered X-ray has a different wavelength. During Thomson scattering however, the incident X-ray causes the electron to oscillate at the frequency of the X-ray and therefore, become a scattering centre with no transfer of energy, rather the scattered X-ray has a different phase.

Upon considering a one atom system, shown in Figure 2.4, an incident X-ray, \mathbf{s}_0 hits an electron, located at position (P), a distance of \mathbf{r} from an arbitrary origin, O. The scattered X-ray, \mathbf{s} , has a different path to the incident X-ray and the total wave scattered from points O and P will of course depend on the phase difference between the two waves.

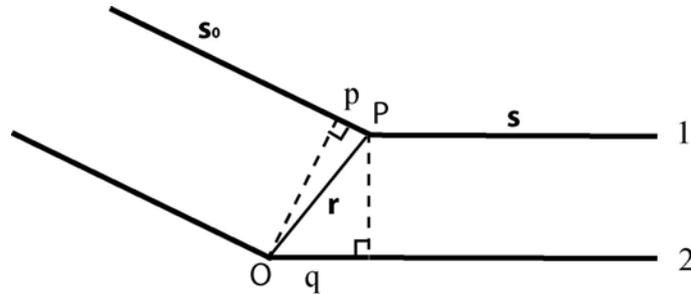


Figure 2. 4 Scattering of the incident X-ray (\mathbf{s}_0) at a point (P), relative to an origin (O). Taken from (Blundell and Johnson, 1976).

The path difference between the two waves (1 and 2) is given by $d = p-q$.

Geometrical analysis states that; $d = r \sin \beta - r \sin \alpha$

Therefore, the phase difference, $\Delta\phi$, is;

$$\Delta\phi = (2\pi/\lambda)r(\sin\beta - \sin\alpha) \quad (\text{equation 2.4})$$

As the wavelength of the X-ray is unchanged upon diffraction, then $\mathbf{s}_0 = \mathbf{s} = 2\pi/\lambda$.

In addition, $\mathbf{s}_0 \cdot \mathbf{r} = 2\pi/\lambda r \sin\alpha$ and $\mathbf{s} \cdot \mathbf{r} = 2\pi/\lambda r \sin\beta$.

The total wave scattered from the atom can now be expressed as

$$F(\mathbf{S}) = \int_{\mathbf{r}} \rho(\mathbf{r}) e^{i2\pi\mathbf{S} \cdot \mathbf{r}} d\mathbf{r} \quad (\text{equation 2.5})$$

Where $\rho(\mathbf{r})$ is the electron density, at a distance \mathbf{r} , from the centre of the atom. It is assumed that the electron density is centrosymmetric, in that $\rho(\mathbf{r}) = -\rho(-\mathbf{r})$. Hence, the imaginary component cancels out giving a real summation. The atomic scattering factor, f , can thus be defined:

$$f = \int_{\mathbf{r}} \rho(\mathbf{r}) e^{i2\pi\mathbf{S} \cdot \mathbf{r}} d\mathbf{r} \quad (\text{equation 2.6})$$

2.4.7. Scattering of X-rays by a crystal

For any atom at a new position within the molecule, \mathbf{r}_1 , its distance from the origin will be $\mathbf{r} + \mathbf{r}_1$. The scattering by the atom relative to the new origin will be

$$\mathbf{f}_1 = \int_V \rho(\mathbf{r}) e^{i2\pi\mathbf{S} \cdot (\mathbf{r} + \mathbf{r}_1)} d\mathbf{r} \quad (\text{equation 2.7})$$

The total wave scattered by all the atoms in the unit cell is a summation of all the vector components i.e.

$$\mathbf{F}(\mathbf{S}) = \sum_{j=1}^N f_j e^{2\pi i \mathbf{r}_j \cdot \mathbf{S}} \quad (\text{equation 2.8})$$

The $\mathbf{F}(\mathbf{S})$ function in equation 2.8 is known as the structure factor, since it is the sum of the scattering by atoms within a unit cell. Scattering is only observed once the phase difference between the waves scattered by successive unit cells is equal to an integral of 2π . This allows the waves to add up constructively to form a significant scattered wave. This gives rise to the Laue equations, shown in equation 2.9

$$\begin{aligned} \mathbf{a} \cdot \mathbf{S} &= h \\ \mathbf{b} \cdot \mathbf{S} &= k \\ \mathbf{c} \cdot \mathbf{S} &= l \end{aligned} \quad (\text{equation 2.9})$$

These equations can be used to rewrite the structure factor equation to represent the molecular transform sampled at the reciprocal lattice points hkl .

$$\mathbf{F}(hkl) = \sum_{j=1}^N f_j e^{2\pi i(hx_j + ky_j + lz_j)} \quad (\text{equation 2.10})$$

The structure factor is a complex number which can be represented by its amplitude ($F(hkl)$) and phase ($\alpha(hkl)$).

$$\mathbf{F}(hkl) = F(hkl) e^{i\alpha(hkl)} \quad (\text{equation 2.11})$$

The real and imaginary components can be represented in an Argand diagram.

2.4.8. The electron density equation

The equation for the calculation of the structure factor (shown in equation 2.11), can be rewritten for a continuous summation over the volume of the unit cell.

$$\mathbf{F}(\mathbf{S}) = \sum_{j=1}^N f_j e^{2\pi i \mathbf{r}_j \cdot \mathbf{S}} \quad (\text{equation 2.12})$$

$$= \int_{\text{Vol. cell}} \rho(\mathbf{r}) e^{-2\pi i \mathbf{r} \cdot \mathbf{S}} dV \quad \text{equation 2.13}$$

Integrating over the entire volume of diffraction space and removal of the integral since $F(\mathbf{S})$ is non continuous and is non-zero only at the reciprocal lattice points, yields equation 2.14.

$$\rho(xyz) = 1/V \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \mathbf{F}(\mathbf{hkl}) e^{-2\pi i(hx + ky + lz)} \quad \text{equation 2.14}$$

Therefore, if the structure factors are known for all reflections h , k and l then the electron density may be calculated. However, as was shown by equation 2.14, the structure factor is a complex number.

2.4.9. The Phase problem

As discussed previously the structure factor is composed of a magnitude ($F_{\mathbf{hkl}}$) and phase ($\alpha_{\mathbf{hkl}}$) and during an X-ray diffraction experiment only the intensities ($F_{\mathbf{hkl}}^2$) are measured, therefore the information on the phase is lost

Several methods exist to derive the phases from an X-ray diffraction experiment. These are: anomalous scattering methods (namely MAD and SAD), isomorphous replacement, *Ab initio* phasing and molecular replacement. Molecular replacement was the only method used to solve the phase problem in this study, the reader is referred to (Blundell and Johnson, 1976) and (Sherwood, 1976) for a thorough discussion of the other methods.

2.4.10. Molecular Replacement

Structurally homologous proteins deposited in databases allow for the determination of the phases based on these models. Six parameters are required to determine a molecule's position and orientation, they are three rotation angles and three translations. Traditionally, this sort of search would be very exhaustive computationally. To maximise efficiency the search is split into two parts, a rotation search then a translation search. As reviewed in (Evans and McCoy, 2008) after the rotation search the correct calculated structure factor (F_c) can not be calculated due to an unknown translation function and so it can not be calculated to the F_o to give an estimation of the validity of the model. This problem is overcome by searching for rotations in Patterson space or the use of Maximum-likelihood as used by the program *PHASER* (Read, 2001). Although the algorithms in *PHASER* (Read, 2001) do not use the Patterson function, for comparison an overview will be given.

Traditionally for small molecule crystallography, the Patterson method was able to determine the phases from the experimental intensities (Patterson, 1934). Effectively the Patterson function is the Fourier transform of the squared intensities with the phases not required. The Patterson map represents interatomic vectors rather than individual atoms, although this is complicated with self-vectors appearing as a result of crystal symmetry. The crowding within a Patterson map makes it impossible to resolve the peaks required for the structure determination of large molecules (typically 20 atoms and above). However, rotating a model will rotate the intermolecular vectors by the same angle and so it is possible to calculate a rotation function even though the translation function of the molecule is not known.

Maximum Likelihood is a statistical approach, which operates within reciprocal space. The probability functions for the structure factors calculated in reciprocal space are represented as Gaussian functions (Bricogne, 1997). Two types of fast rotation functions are used within *PHASER* (Read, 2001), the Wilson and the Sim Maximum Likelihood rotation functions. The model is rotated sequentially on a grid, with at each orientation the lengths of the structure factors are known as well as its symmetry-related orientations, but the relative phases are not known. Therefore, although the total structure factor can not be calculated a probability function can be applied to determine which structure factors are more significant than others and hence to take through for the translation function. The translational search is again applied along a grid in reciprocal space, although at each point the amplitude and phase are known allowing the total structure factor to be calculated. Again, this is along a Gaussian

distribution since the actual position of the model is still not yet known. To find the true maximum probability for the position of the model after rotational and translational searching, the model is rigid body refined. To ensure no models which have a high Likelihood gain form unreasonable overlap and clash, a packing function is applied.

2.4.11. Model building

The generation of the output files from molecular replacement results in the coordinate file (.pdb) of the search model rotated and translated onto the position of the model within the asymmetric unit. The calculated phases and structure factor amplitudes are also written out from molecular replacement in the binary format as an mtz file. If the search model had different residues then analysis of the electron density would reveal an inconsistency between the model and the electron density within the $2F_o - F_c$ map. In addition, due to the refinement within *PHASER* (Read, 2001) a $m2F_o - DF_c$ and a $mF_o - F_c$ weighted difference map will be generated. The weighting terms are derived from sigmaA (Read, 1986) and reduces model bias during refinement. Using programs like *COOT* (Emsley and Cowtan, 2004) enable the visualisation and manipulation of the model within the calculated electron density. Within the program are libraries where the geometry of the molecule (i.e. bond lengths, angles), which allow the refinement of residues and molecules within the electron density. The validity of the structure based upon torsional angles, rotamers and phi and psi angles is made available for the user enabling a model to be finely adjusted during the refinement process. This is particularly important especially when working with molecular replacement due to model bias.

2.4.12. Refinement

The aim of refinement is to adjust the model to find a closer agreement between the calculated and observed structure factors. The process of refinement is monitored by the R factor which allows the subsequent changes in the model to be monitored.

$$R = \frac{\sum_{hkl} ||F_o(h, k, l) - |F_c(h, k, l)||}{\sum_{hkl} |F_o(h, k, l)|} \quad (\text{equation 2.15})$$

A random selection of 5 % of the data prior to refinement is used as a measure to ensure the model is not being over refined, this leads to the generation of an R_{free} factor. The R_{free} is calculated the same way as in equation 2.15 for R, but a random selection of the observed reflections is used.

Various restraints are applied during refinement including that of bond lengths and angles, torsional angles and van der Waals contacts. During refinement it is important to include all the observations, this is particularly important when refining the atomic temperature factors. If the temperature factors are refined with an isotropic weighting then there are three positional parameters and this would have to be reflected in the number of observations. The program *REFMAC5* (Murshudov *et al.*, 1997) has been used in this study to refine the models and relies on maximum likelihood (mentioned in the previous section).

2.4.13. Structure validation

Errors in the model building and refinement processes are easily introduced particularly at low resolution and after using molecular replacement. Most obvious sources of error include the incorrect modelling of electron density. The R and R_{free} factors provide a good indication as to how well the model fits within the electron density. In addition stereochemical checks of the model are used, typically the Ramachandran plot (Ramachandran and Sasisekharan, 1968) in which the dihedral ϕ and ψ angles are plotted for each residue. Other stereochemistry checks are carried out using *PROCHECK* (Laskowski *et al.*, 1993) where in addition to the Ramachandran plot the planarity, bond lengths and angles are analysed against ideal values determined from peptide structures. In addition to *PROCHECK* (Laskowski *et al.*, 1993) the server MolProbity (Davies *et al.*, 2007) was used to analyse any areas of the model which were in disagreement with the above parameters.

2.4.13.1. Omit maps

To determine if regions of the electron density are not a result of over refinement, an omit map can be calculated. In an omit map part of the model is deleted and so does not contribute to the calculation of the phases. Therefore if the fragment which has been removed is real, it will come back at half the height compared to the remainder of the model. During this study composite omit maps were calculated for determining whether substrates were bound within the active site. In a composite omit map simulated annealing is applied to the remainder of the structure.

2.5. Enzymatic analysis

For the measurement of enzymatic activities for the oxidative phosphorylation reaction of the GAPDHs, enzymatic analysis was performed in accordance to the assay developed by Racker (Racker and Krimsky, 1952) and the modified form by (Trentham, 1971). Since the rate determining step is coenzyme exchange and this is dependent on high pHs, the reaction is performed at pH 8.6. The reaction buffer consisted of: 40 mM triethanolamine pH 8.6, 2 mM EDTA. To prevent product inhibition 20 mM arsenate was added to the buffer in accordance with (Trentham, 1971). Arsenate will attack the hemi-thioacetal intermediate in an analogous manner to phosphate. However, the compound rapidly degrades leaving the free acid which does not inhibit the enzyme. When phosphate is used instead of arsenate the reaction rate will reach a maximum due to the build up of product (1,3-bisphosphoglycerate), which inhibits GAPDH.

The reaction is followed by measuring the formation of NAD(P)H at 340 nm for either varying concentrations of substrate and saturating NAD(P)⁺, or varying concentrations of coenzyme at saturating substrate. All assays were measured on a Lambda 40 spectrophotometer at 298 K. The initial rates were plotted in *Kaliedograph* (Synergy Software) and fitted against Michaelis-Menten curve fitting algorithm to give an estimation of K_M and V_{max} . The turnover number for the enzyme, k_{cat} , was determined by calculating the number of moles of NAD(P)H formation through its extinction coefficient at 340 nm of 6220 $M^{-1}cm^{-1}$.

2.6. Methods for the analysis of GAPDHA

2.6.1. Cloning and over expression

The nucleotide sequence encoding GAPDHA was cloned into the pET expression vector pET151/D (Invitrogen). Primers for the PCR amplification of the *gapA* open reading frame (ORF) were: CACCATGCCAATTAGAAT for the forward primer and TATATAGCACAAAATTAA for the reverse primer. The PCR reaction was carried out as described in Chapter 2.2.2 with an extension time of 45 seconds and a primer annealing temperature of 51 °C. Agrose gel electrophoresis confirmed the PCR reaction had been successful, with a band consistent with the size of the *gapA* gene being present once visualised on a transilluminator. The band was excised and the agrose gel removed by purification using a GelExtraction Kit (Qiagen). The topoisomerase reaction was performed as described in Chapter 2.2.3 with linear pET151/D vector as stipulated by the manufacture. The reaction was transformed into chemically competent DH5 α in accordance with the method in Chapter 2.2.4 and selected on LB-agar plates containing 50 μgml^{-1} ampicillin. Colonies were screened by performing colony PCR as described in Chapter 2.2.5. Positively identified colonies were minipreped and sequenced by dideoxy sequencing to confirm the sequence of the *gapA* gene compare to the published sequence in the database. For expression of the *gapA* ORF in *E. coli*, the pET151/D-*gapA* vector was transformed into the expression strain *Rosetta DE3*, which encodes for tRNAs not found within *E. coli*. A single colony was added to 50 ml LB as overnight starter culture, containing 50 μgml^{-1} ampicillin and 34 μgml^{-1} chloramphenicol. The following day the culture was diluted 1:100 into 6 l of 2YT medium (16 g tryptone, 8 g yeast extract, 5 g NaCl per litre of water) containing 50 μgml^{-1} ampicillin and 34 μgml^{-1} chloramphenicol, aliquoted into 2.5 l flasks and incubated with shaking at 30 °C. Once the density of the cultures approached an OD₆₀₀ of 0.7, the flasks were cooled to 18 °C and left overnight without the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) induction. Previous expression tests with un-tagged GAPDHA had shown that expression was "leaky" as no IPTG was required to induce expression. On the contrary, IPTG induction of un-tagged GAPDHA resulted in large amounts of insoluble GAPDHA protein and this was analogous to the expression of His₆-GAPDHA from the pET151/D-*gapA* construct. The cells were harvested the following morning and resuspended in buffer A (20 mM Na₂HPO₄, 500 mM NaCl, 50 mM imidazole pH 7.4) and frozen prior to purification.

2.6.2. GAPDHA purification

Cells were thawed and supplemented with protease inhibitor cocktail VII (Calbiochem). The suspension was sonicated at 12 μ and the cell debris removed by centrifugation at 18 krpm for 30 min. The supernatant was passed through a 0.2 μ m filter to remove any large aggregates and membranes not cleared by centrifugation and loaded onto a 5 ml Hi-Trap Nickel Sepharose column (Amersham Biosciences) equilibrated in buffer A. The column was washed in buffer A to remove protein bound non-specifically to the column resin. His₆-GAPDHA was eluted with a linear gradient of buffer B (20 mM Na₂HPO₄, 500 mM NaCl, 500 mM imidazole, pH 7.4) on an AKTA Purifier (Amersham Biosciences). Pooled fractions of His₆-GAPDHA were dialysed overnight against buffer C (20 mM Na₂HPO₄, 50 mM NaCl, 1 mM DTT pH 7.2) prior to ion exchange chromatography. The theoretical isoelectric point of GAPDHA was calculated to be 8.2, therefore cation exchange chromatography was used. The dialysed sample was loaded onto a 5 ml Hi-Trap Sulfopropyl Sepharose column, equilibrated in buffer C and His₆-GAPDHA was eluted against a linear gradient of buffer D (20 mM Na₂HPO₄, 1 M NaCl, 1 mM DTT, pH 7.2). A single peak collected at ~500 mM NaCl was judged to be ~95 % pure by SDS-PAGE. A typical 6 l preparation of GAPDHA would yield approximately 50 mg protein.

His₆-GAPDHA was concentrated to 12 mgml⁻¹ using a Viva-spin™ centrifugal unit (10 kDa Mwt cutoff; Sartorius) and buffer-exchanged into 20 mM MES, 100 mM NaCl, 1mM DTT pH 6.5 prior to crystallisation. The remainder of the protein was flash frozen and stored at -80 °C prior to further experiments.

2.7. Methods for the analysis of GAPDHB

2.7.1. Cloning and protein over expression

The ORF for the *gapB* gene was amplified out of the *H. pylori* genome in the same way as *gapA* (Chapter 2.6.1). Briefly, the primers for forward and reverse amplification of *gapB* were: CACCATGAAAATTTTTATCATTGGATTG and TTAATAATGATACATAACTGG respectively. The PCR was set up as described in Chapter 2.2.2 with a primer annealing temperature of 44 °C and an extension time of 45 s used. Agrose gel electrophoresis confirmed the PCR had been successful and the band corresponding to the size of the *gapB* gene was excised and purified as described for *gapA* (Chapter 2.6.1). The topoisomerase reaction was performed as described in Chapter 2.2.3. Colonies from the cloning reaction were screened by PCR as described in Chapter 2.2.5 with dideoxy sequencing confirming the sequence integrity of the expression construct. As with the expression of *gapA*, the *gapB* expression construct was transformed into *Rosetta DE3* competent cells. For over expression of the pET151/D-*gapB* construct, 6 l of cells were grown in 2YT, exactly as described for GAPDHA over expression. Once the cell density reached an OD₆₀₀ of 0.6 the cultures were cooled to 18 °C. After the culture had cooled sufficiently IPTG was added to a final concentration of 100 µM to induce over expression and the cultures were left overnight at 18 °C with continuous shaking. The following day the cultures were harvested by centrifugation, resuspended into buffer A, and frozen prior to purification.

2.7.2. GAPDHB purification

Cells were thawed and supplemented with protease inhibitor cocktail VII (Calbiochem). The suspension was sonicated at 12 µ and the cell debris removed by centrifugation at 18 krpm for 30 min. The supernatant was passed through a 0.2 µm filter and loaded onto a 5 ml Hi-Trap Nickel Sepharose column (Amersham Biosciences) equilibrated in buffer A. To remove non-bound protein the column was washed in buffer A on the AKTA Purifier (Amersham Biosciences) until the absorbance at 280 nm (A₂₈₀) stabilised. A gradient elution against buffer B was used to elute His₆-GAPDHB. Fractions were pooled and dialysed overnight against buffer C (20 mM Na₂HPO₄, 50 mM NaCl, 1 mM DTT pH 7.2) for cation exchange chromatography. The dialysed sample was loaded onto a 5 ml Hi-Trap Sulfopropyl Sepharose column, equilibrated in buffer C and His₆-GAPDHB was eluted against a linear gradient of buffer D (20 mM Na₂HPO₄, 1 M NaCl, 1 mM DTT, pH 7.2). A single peak was collected at 600 mM NaCl and judged to be 98 % pure by SDS-PAGE analysis. His₆-

GAPDHB was concentrated to 8 mgml^{-1} using a Viva-spin™ centrifugal unit (10 kDa Mwt cutoff; Sartorius) and buffer-exchanged into 20 mM MES, 100 mM NaCl, 1mM DTT pH 6.5 and flash frozen in liquid nitrogen for storage at $-80 \text{ }^{\circ}\text{C}$.

2.8. Methods for analysing *H. pylori* mutants

2.8.1. Mutagenesis

The generation of active site mutants of *gapA* and *gapB* was achieved by using the overhanging PCR method outlined in Chapter 2.2.6. Active site mutants generated for *gapA* were C149S and C149A. Whilst the active site mutants for *gapB* were C151S and C151A. The PCR reactions were set up as described for the mutagenesis protocol in Chapter 2.2.6. An extension time of 30 s was used for the first PCR reactions with a primer annealing temperature of $50 \text{ }^{\circ}\text{C}$. The PCR reaction was analysed by agarose gel electrophoresis (Chapter 2.2.3) and the DNA excised and purified from the gel and used to set up the next PCR reaction. After agarose analysis of the second PCR reaction the full-length mutant ORF was cloned into the pET151/D vector as described in Chapter 2.6.1. Screening of colonies was performed as described in Chapter 2.2.5. Dideoxy sequencing confirmed the mutagenesis reaction had been successful.

2.8.2. Protein over expression and purification

The expression and purification of GAPDHA and GAPDHB mutants was exactly the same as for the respective wild-type enzymes (Chapter 2.6.2 and 2.7.2). With the exception that after elution of the nickel sepharose column, the proteins were buffer exchanged back into buffer A (20 mM Na_2HPO_4 , 500 mM NaCl, 50 mM imidazole, pH 7.4) prior to being incubated with TEV protease (Invitrogen) overnight. After incubation the protein solution was passed through the nickel sepharose column again, where the mutant enzymes would be collected in the flow-through. Un-tagged GAPDHA and GAPDHB mutants were concentrated as before to 10 mgml^{-1} and buffer exchanged into 20 mM MES, 100 mM NaCl, 1mM DTT pH 6.5 and flash frozen in liquid nitrogen for storage at $-80 \text{ }^{\circ}\text{C}$.

2.9. Methods for analysis of *C. jejuni* GAPDH

2.9.1. Cloning and protein over expression

The nucleotide sequence encoding the only identified GAPDH in the *C. jejuni* genome was cloned from *C. jejuni* strain NCTC1168 genomic DNA provided by Dr Jullian Ketley (University of Leicester). The forward primer used was: CACCATGGCTGTAAAAGTTGCTATAAATGG and the reverse primer: TTAAGCCTTATTTGCAATATATACTGC. Both primers had a melting temperature of 53 °C. The PCR reaction was set up as described in Chapter 2.2.2 and Table 2.1. The primer annealing temperature was 51 °C, with an extension time of 45 s. Agrose gel electrophoresis analysis as described in Chapter 2.2.3 confirmed the PCR had been successful with a band consistent with the size of the *gap* gene from *C. jejuni*. The band was excised from the agrose gel and cloned into the pET151/D topoisomerase vector, as described in Chapter 2.2.3 and analogous to *gapA* in Chapter 2.6.1. Following transformation into chemically competent DH5 α cells and plating onto LB-agar plates containing 50 μgml^{-1} ampicillin, colonies were screened for the corresponding *gap* ORF by colony PCR (Chapter 2.2.5). Dideoxy sequencing confirmed the successful cloning of the *gap* ORF into the expression vector. For over expression of the *C. jejuni gap* gene, the expression construct was transformed into *Rosetta DE3* chemically competent *E. coli*, in the same way as described for GAPDHA and GAPDHB (Chapter 2.6.1 and 2.7.1). Again, a single colony was used to inoculate a starter culture overnight in LB, containing 50 μgml^{-1} ampicillin and 34 μgml^{-1} chloramphenicol. The following day the culture was diluted 1:100 into 4 l of 2YT medium containing 50 μgml^{-1} ampicillin and 34 μgml^{-1} chloramphenicol, aliquoted into 2.5 l flasks and incubated with shaking at 30 °C. Once the cell density reached an OD₆₀₀ of 0.6, the flasks were cooled to 18 °C. Once the cultures had reached thermal equilibrium, IPTG was added to a final concentration of 200 μM to induce over expression of His₆-cGAPDH. The cultures were left overnight after induction at 18 °C with shaking until the following morning. The cells were harvested the following morning and resuspended in buffer A (analogous to His₆-GAPDHA and GAPDHB purifications, Chapter 2.6.2 and 2.7.2) and frozen prior to purification.

2.9.2. Protein purification

Cells were thawed and supplemented with protease inhibitor cocktail VII (Calbiochem). The suspension was sonicated at 12 μ and the cell debris removed by centrifugation at 18 krpm for 30 min. The supernatant was passed through a 0.2 μ m filter and loaded onto a 5 ml Hi-Trap Nickel Sepharose column (Amersham Biosciences) equilibrated in buffer *A*. To remove non-bound protein the column was washed in buffer *A* on the AKTA Purifier (Amersham Biosciences) until the absorbance at 280 nm (A_{280}) stabilised. His₆-cGAPDH was eluted against a linear gradient of buffer *B*. Fractions were pooled and buffer exchanged into buffer *A* using a HiPrep™ 26/10 Desalting Column (Amersham Biosciences) prior to TEV protease treatment. TEV protease (Invitrogen) was added in accordance to the manufactures instructions (1:20 weight:weight ratio of TEV:His₆-cGAPDH) and incubated overnight at room temperature. Removal of TEV protease and impurities from the nickel resin was achieved by passing the TEV-treated sample down a 5 ml Hi-Trap Nickel Sepharose column (Amersham Biosciences) and collecting the flow-through containing un-tagged cGAPDH. Un-tagged cGAPDH was concentrated using a Viva-spin™ centrifugal unit (10 kDa Mwt cutoff; Sartorius) and buffer-exchanged into 20 mM Tris, 100 mM NaCl, 1 mM DTT pH 7.2 and flash frozen in liquid nitrogen prior to storage at -80 °C for further experiments.

Chapter 3
Structures of Glyceraldehyde-3-phosphate Dehydrogenase A

3.1 Introduction

Analysis of the *H. pylori* genome has inferred that the glycolytic pathway is incomplete (Alm *et al.*, 1999, Tomb *et al.*, 1997). In addition, genes that are not annotated have no corresponding enzyme activities being observed *in vitro* (Chalk *et al.*, 1994), (Mendz and Hazell, 1993). Furthermore, experimental data suggests that glucose is the only carbohydrate utilised by the bacterium (Mendz *et al.*, 1993, Mendz *et al.*, 1995a, Mendz and Hazell, 1993). The Entner-Doudoroff pathway, is however, constitutively active (Mendz *et al.*, 1994a). Although glucose can be metabolised by the bacteria through the above mechanism, growth requirements of the bacteria suggests that amino acid metabolism is the primary energy source (Ashdown, 1978, Nagata *et al.*, 2003). Nevertheless, as described in Chapter 1.7.1.1, the understanding of metabolism within *H. pylori* is important for the development of new avenues for drug discovery. Within the carbohydrate metabolic pathways, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was identified as a possible target owing to the divergence of several pathways at the substrate, glyceraldehyde-3-phosphate (g3p). There are two annotated genes encoding for two GAPDHs within *H. pylori*; *gapA* and *gapB*.

This chapter analyses the *gapA* gene and the correspondingly expressed GAPDHA enzyme. Purified GAPDHA is shown to be a phosphorylating, NADP-dependent enzyme. The first GAPDH outside of chloroplast containing organisms to be shown to be NADP⁺ dependent. Analyses of apo and holo structures of GAPDHA reveal conformational changes upon nucleotide binding. In addition, the molecular specificity of GAPDHA for NADP⁺ over NAD⁺ is determined and the characteristic "fingerprint" for NADP⁺ specificity is applied in the *Campylobacterales* order, revealing further putative NADP⁺-dependent GAPDHs, of which *C. jejuni* is addressed in Chapter 4.

3.2 GAPDHA cloning and expression

Crystals of un-tagged *H. pylori* GAPDHA had been obtained previously and only grew in the presence of NAD^+ . Preliminary analysis of the electron density had shown that the enzyme was in the apo form and further attempts to reproduce these crystallisation conditions failed. To investigate GAPDHA further, the *gapA* gene (gi: 6626253) was re-cloned out of its original expression vector; pDESTTM-14 as described in Chapter 2.6.1. Re-cloning of *gapA* into the pET151/D vector (Invitrogen) resulted in the addition of an N-terminal hexa-histidine affinity tag. The addition of the hexa-histidine (His_6) tag resulted in increased yields of pure protein. His_6 -GAPDHA was expressed and purified as described in Chapter 2.6.2, with a typical chromatogram from the nickel affinity chromatography shown in Figure 3.1.

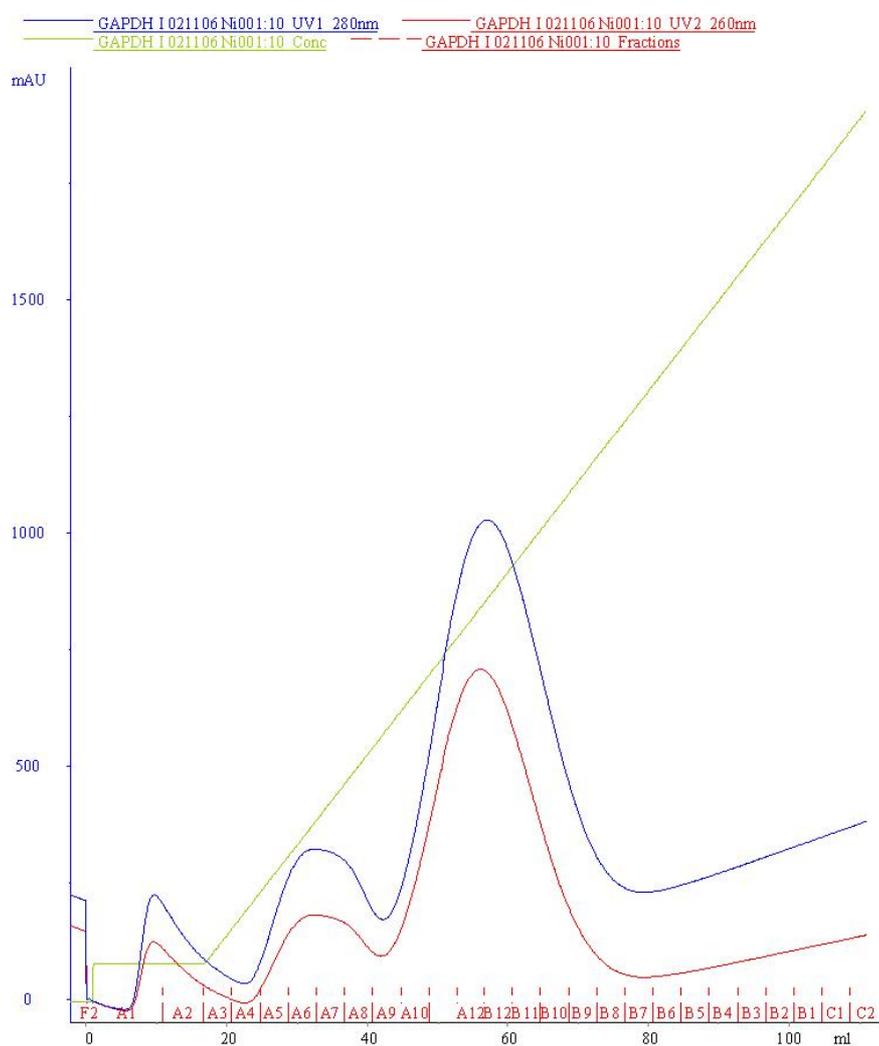


Figure 3. 1 Typical chromatogram for the purification of His_6 -GAPDHA. The absorption at 280 and 260 nm (blue and red respectively) of protein is shown plotted against volume of elution with a gradient of buffer containing 500 mM imidazole (represented in green) (as described in Chapter 2.6.2). The corresponding fractions are also shown. Fractions A11-B9 contained His_6 -GAPDHA as confirmed by SDS-PAGE analysis and were pooled.

3.3 Enzymatic analysis

Enzymatic analysis as described in Chapter 2.5 demonstrated that GAPDHA is a phosphorylating NADP⁺-dependent GAPDH, with a K_M and k_{cat} for glyceraldehyde-3-phosphate (g3p) of 0.86 mM and 22.83 s⁻¹ respectively, shown in Table 3.1. There was a detectable activity for g3p with NAD⁺ as a coenzyme and not surprisingly the K_M for g3p binding to the NAD⁺/NADP⁺ GAPDHA is similar. This would be expected due to there being little conformational change upon binding of the substrate to the active site formed by the bound nucleotide as discussed in Chapter 6.9 and has been previously demonstrated in *Bacillus stearothermophilus* (Didierjean *et al.*, 2003).

	K_M (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_M (mM ⁻¹ s ⁻¹)
NADP ⁺	0.86 ± 0.17	22.83 ± 1.10	26.5 ± 6.5
NAD ⁺	0.36 ± 0.10	8.43 ± 0.70	23.4 ± 7.0

Table 3. 1 Enzymatic parameters of GAPDHA for varying concentrations of substrate at saturating NAD⁺/NADP⁺ concentrations (3mM), measured as described in Chapter 2.5.

Therefore, once coenzyme is bound then the type of cofactor bound does not influence substrate binding. There was a negligible turnover by GAPDHA when erythrose-4-phosphate (e4p) was a substrate. The context of which is discussed in Chapter 5.7. Owing to the fact that no detectable activity with e4p was observed in the absence of phosphate or arsenate, this low turnover is consistent with the presence of low levels of contaminating g3p in the e4p preparation and is discussed further in Chapter 5.7.

Analysis on the ability of GAPDHA to bind either NAD⁺ or NADP⁺ as a coenzyme was investigated at a fixed g3p concentration of 2.4 mM, which, is not quite at saturation, however, was the highest concentration of substrate that could be used in the reaction owing to limiting amounts of g3p. Initial rate enzymatic parameters are shown in Table 3.2. It is clear that there is a greater specificity for NADP⁺ over NAD⁺ with the k_{cat}/K_M values suggesting that GAPDHA binds NADP⁺ 60 times better than it binds NAD⁺. In Chapter 4.3, Table 4.3 compares the difference between NADP and NAD⁺ binding within other NADP⁺-dependent GAPDHs and also with *C. jejuni* GAPDH, which appears to display dual coenzyme specificity. This and *C. jejuni* GAPDH described in Chapter 4 are the first GAPDHs to be purified to homogeneity that utilise NADP⁺ as a coenzyme to oxidatively phosphorylate g3p outside of *Archaea* and chloroplast-containing plants. Analysis of the

enzymatic activity is consistent with GAPDHA being a phosphorylating NADP⁺-dependent GAPDH and is simply not being activated in the presence of phosphate to catalyse the irreversible oxidation of g3p in a phosphorylation-independent reaction (i.e. as an aldehyde dehydrogenase discussed in Chapter 1.7.1.3) (Ettema *et al.*, 2008, Pailot *et al.*, 2006).

The incubation of arsenate instead of phosphate according to the method of (Racker and Krimsky, 1952) results in rapid breakdown of the 1-arseno,3-phosphoglycerate reducing product inhibition which is seen when phosphate is used as a substrate and is discussed in Chapter 2.5.

	NAD ⁺			NADP ⁺		
	K _M (mM)	k _{cat} (s ⁻¹)	k _{cat} /K _M (mM ⁻¹ s ⁻¹)	K _M (mM)	k _{cat} (s ⁻¹)	k _{cat} /K _M (mM ⁻¹ s ⁻¹)
GAPDHA	1.93 ± 0.29	8.06 ± 0.68	4.2 ± 2.3	0.07 ± 0.01	18.59 ± 0.75	265.6 ± 75.0

Table 3. 2 Enzymatic parameters of GAPDHA with varying concentrations of NAD⁺ and NADP⁺ at a fixed concentration of 2.4 mM g3p.

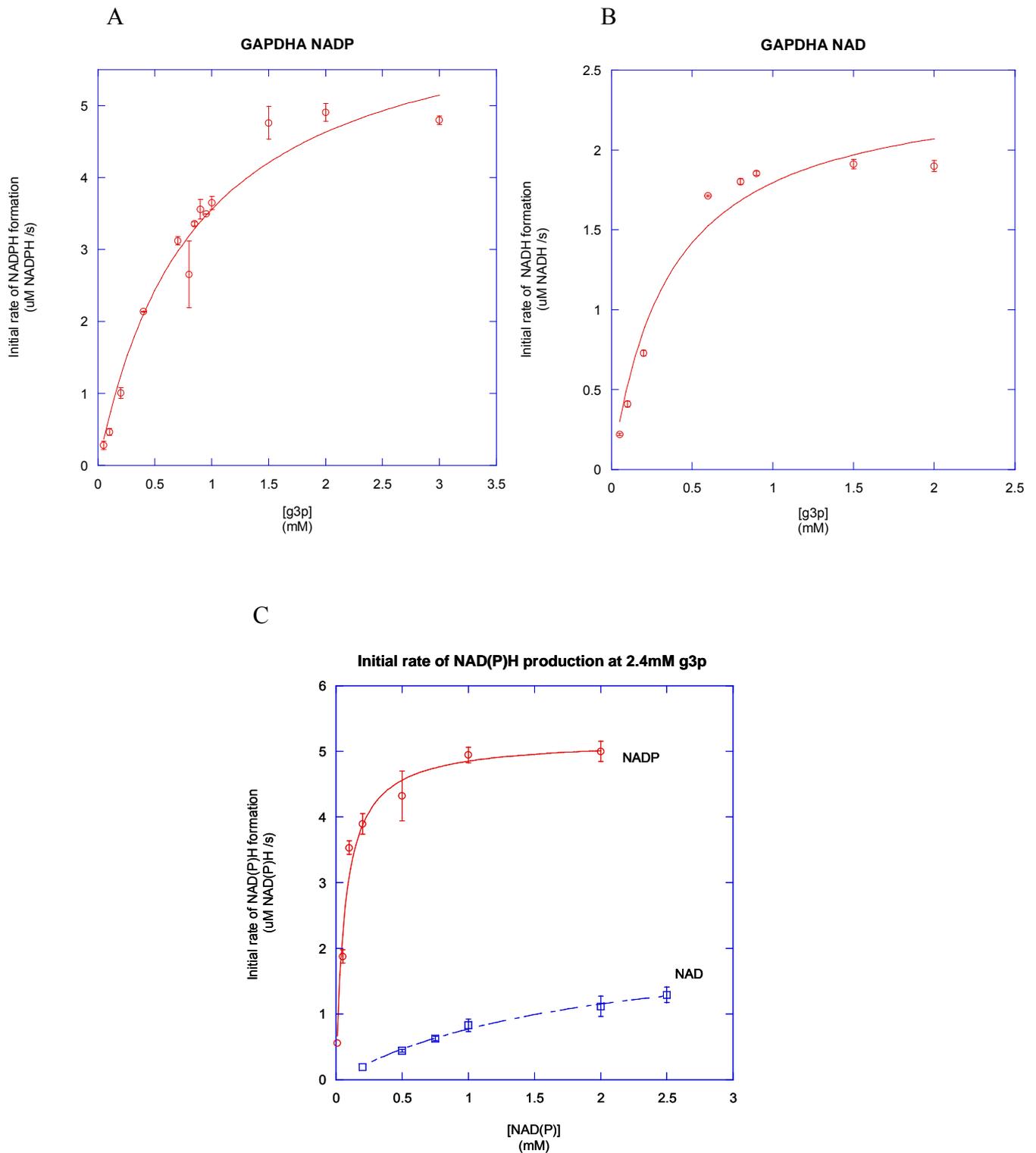


Figure 3. 2 Enzymatic analysis of GAPDHA, performed as described in Chapter 2.5. Panel A and B show the plot of initial rate for varying concentrations of g3p substrate at a fixed 3 mM concentration of either NADP^+ or NAD^+ . Panel C shows the rate for varied concentrations of either NADP^+ or NAD^+ at a fixed concentration of 2.4 mM g3p.

3.4 Crystallisation of NADP⁺-bound GAPDHA

As described in Chapter 3.2, previous attempts to obtain crystals of GAPDHA were successful only in the presence of NAD⁺. With GAPDHA being identified as a NADP⁺-dependent enzyme, (which confirmed the absence of NAD⁺ within the electron density of the previously crystallised GAPDHA). It was therefore important to obtain crystals in the presence of NADP⁺ and investigate the molecular mechanisms of NADP⁺ binding to GAPDHA. To explore crystallisation conditions for GAPDHA sparse matrix screening (Jancarik and Kim, 1991) using Wizard™ and Cryo™ I & II screens (Emerald BioSystems) was carried out in utilising the crystallisation robots described in Chapter 2.4.1. Briefly, 100 nl drops of protein solution were mixed with an equal volume of precipitant dispensed from a Genomic Solutions Cartesian Honeybee 8+1 (Harvard Bioscience) onto 96-well MRC plates (Innovadyne) in a humidified chamber. Plates were sealed with transparent tape and crystal growth was monitored using CrystalProHT (TriTek) plate storage and imaging systems at 293 K and 277 K. His6-GAPDHA of 12 mgml⁻¹ with 2 mM NADP⁺ gave initial crystals from 15% (v/v) EtOH, 100 mM Tris pH 7.0 at 293 K. Further optimisation of the conditions exploring different alcohols as precipitants and varying pHs using 500 nl drops of protein and an equal volume of reservoir, setup as described above, gave MeOH as the optimal precipitant with conditions around 15% (v/v) MeOH, 100 mM Tris pH 7.2 as most optimal. Crystals of 200 µm in diameter would grow to full-size after two days. The pH of the crystallisation buffer was found to strongly influence the morphology and quality of the crystals, as shown in Figure 3.3 with a significant reduction in crystal quality at one pH unit either side of the optimal condition.

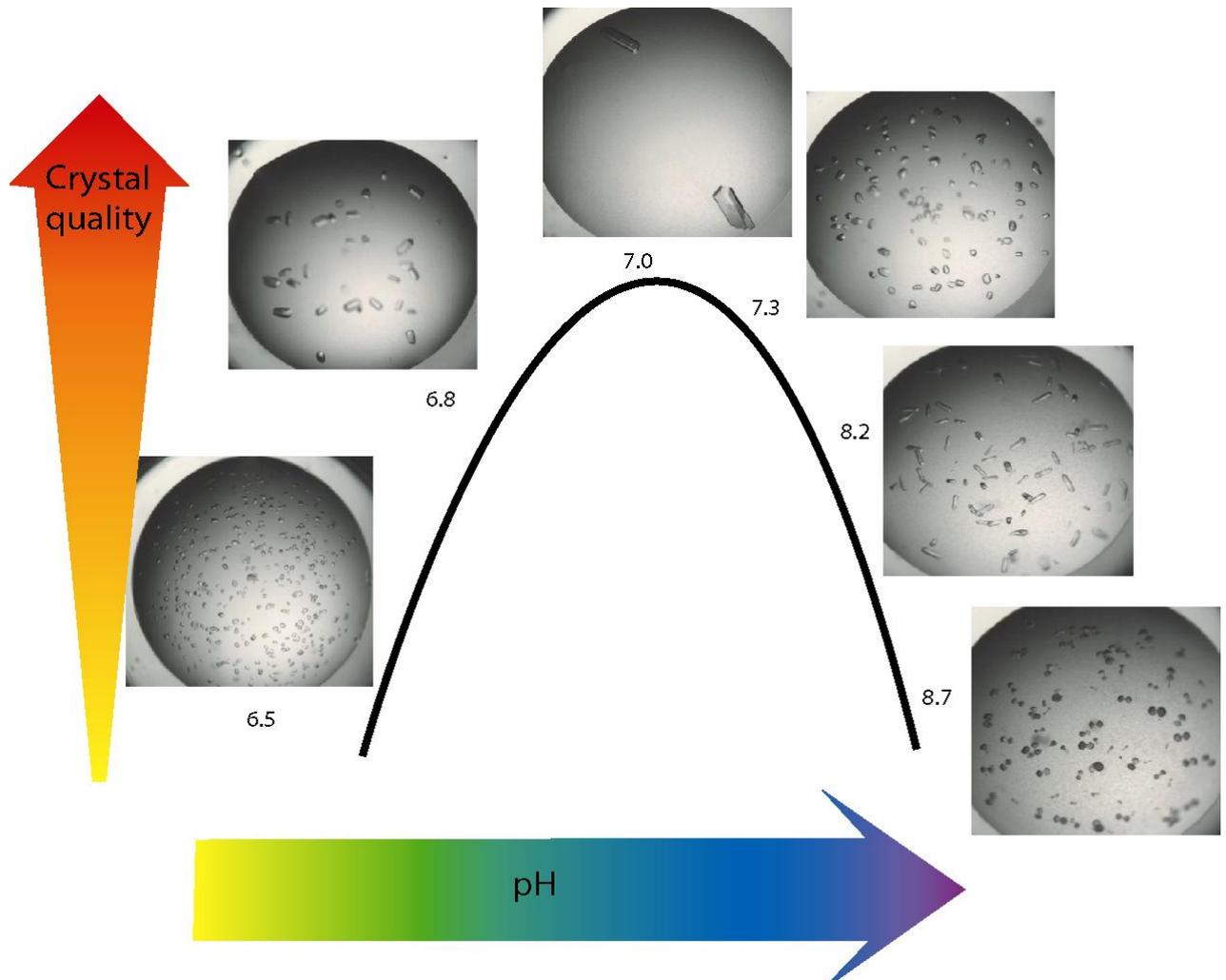


Figure 3. 3 Crystals of His-GAPDHA grown in the presence of NADP^+ . Crystals are grown out of 15% (v/v) MeOH by sitting drop vapour diffusion. As the pH of the crystallisation buffer increases from pH 6.5 there is an improvement in crystal morphology until an optimum is reached, after which the crystal quality deteriorates with further growth at the edges of the crystals becoming apparent.

3.5 Data collection of His₆-GAPDHA

Crystals of His₆-GAPDHA grown in the presence of NADP⁺ (holo-GAPDHA) were tested for cryo protection by transferring a single crystal into different cryo solutions as described in Chapter 2.4.4. Breaking the seal around the crystallisation drop would result in disintegration of the crystals due to the evaporation of methanol. Attempts to reduce the rate of evaporation by harvesting the crystals at a lower temperature or harvesting in a methanol-saturated environment did not alleviate this problem. To overcome the evaporation of methanol, different cryo mixtures were injected onto the drops by piercing the seal with a Hamilton syringe. Crystals were harvested and frozen in a stream of boiled off liquid nitrogen and were tested for their ability to diffract as a result of the different cryo mixtures by collecting two orientations on the home X-ray source (Rigaku RU2HB, Xenocs optics, R-Axis IV detector). Out of cryo solutions of paratone-N and 15-30% (v/v) glycerol mixture in the crystallisation condition, a cryo solution of 35% (v/v) MPD, 100 mM Tris pH 7.2 resulted in a significant improvement in diffraction compared to the other cryo protectants and allowed data to be collected.

A complete data set was collected from a single cryo-cooled crystal at beam line ID14-1 at the ESRF, Grenoble, using a fixed wavelength of 0.934 Å and an ADSCQ4R CCD detector with a crystal to detector distance of 300.68 mm. A typical diffraction image for holo-GAPDHA is shown in Figure 3.4, 120 images were collected at 0.5° oscillation width.

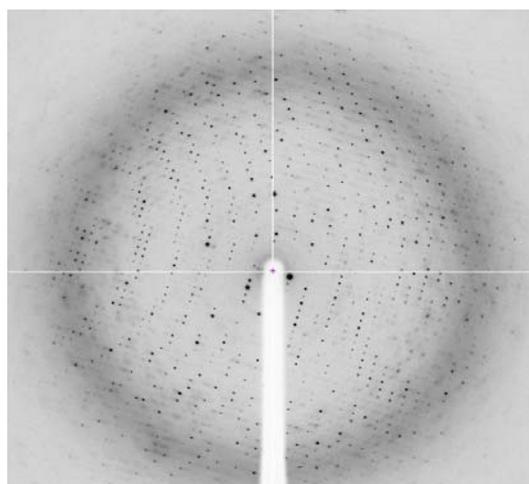


Figure 3. 4 Typical diffraction image of holo-GAPDHA grown in the presence of NADP⁺. 120 images were collected at a fixed wavelength of 0.934 Å and a crystal to detector distance of 300.68 mm.

3.6 Data processing

In the case of holo-GAPDHA (His₆-GAPDHA grown in the presence of NADP⁺) analysis of the autoindexing routines of *MOSFLM* (Leslie, 1992) suggested a primitive trigonal cell. The data was initially scaled using *SCALA* (Evans, 2006) in *P3* and examination of the distribution of intensities along the principal axes indicated either a 3₁ or a 3₂ screw axis. The Matthews coefficient suggested the presence of a tetramer in the asymmetric unit ($V_m = 2.5 \text{ \AA}^3 \text{ Da}^{-1}$; $V_s = 50.9\%$ solvent). Molecular replacement used GAPDH from *Bacillus stearoathermophilus* 1GD1, of sequence identity 46% (Skarzynski *et al.*, 1987) as a search model in *PHASER* (Read, 2001). Searching using a monomer of 1GD1 with NAD⁺ and the water molecules removed, and searching for four copies in the asymmetric unit gave a solution consistent only with the space group *P3*₂. The data was reprocessed in *P3*₂ giving an R_{merge} of 4.5% and $\{I/\sigma I\}$ of 17.9, with the space group and cell dimensions shown in Table 3.3.

	holo-GAPDHA
Space group	<i>P3</i> ₂
Unit cell parameters (a,b,c, Å)	116.88, 116.88, 95.43
($\alpha, \beta, \gamma, ^\circ$)	90.0, 90.0, 120.0
Resolution Limits	69.5-2.6 (2.74-2.60)
Number Observations	158217 (15086)
Number Unique Observations	43460 (5857)
Average Redundancy	3.6 (2.6)
Completeness	96.8 (88.3)
($I/\sigma(I)$)	17.9 (5.7)
R_{merge}^\dagger	0.045 (0.193)

Table 3. 3 Summary of data reduction statistics for holo-GAPDHA. Values shown in parentheses are for the highest resolution shell. $R_{\text{merge}}^\dagger = \frac{\sum_{hkl} |\sum_i I_{hkl,i} - \langle I_{hkl} \rangle|}{\sum_{hkl,i} \langle I_{hkl} \rangle}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices $h, k,$ and l and $\langle I_{hkl} \rangle$ is the mean intensity of that reflection

Data had been previously collected for apo-GAPDHA (un-tagged GAPDHA grown in the absence of NADP⁺ but in the presence of NAD⁺) at ID14-2 at the ESRF, Grenoble. For apo- GAPDHA 360 images had been collected at a 0.5° oscillation width, at a fixed wavelength of 0.933 Å. Intensities were measured using *DENZO* (Otwinowski and Minor, 1997) and scaled using *SCALEPACK* (Otwinowski and Minor, 1997) for a primitive monoclinic cell, the data collection statistics are shown in Table 3.4.

	apo-GAPDHA
Space group	$P2_1$
Unit cell parameters (a,b,c, Å)	75.16, 100.62, 97.79
($\alpha, \beta, \gamma, ^\circ$)	90.0, 93.7, 90.0
Resolution Limits	95.4-1.74 (1.82-1.74)
Number Observations	308994 (15449)
Number Unique Observations	147140 (7357)
Average Redundancy	2.1 (2.1)
Completeness	99.0 (99.0)
(I/ σ (I))	19.0 (2.2)
$\dagger R_{\text{merge}}$	0.161 (0.534)

Table 3. 4 Summary of data reduction statistics for apo-GAPDHA. Values shown in parentheses are for the highest resolution shell. $\dagger R_{\text{merge}} = \frac{\sum_{hkl} |\sum_i I_{hkl,i} - \langle I_{hkl} \rangle|}{\sum_{hkl,i} \langle I_{hkl} \rangle}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices h, k , and l and $\langle I_{hkl} \rangle$ is the mean intensity of that reflection.

3.7 Structure determination of GAPDHA

3.7.1 Apo-GAPDHA

Molecular replacement using NAD^+ -bound GAPDH from *Bacillus stearoathermophilus* (PDB accession number, 1GD1) gave an interpretable map at 1.8 Å in space group $P2_1$. Even after refinement, areas of the map could not be interpreted with great confidence, particularly within the regions around the S-loop. This partially built model was used as a search model for determining the phases for holo-GAPDHA (described below). Once holo-GAPDHA had been rebuilt and refined with the previously un-interpretable regions built into, the refined structure was subsequently used as a starting model to generate a map for the apo enzyme. Refinement was carried out with *REFMAC5* (Murshudov *et al.*, 1997) after selecting a random 5% of the measured reflections used for the generation of the R_{Free} flag to ensure the model is not "over refined" during the refinement process.

Restrained refinement with isotropic temperature factors (described in Chapter 2.4.9) was used, with subsequent rebuilding in *COOT* (Emsley and Cowtan, 2004) and further rounds of refinement, until no further progress in reduction of the R-factor could be made and all electron density was modelled into. Residues 178-190 which correspond to the S-loop responsible for conferring conformational change within the molecule (Murthy *et al.*, 1980, Skarzynski *et al.*, 1987) and also responsible for forming a binding pocket for the NAD^+ were still absent from the final model.

Waters were added using *COOT* (Emsley and Cowtan, 2004) at positions where peaks were greater than 1.0σ in the $m2F_o-F_c$ weighted map and greater than 3.0σ in the mF_o-F_c weighted difference map. Waters were manually inspected after further refinement and removed if they had a B-factor larger than 60 \AA^2 or were greater than 3.5 \AA distance from hydrogen bond donors or acceptors. The final R-factor for the apo model was 21.6% and R_{free} 25.4%, as summarised in Table 3.5.

	apo-GAPDHA
Resolution range (\AA)	50.06-1.75
R^{\S}	21.6
R_{free}	25.4
RMSD bond lengths (\AA)	0.009
RMSD bond angles ($^{\circ}$)	1.21
No. protein atoms	9615
No. solvent atoms	871
Average B-factors (\AA^2)	
Overall	25
Main chain atoms	23
Side chain atoms	25
Water molecules	32

Table 3. 5 Refinement statistics for apo-GAPDHA for Chains A, B, C and D. $\S R = \frac{\sum_{hkl} ||F_o(h, k, l) - |F_c(h, k, l)||}{\sum_{hkl} |F_o(h, k, l)|}$. R_{free} is defined as R calculated on a random 5 % non refined data.

3.7.2 Holo-GAPDHA

Molecular replacement with a partial-refined apo-GAPDHA as a search model, was carried out with *PHASER* (Read, 2001) as described in Chapter 3.7.1. Analysis of the solvent content in the asymmetric unit gave a Matthews coefficient of 2.56 (solvent content 51.6%), consistent with four molecules being present in the asymmetric unit. Molecular replacement was therefore carried out using chain A of apo-GAPDHA as a search model, searching for four monomers in the space group *P3* investigating, which may be the correct screw axis. A solution for *P3*₂ was obtained with no solution for *P3*₁ being found. As expected four copies of the monomer had been found which were arranged as a typical GAPDH, displaying the 222 symmetry as with the majority of previously determined structures. Subsequent cycling between model building in *COOT* (Emsley and Cowtan, 2004) and refinement in *REFMAC5* (Murshudov *et al.*, 1997) (as mentioned in Chapter 3.7.1) gave a model where the S-loop had defined electron density and NADP⁺ was clearly defined as is shown in Figure 3.5. Waters were added and inspected as described in Chapter 3.7.1. The final R-factor for the holo-model was 18% and R_{free} 27%, as summarised in Table 3.6.

	holo-GAPDHA
Resolution range (Å)	69.5-2.6
R [§]	18.0
R _{free}	27.3
RMSD bond lengths (Å)	0.013
RMSD bond angles (°)	2.04
No. protein atoms	10233
No. solvent atoms	233
Average B-factors (Å ²)	
Overall	42
Main chain atoms	50
Side chain atoms	51
NADP	47
Water molecules	42

Table 3. 6 Refinement statistics for holo-GAPDHA for chains A, B, C and D. § $R = \sum_{hkl} ||F_o(h, k, l) - |F_c(h, k, l)|| / \sum_{hkl} |F_o(h, k, l)|$. R_{free} is defined as R calculated on a random 5 % non refined data.

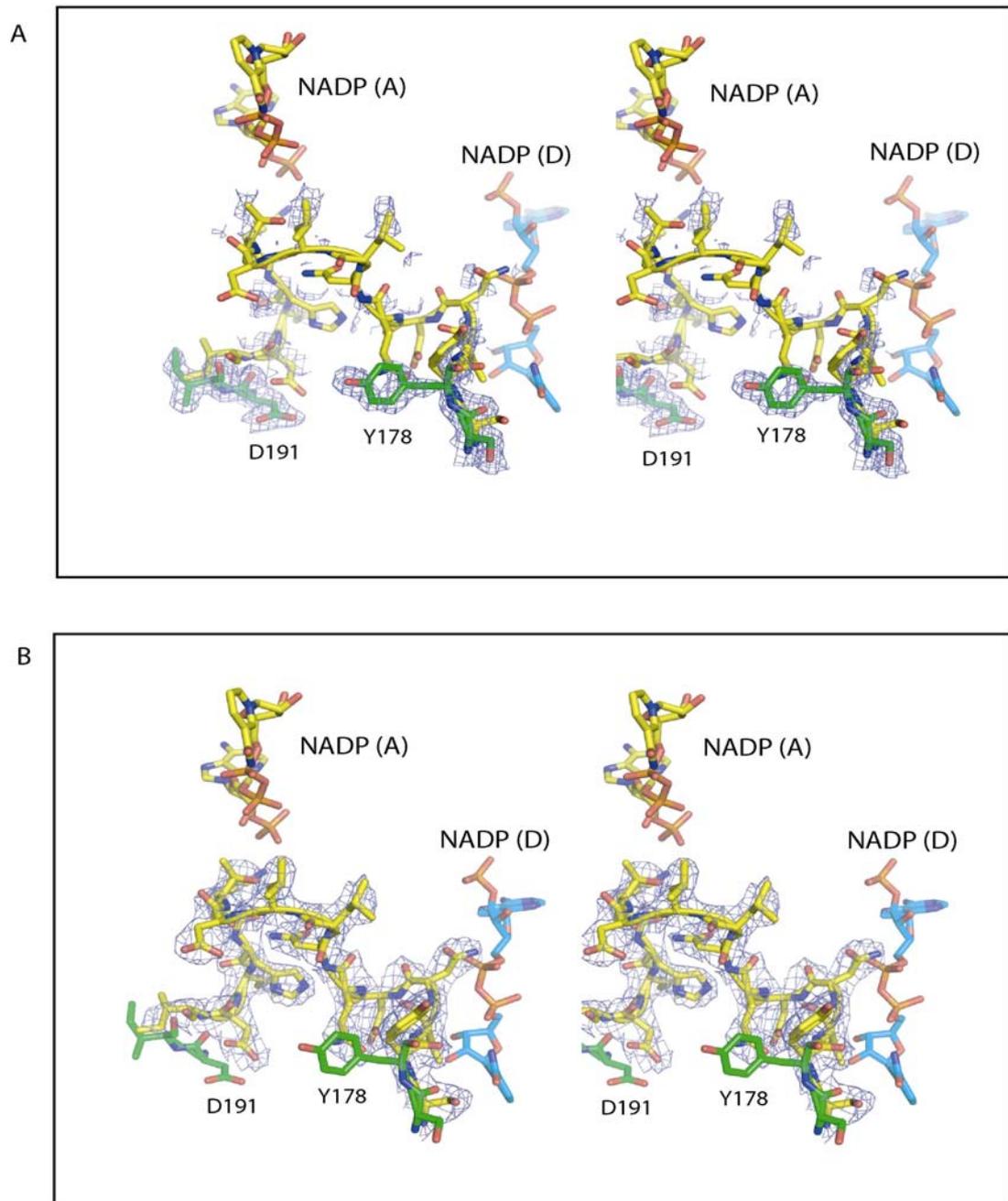


Figure 3. 5 Stereo pairs of the region from residues Y178 to D191 corresponding to the S-loop in GAPDHA (Chain A). Holo-GAPDHA is shown as yellow carbon atoms with the NADP^+ molecule belonging to chain A shown as yellow carbon atoms. The NADP^+ from the R-axis related subunit (Chain D) is shown with light blue carbon atoms. Apo-GAPDHA is shown with green carbon atoms and does not include residues 179-190 as there was no interpretable electron density for them in the refined structure. Apo-GAPDHA is superimposed onto holo-GAPDHA using the program *LSQKAB* (Kabsch, 1976) aligning the entire monomer. Panel A shows the electron density at 1.0σ for the apo-GAPDHA model from residues 177-192. Panel B shows the electron density at 1.0σ for the holo-GAPDHA model for the same residue range. For clarity the electron density for NADP^+ in panel B is not shown.

3.8 Analysis of apo- and holo-GAPDHA

3.8.1 Apo-GAPDHA

The final structure for apo-GAPDHA contains four molecules in a 222 tetramer as found for other GAPDHs. Each monomer model is composed of 314 residues out of a total of 328 in the annotated sequence, with the N-terminal methionine either unstructured or physically absent from the protein. Residues 179-190 of the S-loop have no defined electron density, which is clear from Figure 3.5. Due to the crystal packing, some chains have better defined electron density; in particular, chains A and C have some clearly defined regions, which are not so apparent in chains B and D. Although non-crystallographic symmetry (NCS) averaging was not used in refinement since any regions that had poorly defined electron density, tended to be poorly defined across all the chains and hence provided no further improvement to the electron density maps. The use of NCS averaging in refinement would also remove any asymmetry, of which has been reported for some GAPDHs (Moras *et al.*, 1975). The overall conformation of each monomer is similar with respect to each other with an RMSD of 0.274 Å for all C $_{\alpha}$ atoms, therefore there is no apparent asymmetry with respect to individual chains. A summary of the refinement parameters is shown in Table 3.5, the average B-factor for the enzyme is 25 Å² overall and 32 Å² for all 871 water molecules.

Analysis of the stereochemical quality of the model was carried out using *PROCHECK* (Laskowski *et al.*, 1993) and MolProbity (Davies *et al.*, 2007). Overall, the RMSD from standard bond lengths is 0.009 Å and 1.22° for the angles. The Ramachandran plot given in Figure 3.6 is that from Molprobity (Davies *et al.*, 2007) and shows that 99.2% of non-glycine residues have conformational angles (ϕ, ψ) in permitted regions, of which 88.8% is in the most favoured regions. All residues that fall within the disallowed and generously allowed regions have clearly defined electron density as shown in Figure 3.7.

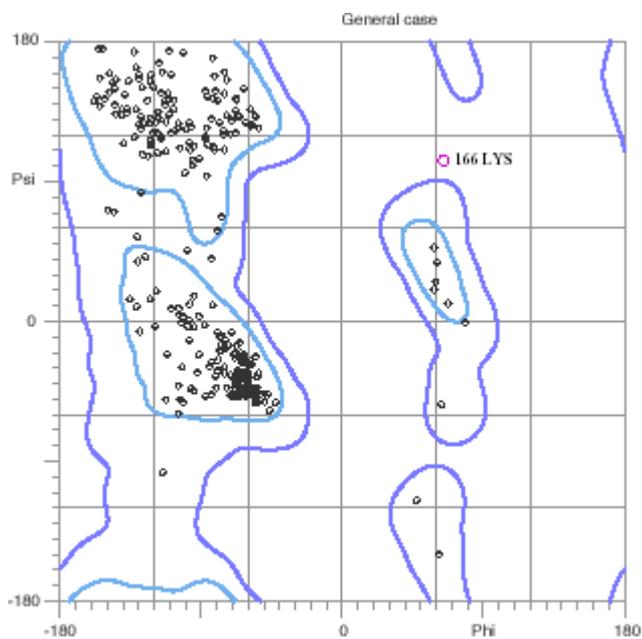
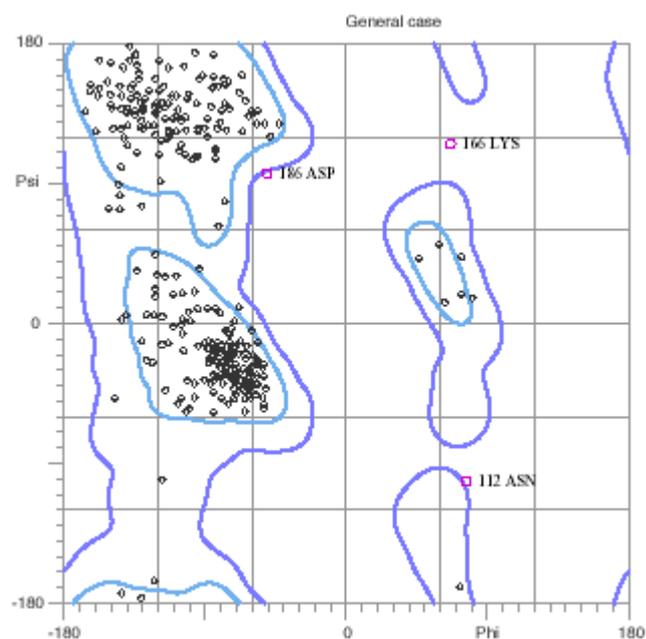
Chain A apo-
GAPDHAChain A holo-
GAPDHA

Figure 3. 6 Ramachandran plots for the models of apo-GAPDHA, top and holo-GAPDHA, bottom, generated from MolProbity (Davies *et al.*, 2007). The area encompassed by light blue lines is the allowed regions and the areas encompassed by the dark blue lines are those which are generously allowed. Labeled residues are those which are in the disallowed regions.

3.8.2 Holo-GAPDHA

Although the resolution of the data for the holo enzyme is lower than that of the apo enzyme (1.8Å data for the apo with 2.6Å data for the holo enzyme), the structure has more interpretable features, in particular there is interpretable electron density for the S-loop (residues 179-190), unstructured in the apo-structure, as is clear from Figure 3.5. In total, there were four residues where no interpretable density was found and these were at the N-terminus of Chains B and D. From Table 3.6 the average B-factor for the enzyme is 50 Å² overall and 42 Å² for all 256 water molecules.

Analysis of the stereochemical quality of the model was again carried out using *PROCHECK* (Laskowski *et al.*, 1993) and MolProbity (Davies *et al.*, 2007). Overall, the RMSD from standard bond lengths is 0.013 Å and 1.64° from the angles. The Ramachandran plot given in Figure 3.6 shows that 99.7% of non-glycine residues have conformational angles (ϕ, ψ) in permitted regions, of which 82.2% is in the most favoured regions. All residues in the disallowed regions have clearly defined electron density, as shown in Figure 3.7.

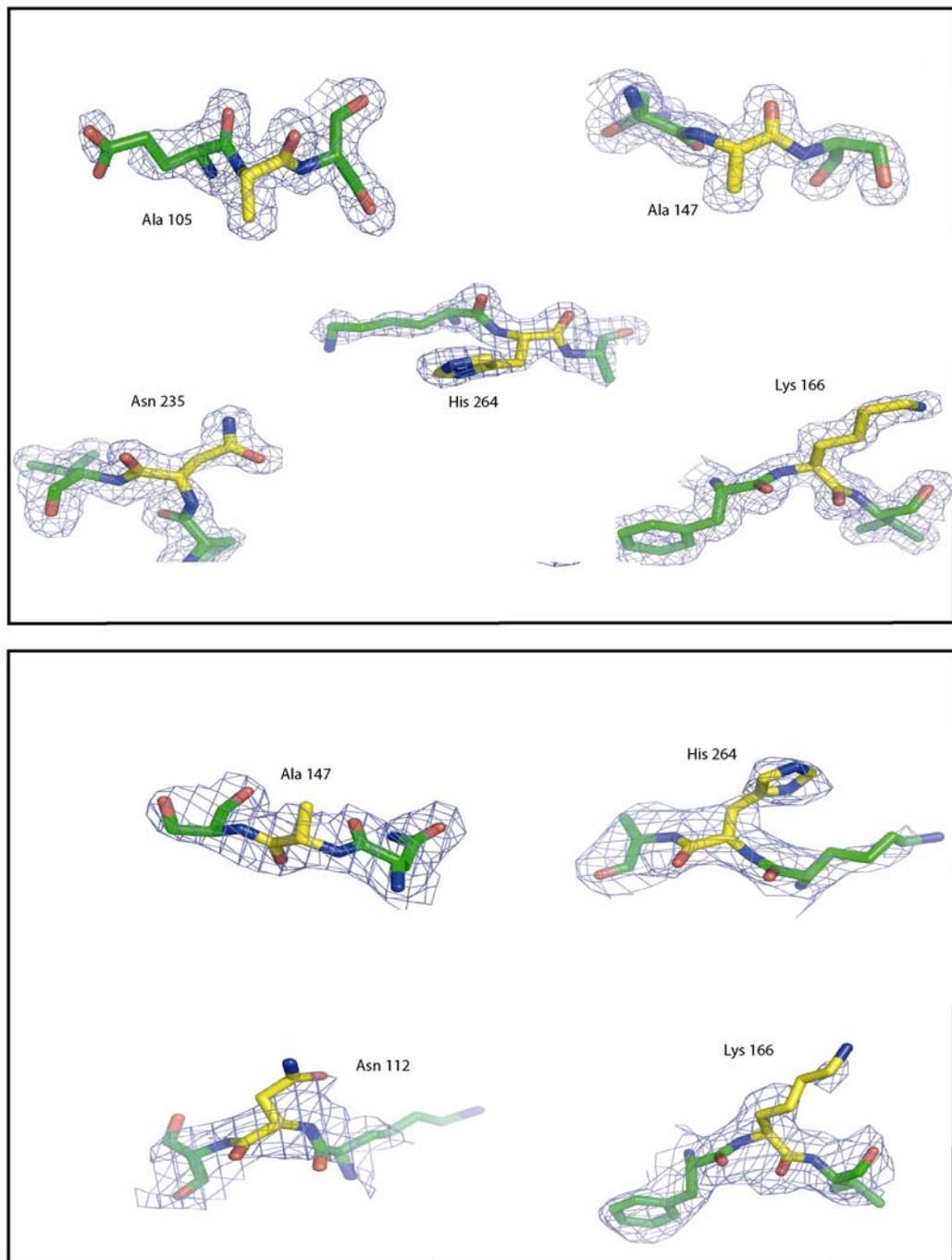


Figure 3. 7 Ramachandran outliers and generously allowed residues from the models; apo-GAPDHA, A, and holo-GAPDHA, B. The residue of interest is shown in yellow within a $2F_o - F_c$ map contoured at 1σ shown in blue. All residues are taken from Chain A and are identical with respect to conformation and position on the Ramachdran plot.

3.9 Structural analysis

The overall structure of GAPDHA is a tetramer with 222 symmetry, the two-fold axes are denoted P, Q and R, in agreement with the original GAPDH nomenclature for O, P, Q, R as the subunits (Moras *et al.*, 1975, Skarzynski *et al.*, 1987). The chains have been annotated as, A through to D in-line with current convention but the notation for the axes will be kept to avoid confusion and allow direct comparison to other GAPDHs.

3.9.1 Holo-GAPDHA

During the purification of His₆-GAPDHA, the hexa histidine tag was not cleaved prior to crystallisation (as described in Chapter 3.4). Analysis of the N-terminus of holo-GAPDHA showed extra electron density corresponding to 12 residues of the TEV protease site downstream of the histidines and V5 epitope encoded by the pET151/D vector. These ordered residues make several contacts between the symmetry related molecules and hence are involved in crystal packing. This may explain the differing space groups between the apo, untagged GAPDHA of space group $P2_1$ and that of holo, His₆-tagged GAPDHA of space group $P3_2$. Thus, the His₆-tag of the holo-enzyme favours an alternative packing. The 12 residues from the TEV protease site form a short turn, which is stabilised by intermolecular contacts within the TEV site and intramolecular contacts between the TEV site and the symmetry related molecule, as shown in Figure 3.8.

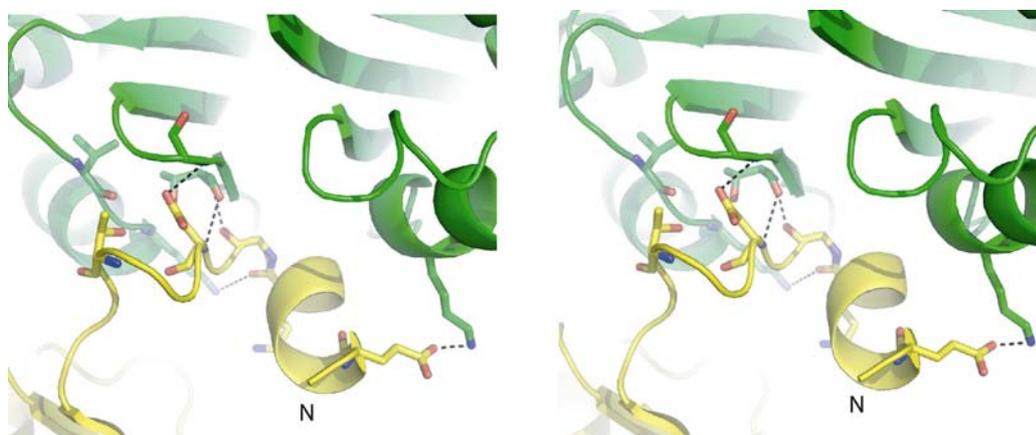


Figure 3. 8 Stereo view of the TEV cleavage site of holo-GAPDHA, shown as yellow carbon atoms, interacting with the symmetry equivalent chain in the crystal, shown as green carbon atoms. Key residues from both molecules that make contact are shown in stick representation. The secondary structure of the TEV site is also apparent with the slight turn which is stabilised through interactions with the symmetry equivalent chain and intramolecular interactions. No interactions are observed between the TEV cleavage site and its own chain.

3.9.2 Crystal Packing

Analysis of holo-GAPDHA reveals that chains B and D have a high B-factor with an average B-factor of 58 \AA^2 , compared with an average B-factor of 43 \AA^2 for chains A and C, as shown in Figure 3.9. The higher B-factor corresponds to this part of the tetramer being more disordered with respect to Chains A and C. Mapping of the B-factor onto the structure, illustrated in Figure 3.10, highlights the different packing of the tetramer in the crystal, revealing a larger solvent channel at chains B and D and fewer contacts between the symmetry related molecule within the asymmetric unit.

It should be noted that the hexa-histidine affinity tag protrudes into this solvent channel from chains B and D and that no interpretable density for the TEV site is present. The lack of interpretable density for the TEV site is therefore likely due to the fewer contacts and thus higher mobility of chains B and D, which would not allow stabilisation of the residues that form the short helical strand, observed in chains A and C.

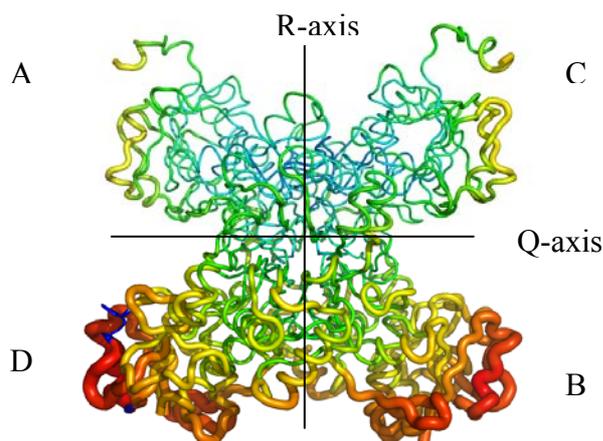


Figure 3.9 B-factor analysis of holo-GAPDHA down the P-axis. The main chain is displayed, representing the secondary structure and is coloured in a spectrum according to its B-factor with blue being the lowest and red being the highest B-factor. Areas with high B-factors are also shown in a thicker cartoon format in proportion to areas of lower B-factor values, giving further emphasis to these regions.

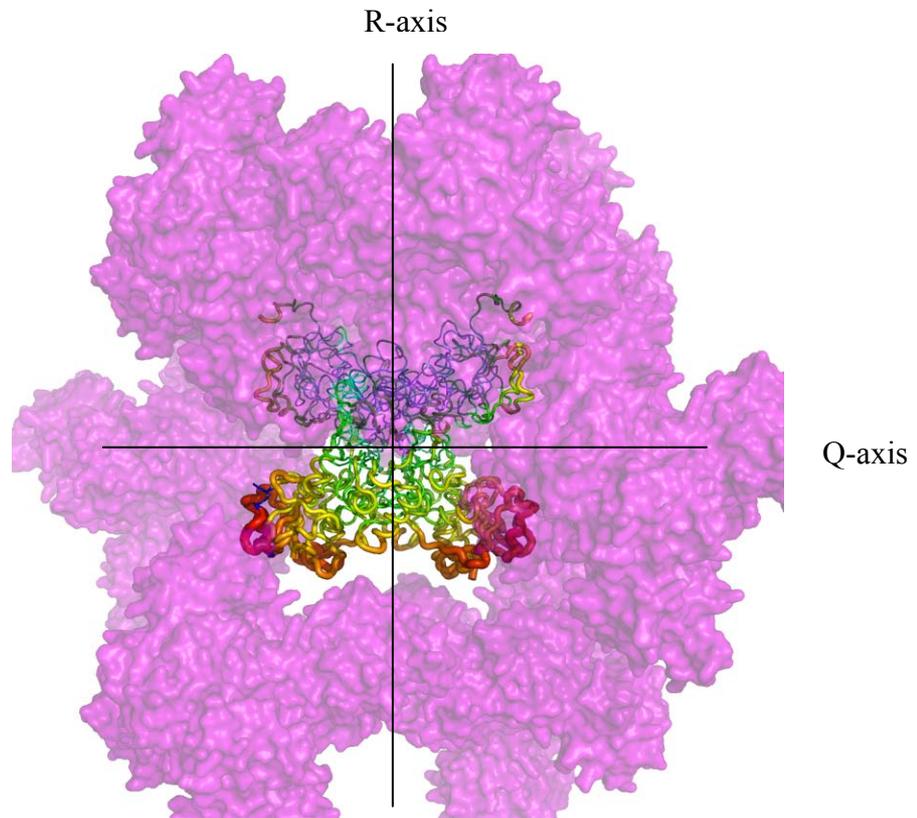


Figure 3. 10 As in Figure 3.9 the tetramer is shown but displaying a lilac surface for all tetramers within a five angstrom radius of the molecule within the asymmetric unit. As is apparent there are more closer contacts with chains A and C within the crystal, than chains B and D which have a larger surface exposed surface and are therefore more disordered, hence a higher B-factor.

3.9.3 NADP⁺ binding

GAPDHA is NADP⁺-dependent with a k_{cat}/K_M difference of approximately 20 times greater than that with NAD⁺ as demonstrated in Chapter 3.2. This distinguishes *H. pylori*'s GAPDHA from other organisms, which utilise NAD⁺, with the exception of chloroplast-containing organisms and *Archaea*. A composite omit map, generated as described in Chapter 2.4.13.1 with NADP⁺ omitted from the structure is shown in Figure 3.11 for holo GAPDHA. Contacts from Thr35 OG1, Arg78 NH1 and NH2, from within the monomer and Lys188 NZ from the S-loop of the R-axis-related monomer, coordinate O1X and O3X of the 2'-phosphate oxygen of NADP⁺.

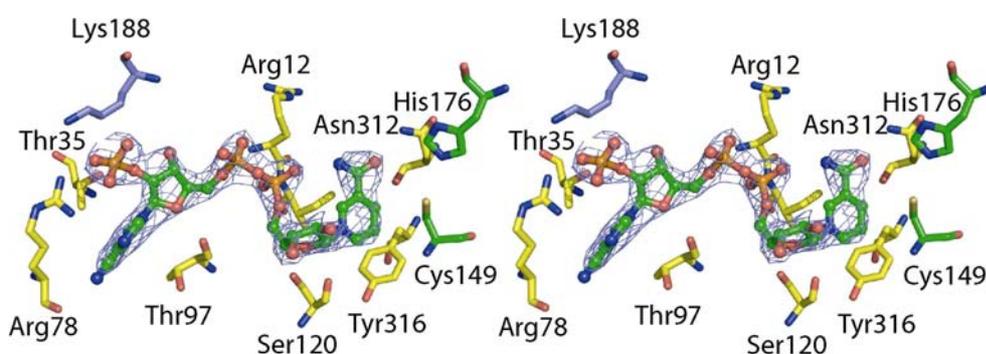


Figure 3. 11 Stereo representation of NADP⁺ bound in His₆-GAPDHA. The NADP⁺ is shown as ball-and-stick enclosed by an omit map calculated after NADP had been removed (described in Chapter 2.4.13.1) contoured at 1.2 σ with residues responsible for coordinating NADP⁺ are shown as sticks. The catalytic Cys149 and His176 are shown with green carbons as reference to the active site.

The binding of NADP⁺ within holo-GAPDHA is similar to the binding of any other NAD(P)⁺ molecule in a GAPDH in that the NADP⁺ is bound in an extended conformation as is seen in Figure 3.11 and Figure 3.12. The nicotinamide ring moiety is in a *syn* conformation due to steric hindrance from Tyr316, the *syn* conformation is further stabilised through an internal hydrogen bond between O1N and N7N of NADP⁺. Further stabilisation of the extended conformation is provided through X₃ of the characteristic motif; G-X₁-G-X₂-X₃-G, present in all nucleotide binding domains (Wierenga and Hol, 1983, Bork and Grunwald, 1990) where X₃ in GAPDHA is an isoleucine (Ile13) and is typically a hydrophobic residue in all known GAPDHs. The adenine ring is nearly perpendicular with respect to the plane of the

adenine ribose and is in an *anti* conformation. Residues of the NADP⁺ binding domain that coordinate the NADP⁺ molecule are shown in Figure 3.12 and are summarised in Table 3.7

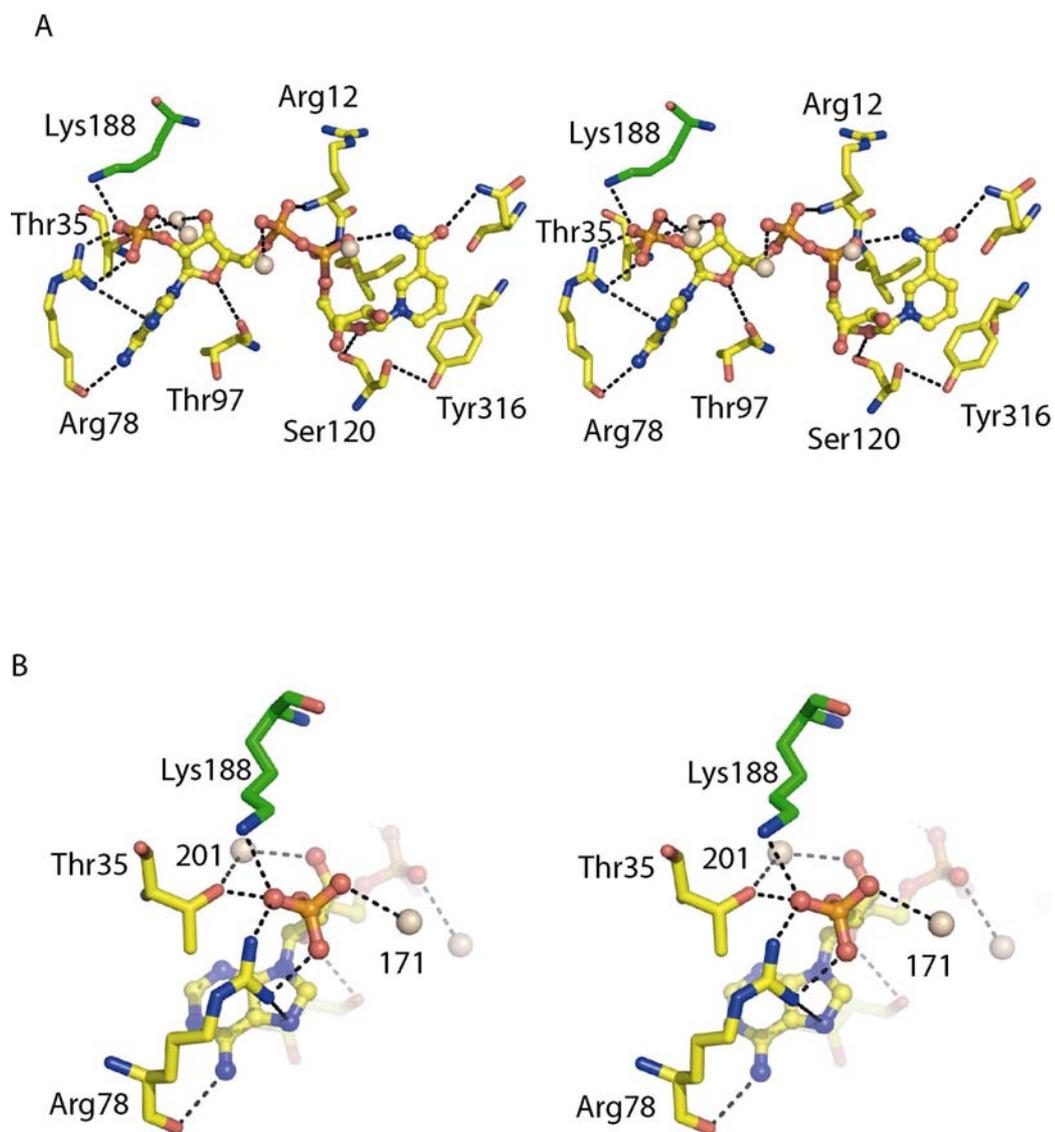
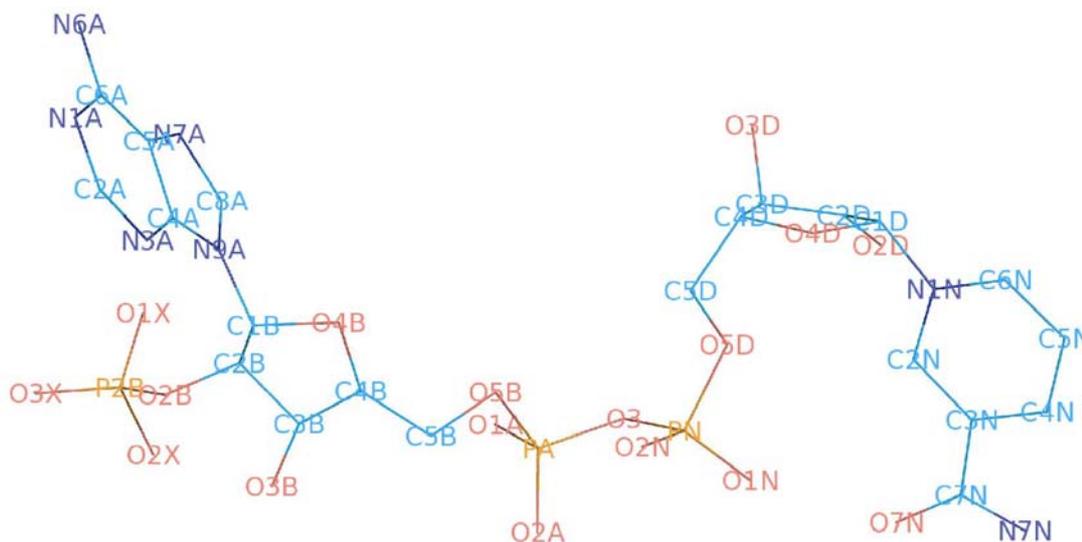


Figure 3. 12 A. Stereo view of contacts between the residues within in the NADP⁺-binding domain of GAPDHA, shown in yellow and the NADP⁺ molecule represented as ball-and-sticks. B shows the residues responsible for the coordination of the 2'-phosphate group. Waters are shown as spheres and corresponding number. Distances between the interacting atoms are shown in Table 3.7.

Residue	Distance (Å)	NADP atom
OG 35Thr	2.6	O3
NH2 78Arg	2.5	O3X
NH1 78Arg	3.0	O1X
171 H ₂ O	3.1	O2X
O 78Arg	3.0	N6A
O 97Thr	3.4	O4B
201 H ₂ O	2.6	O3B
98 H ₂ O	3.0	O1A
N 12Arg	3.0	O2A
N 13Ile	3.1	O2N
159 H ₂ O	2.7	O1N
N7N NADP	2.9	O1N
O 120Ser	2.7	O4D
ND2 312Asn	2.7	O7N
NZ 188Lys(D)	3.3	O3X

Table 3. 7 Summary of the distances between NADP⁺ (atom nomenclature shown below) and the atoms of the corresponding contacting residues. All atoms are labelled in the PDB format. Values are taken as an average from all chains within the tetramer. A schematic of NADP⁺ is shown to below for clarity of the NADP⁺ atom nomenclature.



3.10 Determinants of NADP⁺-binding

H. pylori GAPDHA contains a Rossmann fold (as for all GAPDHs); a fold highly conserved in all three domains of life. The Rossmann fold is a structural element that permits binding of either NAD⁺ or NADP⁺, but in nearly all cases shows preference for one type of nicotinamide adenine dinucleotide; phosphorylated or non-phosphorylated. To analyse the determinant of NADP⁺ binding comparisons will be made between GAPDHA and other NADP⁺-dependent GAPDHs and then with other NADP⁺-dependent dehydrogenases.

3.10.1 Comparison with other NADP-dependent GAPDHs

A comparison with the other published structures of NADP⁺-bound GAPDHs, notably the Archaeobacterium *Methanothermus fervidus* (PDB accession 1cf2) and the GAPDH from *Spinacia oleracea* (Spinach chloroplast) (PDB accession 1JN0), reveals a common signature residue for determining specificity towards NADP⁺. As shown in Figure 3.13 OG1 of Thr33 (1JN0), Thr34 (1CF2) and Thr35(holo-GAPDHA) coordinates the 2' phosphate of NADP⁺. In NAD⁺-dependent GAPDHs the residue at this position is typically a Leu (1GD1) or Pro (GAPDHB see Chapter 4) that would protrude into the site, which the 2' phosphate would occupy. A least squares fit of the NAD⁺ domain of 1GD1 onto holo-GAPDHA as in Figure 3.14, shows that Leu33 would clash with the 2' phosphate of NADP⁺. The residue adjacent to Leu33 in 1GD1 is Asp32 and it has been proposed by (Didierjean et al., 1997) that Asp32 is responsible for preventing the 2' phosphate of NADP⁺ from binding. This is due to the electrostatic repulsion between the acid group of Asp32 and the negatively charged phosphate (Charron *et al.*, 2000), in 1GD1 Asp32 forms a hydrogen bond to the 2' hydroxyl group of the ribose stabilising the bound NAD⁺. Although this residue in the holo-GAPDHA structure is a serine, rather than an aspartate, both 1CF2 and 1JN0 have the aspartate conserved at this position, therefore the aspartate can move out in these structures to accommodate the 2' phosphate and is not a key determinant for NADP⁺ binding.

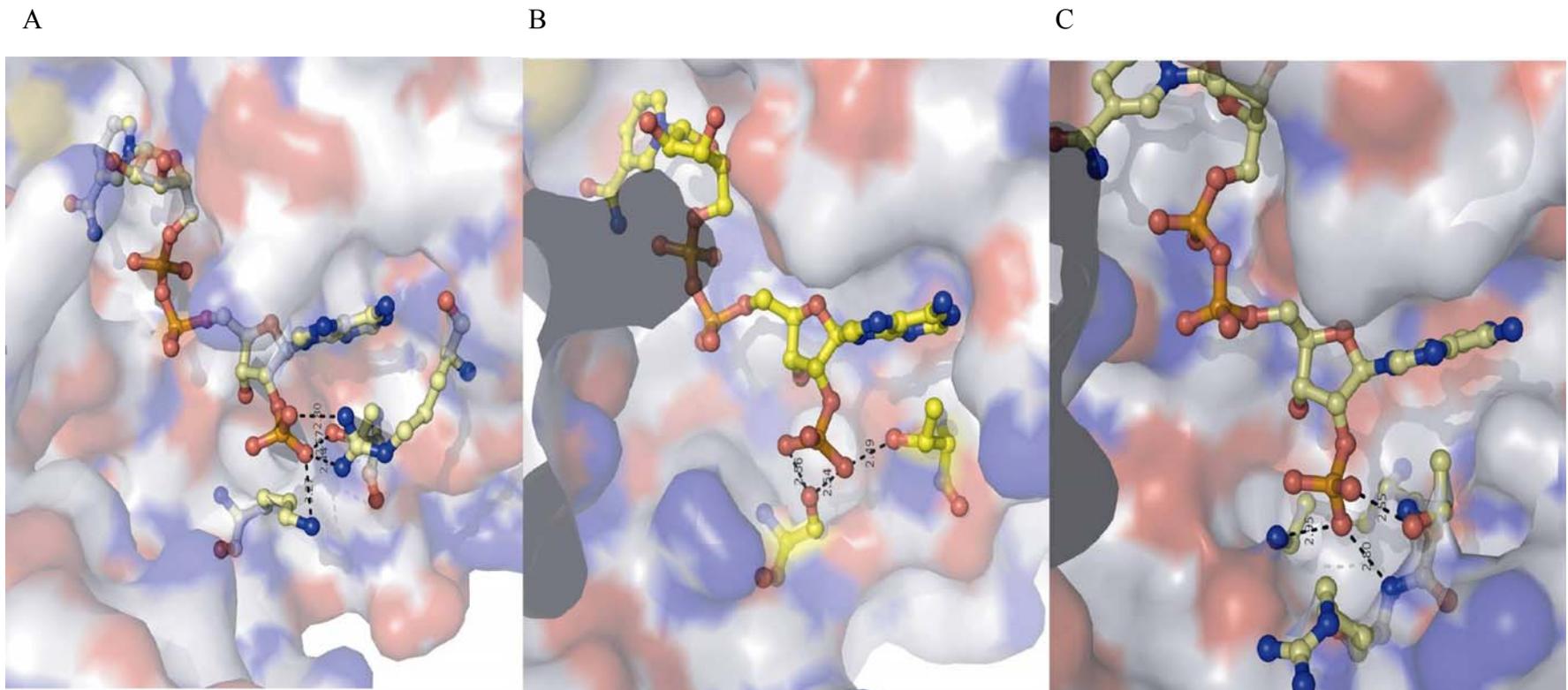


Figure 3. 13 Comparison between NADP⁺ in several other GAPDHs. In each case NADP⁺'s 2' phosphate is shown coordinated by the residues in each GAPDH. The remainder of the NADP⁺ binding pocket is represented as a surface view. A, *Methanothermus fervidus* (PDB accession 1cf2). B, holo-GAPDHA. C, *Spinacia oleracea* (spinach chloroplast) (1jn0).

Examination of the superposition of holo-GAPDHA and 1GD1 as shown in Figure 3.14 shows that there is a 1.5\AA movement between the 2' carbon of the ribose ring of NADP^+ with respect to NAD^+ with the 2' phosphate of NADP^+ facing away from the aspartate group.

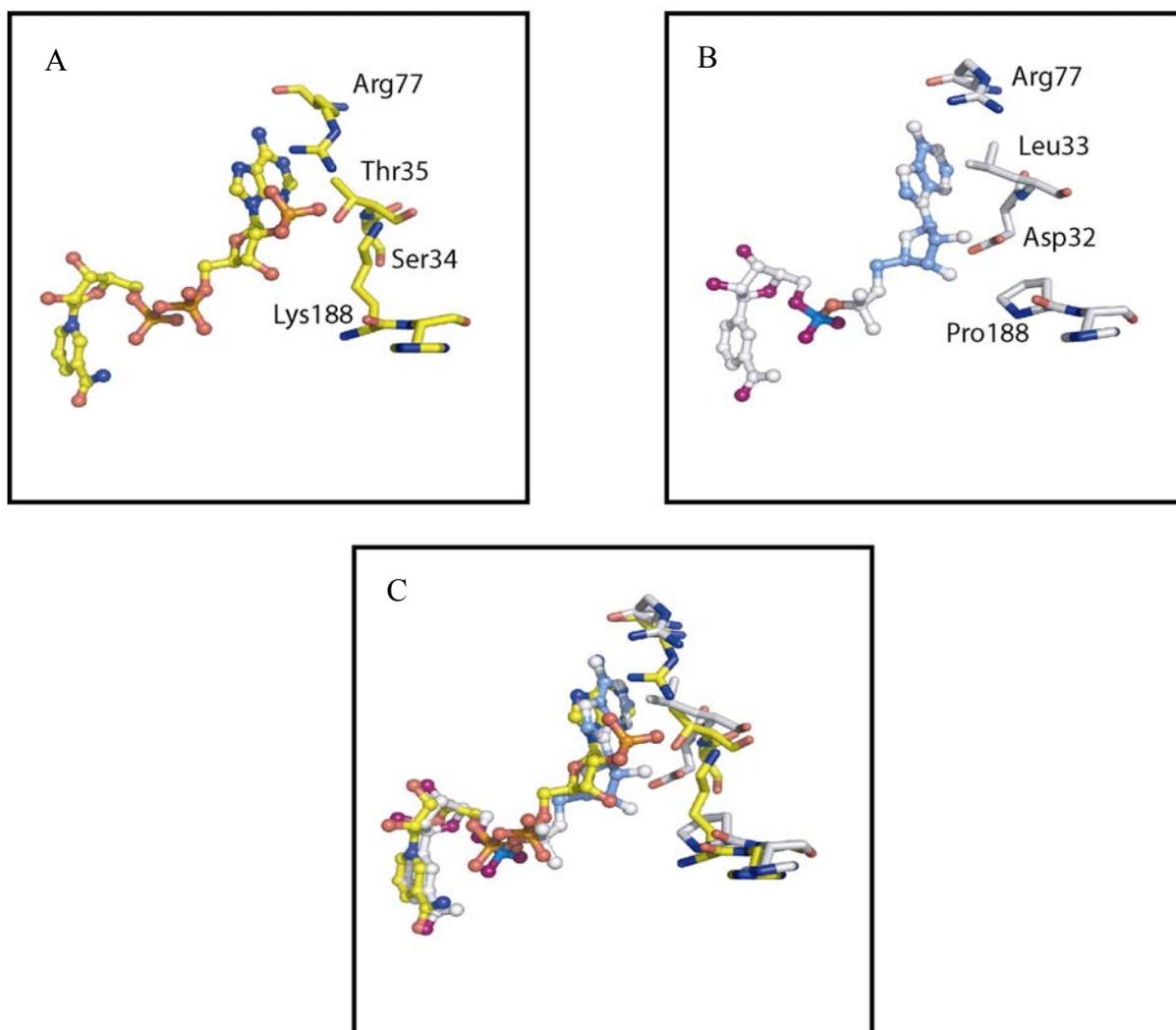


Figure 3. 14 Superposition of 1GD1 with holo-GAPDHA using LSQKAB (Kabsch, 1976) A. Residues involved in binding to NADP^+ are shown as yellow carbon atoms. B. Residues involved in permitting NAD^+ binding are shown with the carbon atoms shown in white. C. Superposition of A and B showing the 1.5\AA movement of the NADP^+ nucleotide ribose with respect to the NAD^+ ribose.

Another interesting feature is Arg77, which, in the holo-GAPDHA structure forms a salt bridge to the 2' phosphate of NADP⁺, this would therefore be seen as another key residue for specifying NADP⁺ over NAD⁺ binding. Arginine 77 is however, found in 1GD1 and is well conserved amongst other GAPDHs in particular 1CF2 and 1JN0, but in neither of these structures containing NADP⁺ is Arg77 coordinating the 2' phosphate, only in holo-GAPDHA. One final point is the involvement of the R-axis-related monomer in binding of NADP⁺. In holo-GAPDHA and 1JN0 (Figure 3.13 and 3.14 respectively) residue 188 from the S-loop of the R-axis-related monomer coordinates the 2' phosphate of NADP⁺. In the case of 1JN0 Ser188 from chain B forms a hydrogen bond to the 2' phosphate and in holo-GAPDHA Lys188 forms a salt-bridge to the 2' phosphate of NADP⁺. In 1CF2 however, the S-loop is considerably shorter and does not make any contacts with the NADP⁺. In all other NAD⁺-binding GAPDHs this residue is a proline, as in 1GD1 and shown in Figure 3.15. Whilst in 1JN0 this position is Ser188 and in *H. pylori* GAPDHB an alanine occupies this position (described in Chapter 5.6.5). Therefore, residue 188 appears another key determinant for NADP⁺ binding, specifying either a lysine or a serine. Other, more subtle variations may be required, in particular within the S-loop to permit flexibility for the loop to be able to interact with NADP⁺. As this sequence is disordered within the apo-GAPDHA structure only the inference that it becomes less flexible upon binding to coenzyme can be made.

Previous mutagenic approaches with *Bacillus stearothermophilus* have had some success in reverting cofactor specificity from NAD⁺ to NADP⁺ (Didierjean *et al.*, 1997). The D32G mutation removed the electrostatic repulsion from the aspartate allowing accommodation of the 2' phosphate of NADP⁺ and L187A removed steric hindrance from the leucine and P188S formed the hydrogen bond to the 2' phosphate as in 1JN0. However, Arg77 did not coordinate NADP⁺ as in the case of *H. pylori* holo-GAPDHA. Residues responsible for conferring NADP⁺ binding over NAD⁺ are highlighted in the multiple sequence alignment in Figure 3.15. However, this analysis suggests that the only key residue in determining NADP⁺ specificity in GAPDHs is a polar residue at position 35 i.e. Thr35 or Ser35. The other residues that do make contact provide better coordination of the 2' phosphate and vary from enzyme to enzyme, depending (probably) on the global function of the GAPDH within the organism with respect to the need to control the NAD⁺/NADP⁺ ratio.

```

1gd1      -AVKVGINGFGRIGRNVFRAAL----KNPDIEVVAVNDL-TDANTLAHLLKYDSVHGRLD  54
1jn0     -KLKVAINGFGRIGRNFLRCWHG---KDSPLDVVVINDT-GGVKQASHLLKYDSILGTFD  55
gapA     MPIRIAINGTGRIGLCAIRVAS----QRKDIEIVAINST-AELETLLHLIRHDSVHGHE  55
gapB     --MKIFINGFGRIGRCVLRAILERNDTNPKEVIGINDP-ANWEILAYLLEHDSVHGLLP  57
1cf2     -MKAVAINGYGTVGKRVADAIQA----QDDMKVIGVSKTRPDFEARMALKKGYDLYVAIP  55
          :  *** * :*                ::: ...      :  * .  .:  :

1gd1     AEVSVNGNNLVVNGKEIIVKAEKDPENLAWGEIGVDIVVESTGRFTKREDAAKHLEAGAK  114
1jn0     ADVKTAGDSAISVGKVIKVVSDRNPNVNLPGWDMGIDLVIIEGTGVFVDRDGAGKHLQAGAK  115
gapA     AQLNADRTLNIHGSKNILVLSERDINKLDFSAANAEEIIIECTGKFNSLEASSAHLKNSVK  115
gapB     KEVRYSNYKLIIGSLEIPVFN--SIKDLK----GVDVIECSGKFLEPKTLENYLLLGAK  111
1cf2     ERVKLFEKAGIEVAGTVDDMLD-----EADIVIDCTPEGIGAKNLKMYKEKGIK  104
          :      : . :                ::::: :      . : . *

1gd1     KVIIISAP-----AKNEDITIVMGVNQDKYDPKAHHVI-  146
1jn0     KVLITAP-----GKGDIPTYVVGVNNEEGYT-HADTII-  146
gapA     KVIIISAP-----AQN-TPTFVYGVNHKNYH--NESVIS  145
gapB     KVLISAPFMGEYDEKQYPTLVYGVNHFCYQ--NQAIVS  147
1cf2     AIFQGGE---KHEDIGLSFNLSNYEESYGKDYTRVV-  138
          :: .      .      .      *      ::

```

Figure 3. 15 Alignment of the nucleotide binding domains of NADP⁺-dependent GAPDHs; 1jn0,1cf2, gapA and NAD⁺-dependent GAPDHs; 1gd1 and gapB. Residues that determine NADP⁺ binding over NAD⁺ binding are shown in yellow, as a comparison residues wrongly predicted by (Didierjean *et al.*, 1997) to be involved in NAD⁺ binding over NADP⁺ are shown in red. 1gd1, 1jn0, 1cf2 are all PDB accession numbers, whilst gapA and gapB are GAPDHA and GAPDHB respectively.

3.10.2 Comparison with NADP⁺ binding in other Rossmann folds

The analysis of specificity between nicotinamide coenzymes has been the subject of intense investigation within the dehydrogenase super family, which have the characteristic Rossmann fold for the nucleotide binding domain. Investigations have inferred that in most cases for NADP⁺ binding a positively charged arginine is present which can coordinate the 2' phosphate and also stack against the adenine ring further stabilising its *anti* conformation (Carugo and Argos, 1997). Whilst Arg77 of GAPDHA forms a salt bridge to the 2' phosphate group of NADP⁺, it is not the case in 1JNO and because of its presence in NAD⁺-binding GAPDHs then it is assumed that an arginine at this position is not a key determinant in NADP⁺ binding over NAD⁺.

Other analysis has focused upon the importance of a hydroxylated side chain at position 35 in GAPDHA. Studies on several oxidoreductases have shown the importance of this position for conferring NADP⁺ binding. In trypanothione reductase a tyrosine residue, whose phenol group performs the same function as threonine or serine (Bailey *et al.*, 1994), occupies this position. Although in GAPDH from *C. jejuni* Thr34 does not directly co-ordinate the 2' phosphate (discussed in Chapter 4.8.3).

A least squares fit of NADP⁺ molecules in a variety of dehydrogenases from the aldehyde superfamily are shown below as comparison to GAPDHA.

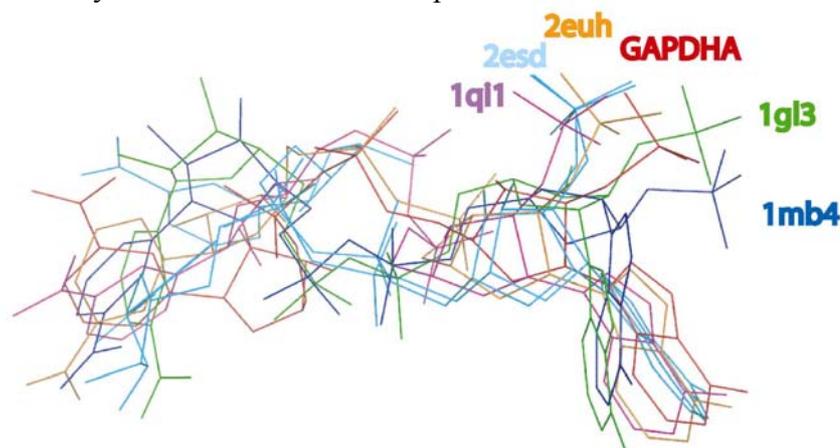


Figure 3. 16 LSQFIT of NADP⁺ from within the aldehyde dehydrogenase superfamily onto holo-GAPDHA. This highlights the diversity between different members of the superfamily; 1gl3, aspartate semialdehyde dehydrogenase (Hadfield *et al.*, 2001); 1mb4, Aspartate-beta-semialdehyde Dehydrogenase (Blanco *et al.*, 2003); 1q11, NADP⁺-dependent aldehyde dehydrogenase (Cobessi *et al.*, 2000); 2esd (D' Ambrosio *et al.*, 2006); 2euh, (Cobessi *et al.*, 1999); 2gz1, aspartate-beta-semialdehyde dehydrogenase, (Faehnle *et al.*, 2006).

3.10.3 Effects of NADP⁺ binding

Examination of the intra and inter domain rearrangements upon binding of NADP⁺ can be shown by comparison of the apo and holo structures. As discussed in Chapters 1.8.1.1.1 and 3.9 each chain is composed of an NAD(P)⁺ binding domain and an catalytic domain. Analysis of the separate domain movements using *LSQKAB* (Kabsch, 1976) to superimpose residues belonging to the NAD(P)⁺ domain; residues 1-147, 311-328 and the catalytic domain; residues 148-310. Reveals that there is 0.3 Å movement of the NAD(P)⁺ binding domain upon a least-squares fit of the NAD(P)⁺ domain of the apo structure upon the holo structure and an 0.6 Å movement when moving the catalytic domain of the apo structure onto the holo structure. Figure 3.17 shows that most of the intra domain movements occur within the NAD(P)⁺ domain, as would be expected.

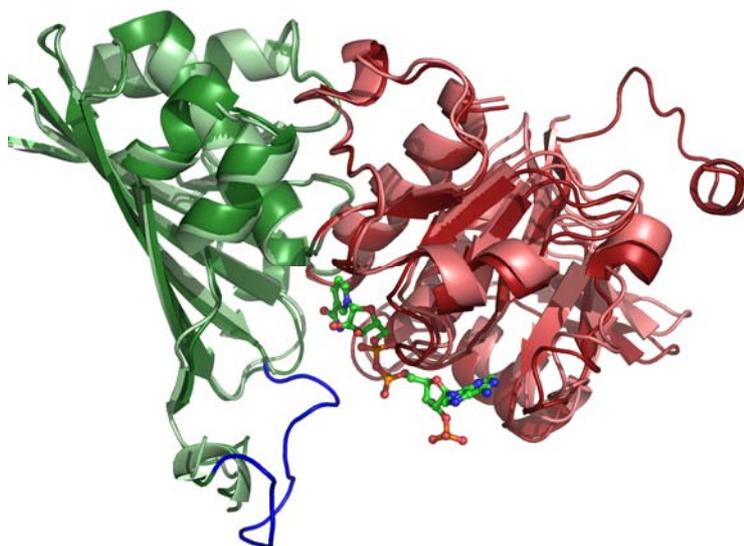


Figure 3. 17 Overlay of apo-GAPDHA NAD⁺ domain (salmon pink) and catalytic domain (light green) of chain A with the NADP⁺ domain (dark red) and catalytic domain (dark green) of holo-GAPDHA (*LSQKAB* (Kabsch, 1976)). NADP is shown as ball and stick representation, with the unstructured region of apo-GAPDHA, present in holo-GAPDHA shown in blue. This region corresponds to the S-loop which has been shown to be responsible for nucleotide binding in the subunit along the R-axis-related subunit.

It has been shown in *Bacillus stearothermophilus* GAPDH that upon binding of NAD⁺ there is rotation of the NAD⁺-binding domain by 4° with respect to the catalytic domain (Leslie and Wonacott, 1984) and more recently (Duce *et al.*, 1996) in support of inter domain movements observed within *H. pylori* GAPDHA.

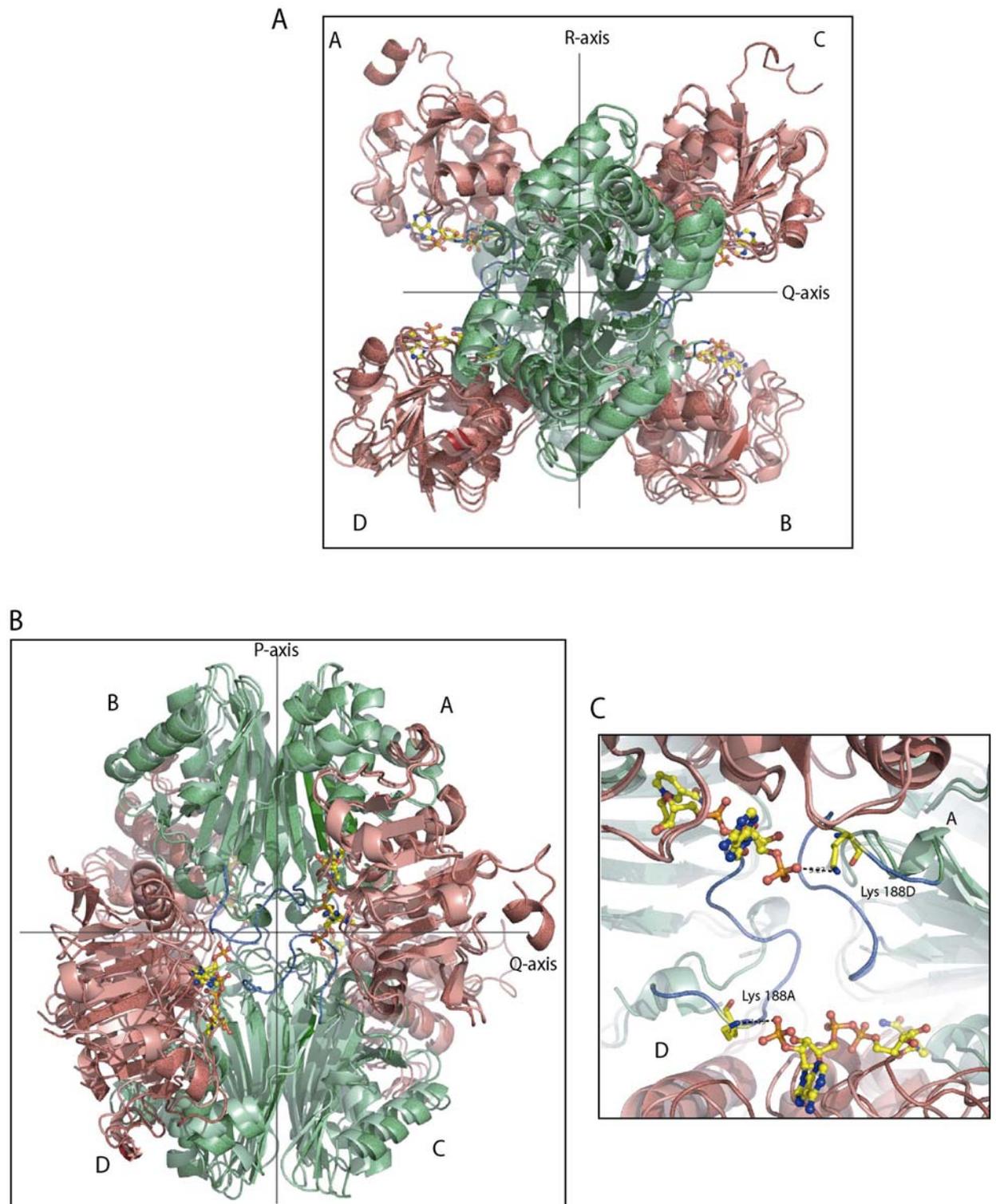


Figure 3. 18 Analysis of the tetrameric arrangement of holo- and apo-GAPDHA. As in Figure 3.9, the apo NADP domain is shown in salmon pink, with the catalytic domain in light green. Whilst the holo NADP domain is shown in dark red and the catalytic domain as dark green. The S-loop of the holo structure is shown in blue. A, view of the tetramer along the P-axis. B, view along the R-axis . C, View of chains A and D along the R-axis showing interaction of Lysine 188 from the R-axis-related monomer to the adjacent NADP⁺ molecule.

Nucleotide binding to GAPDH is cooperative, with strong negative cooperativity for NAD⁺ binding to the rabbit muscle GAPDH (Peczon and Spivey, 1972) although positive cooperativity has been observed for the yeast GAPDH (Gennis, 1976). It has been shown that the P-interface of the monomer is responsible for transmitting conformational changes required for cooperativity (Roitel *et al.*, 2003) since the majority of the inter-subunit interactions occur across the P-axis interface then this supports the motion for conferring cooperativity. It is therefore likely that Lys188 of the R-axis-related monomer provides specificity for binding of NADP⁺, as is shown in Figure 3.12C, over NAD⁺, rather than acts as a conformational switch to the other monomers.

3.11 Implications for *H. pylori* carbohydrate metabolism

The structure of the holo-GAPDHA is the first structure of a GAPDH outside of chloroplast containing organisms that uses NADP⁺ as the primary co-enzyme, in a phosphate-dependent reaction. As will be presented in Chapter 4, *H. pylori* is not the only unique bacterium to catalyse the NADP⁺, phosphate-dependent oxidative phosphorylation of glyceraldehyde-3-phosphate. Without any molecular biology data to demonstrate the expression levels of *gapA* within *H. pylori* during activation of glycolysis (a glucose rich medium) or gluconeogenesis (an amino acid rich medium) and knock-out studies to try and identify specific pathways involved, a true understanding of the significance (or insignificance) of this finding can not be made. Only inferences based upon existing models and genomic analysis can be made. Based upon current understanding there are two main hypothesis that can be derived.

Firstly, GAPDHA catalyses the oxidative reaction of glyceraldehyde-3-phosphate yielding NADPH and 1,3-bisphosphoglycerate. The generated NADPH serves not as an electron donor for energy production (the normal fate of NADH) but is used instead for biosynthetic pathways. However, there are reports that NADPH could be the main electron donor for the NADH-quinone oxidoreductase (Complex I) (Finel, 1998).

Secondly, GAPDHA catalyses the reverse reaction, reduction of 1,3-bisphosphoglycerate to yield glyceraldehyde-3-phosphate and NADP⁺. Glyceraldehyde-3-phosphate is then utilised in the Pentose Phosphate pathway (Chapter 1.7.1.1) for nucleotide biosynthesis and further NADPH production from the Pentose Phosphate pathway.

3.12 NADP⁺-dependent GAPDH's within the *Campylobacterales* order

For any significant hypothesis of GAPDHA's function within *H. pylori* to be made, levels of *gapA* expression in *H. pylori* would be required. Knowledge of other GAPDH's putative co-factor specificities within the *Campylobacterales* order would suggest if this NADP⁺ specificity is a common feature amongst the epsilon proteobacteria, or reserved solely to *H. pylori*. Searching the *Campylobacterales* order for GAPDH sequences, which contained a hydroxylated residue at the equivalent position to Thr34, revealed other species that contain a putative NADP⁺-dependent *gap* gene. Figure 3.19 reveals there is a clear link within the *Campylobacterales* order between potential NADP⁺-dependent GAPDHs and the NADP⁺-dependent GAPDHA identified in this chapter. All species, with the exception of *Campylobacter*, contain two GAPDH encoding genes. In all cases, this is analogous to *H. pylori* with one gene being a putative NADP⁺-dependent GAPDH and the other NAD⁺-dependent. The phylogenic tree in Figure 3.19B shows how the *gap* genes can be divided into two classes, NADP⁺-dependent and NAD⁺-dependent. One interesting feature is all of the species within the *Campylobacter* genus only contain one *gap* gene (based upon sequence homology) and these are predicted to be NADP⁺-dependent. The next Chapter will address whether *C. jejuni*'s GAPDH is NADP⁺-dependent.

A

```

Wolinella_succinogenes-A      -MALKVAINGTGRIGLCACRVIGNR----DDIELVALNTT-APIDTLVHL 44
Helicobaacter_hepaticus-A     -MAIKLAINGTGRIGLCAARIIGQR----DDVELVAMNTT-ADIDTLVHL 44
Helicobacter_acinonychis-A   -MPIKIAINGTGRIGLCTIRVASQR----KDIEIVAINST-AEIETLLHL 44
GAPDHA                        -MPIRIAINGTGRIGLCAIRVASQR----KDIEIVAINST-AELETLLHL 44
Sulfurimonas_denitrificans-A -MALKIAINGFGRIGRVCARIAATR----DDVEIVAINM-ASMDMMLYL 44
Arcobacter_butzleri-A        -MAVKVAINGLGRIGRVCARIASR----NDVELVAANAS-GSEEMIQYN 44
Campylobacter_coli-A         -MAVKVAINGFGRIGRVCARIISKR----DDIELVAINDT-TDIELTKYL 44
Campylobacter_jejuni-A       -MAVKVAINGFGRIGRVCARIILER----NDIELVAINDT-TDIELTKYL 44
Campylobacter_lari-A         -MAVKVAINGFGRIGRVCARIIMKR----DDIELVAINDT-TDIELTKYL 44
Campylobacter_upsalienis-A   -MAIKVAINGFGRIGRVCARIAER----DDLELVAINDT-TELELTKYL 44
Campylobacter_curvus-A       -MSVKVAINGFGRIGRCAARIILGR----DDVELVAINDT-AKRDMTRYL 44
Campylobacter_fetus-A        -MALKIAINGFGRIGRCAARIILNS----SEVELAIINDT-AERNMTRYL 44
Thiomicrospira_crunogena-A   -MTIKVINGFGRIGRMAFRAAAKDF----ANIEVVAINDL-LDPEYLAYM 45
Wolinella_succinogenes-B     MSPLRIAINGFGRIGRSVARVIAQR----SDVELVAINDL-ASPETLAYL 45
Helicobacter_hepaticus-B     --MINIAINGFGRIGRSIMRVALQHKYK-EHISTVAINDL-NDWEILSYL 46
Helicobacter_acinonychis-B   ---MKIFINGFGRIGRCLRAILERNDTNHPLEVVGINDB-ANWEILAYL 46
Arcobacter_butzleri-B        -MSTKILLNAGRIGKAVLKQLEN----KDFEIVINEINPYIENIVYS 45
      .: :** ***** :                .: * : :

```

B

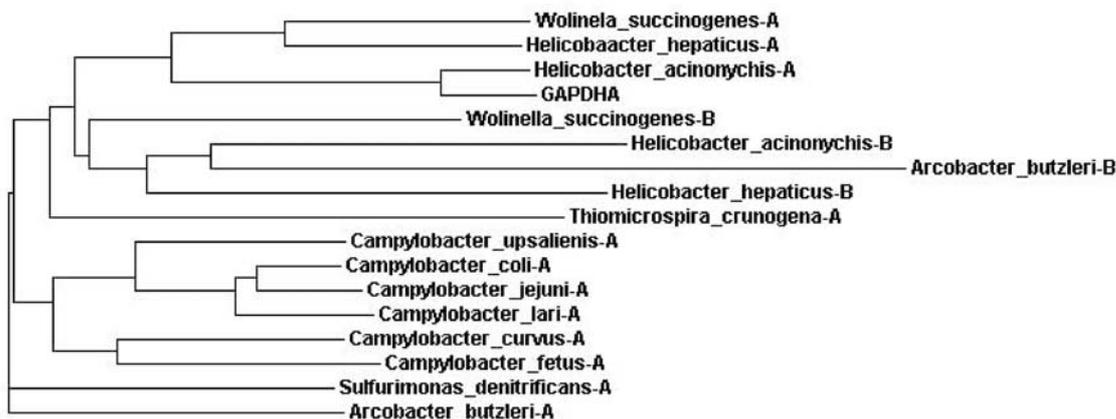


Figure 3. 19 Sequence analysis of the NAD(P)⁺ domain from the *Campylobacterales* order. Panel A shows a sequence alignment with residues at position 34 highlighted to determine if the enzyme is more likely to be NADP⁺ dependent rather than NAD⁺ dependent. Residues coloured in red is the invariant Thr34 for NADP-dependent GAPDHA. Residues in magenta are more likely to be NAD⁺-dependent based upon GAPDHB and 1gd1. Residues containing a serine or methionine are coloured in cyan as these fall out of the prediction. Panel B shows the phylogenetic relationship between the different species and also the difference between the two GAPDH forms in species containing more than one GAPDH. In these cases the sequences are dissimilar enough to allow grouping into the appropriate nucleotide dependence.

3.13 Conclusion

This Chapter has addressed the cloning, expression, purification and crystallisation of the *gapA* gene from *H. pylori*, the first NADP⁺-dependent phosphorylating GAPDH outside of chloroplast containing organisms. An analysis of the apo and holo-GAPDHA reveals the S-loop in the apo structure is disordered. Upon binding of NADP⁺ to GAPDH, the S-loop becomes structured with Lys188 of the R-axis related monomer coordinating the 2' phosphate of NADP⁺. The residues providing specificity for NADP⁺ were analysed in context of current mutagenic studies by Branlant and co-workers (Didierjean *et al.*, 1997, Duee *et al.*, 1996) and a revised method for analysing NADP⁺ specificity was proposed. This approach was used to analyse other GAPDH's within the *Campylobacterales* order and identified a similar feature to *H. pylori*. Based upon sequence analysis only one *gap* gene was identified in the *Campylobacter* genera and this was predicted to be NADP⁺-dependent. *C. jejuni*, like *H. pylori* is a human pathogen infecting the gastrointestinal tract. To extend the discussion about NADP⁺-dependent GAPDHs further within the *Campylobacterales* order, the next Chapter will address the *gap* gene from *C. jejuni*. In addition, the role of GAPDHA within *H. pylori* has been speculated and is discussed further, in context with GAPDHB (Chapter 5) in Chapter 7.

Chapter 4.
***Glyceraldehyde-3-phosphate Dehydrogenase from
Campylobacter jejuni***

4.1. Introduction

As described in Chapter 1.2.2 *C. jejuni* belongs to the same *Campylobacterales* order as *H. pylori*. Like *H. pylori*, *C. jejuni* colonises the gastrointestinal tract of humans, but unlike *H. pylori* is regarded as an "accidental tourist" (Morris and Nicholson, 1987). *C. jejuni* is a commensal organism of poultry infecting the gastrointestinal tract as a result of ingesting under cooked poultry. In the majority of patients infected with *C. jejuni*, illness is short and typically involves diarrhoea, abdominal pains and fever. In addition to the chronic gastroenteritis, *C. jejuni* infection can result in the immune-related Guillain-Barré Syndrome, which affects the peripheral nervous system. The high incidence of infection and illness (two and a half million cases within the USA annually) results in unnecessary economical loss through time of work and hospital treatment in the serious of cases. Therefore *C. jejuni*, like *H. pylori* is regarded as a serious risk factor and warrants further investigation.

The genomic sequence of two species of *Campylobacter jejuni* have been reported, *C. jejuni* NCTC11168 and *C. jejuni* RM1221. The overall genetic make-up of *C. jejuni* is similar to *H. pylori*, with regards to genomic size (only 1.64 Mbp) and contains incomplete metabolic pathways. As would be expected for the bacteria in the same order as *H. pylori*, *C. jejuni* is adapted to its host-specific interactions. This is reflected in its requirements for certain amino acids as precursors to metabolism and its lack of carbohydrate transporters along with growth requirements showing its lack of dependence on glucose for metabolism (Velayudhan *et al.*, 2004). Post-genomic analysis of *C. jejuni*, highlights the apparent lack of key enzymes for glycolysis although gluconeogenesis appears complete (Parkhill *et al.*, 2000).

One surprising feature of *C. jejuni's* gluconeogenic/glycolytic pathway is the presence of only one gene encoding a glyceraldehyde-3-phosphate dehydrogenase, in contrast to *H. pylori*, which contains two genes encoding GAPDH (Chapter 1.7.1.1). In Chapter 3 it was demonstrated that one of the *gap* genes from *H. pylori* (*gapA*) encoded a NADP⁺-dependent GAPDH. Structural analysis of GAPDHA identified a key residue for predicting possible NADP⁺ specificity of GAPDHs. This information was applied to analyse other GAPDH sequences within the *Campylobacterales* order and other putative NADP⁺-specific GAPDHs were identified (Chapter 3.12). Within *C. jejuni* the only identified GAPDH was predicted to be NADP⁺-dependent. Thus this chapter will address the cloning, expression, purification, kinetic and structural analysis of the *gap* gene from *C. jejuni* NCTC11168 and demonstrate that the gene-encoding *cgap*, encodes an NADP⁺-dependent GAPDH, termed cGAPDH. As will be presented, cGAPDH appears to display dual coenzyme specificity between NAD⁺ and NADP⁺. Structural analysis of holo-cGAPDH reveals subtle differences in the coordination of the 2' phosphate of NADP⁺, which may account for the differences in coenzyme binding.

4.2. Cloning, over expression and purification of cGAPDH

The gene encoding cGAPDH was cloned from *C. jejuni* NCTC11168 genomic DNA into the pET151/D vector (Invitrogen) facilitated by TOPO isomerase as described in Chapter 2.7.1 of Materials and Methods resulting in the addition of an N-terminal hexa-histidine affinity tag and a TEV cleavage site at the 5' end of the coding sequence for *cgap*. Over expression of the intact *gap* gene was achieved in the *E. coli* strain Rosetta DE3 with the addition of 200 μ M IPTG to cells grown to a density (O.D.₆₀₀) of 0.7 and cooled to 18 °C prior to induction. Cells were left over night at 18 °C prior to harvesting.

His₆-cGAPDH was purified from cells by nickel affinity chromatography as described in Chapter 2.7.2. The hexa-histidine tag was removed by incubation with TEV protease overnight. Cleaved protein was separated from the TEV protease and non-specific proteins from the nickel affinity purification by a subsequent nickel affinity chromatographic step, where cleaved cGAPDH was collected in the flow-through. Purity was judged at 95% based on SDS-PAGE analysis. A typical chromatograph of His₆-cGAPDH from the nickel affinity column is shown below in Figure 4.1

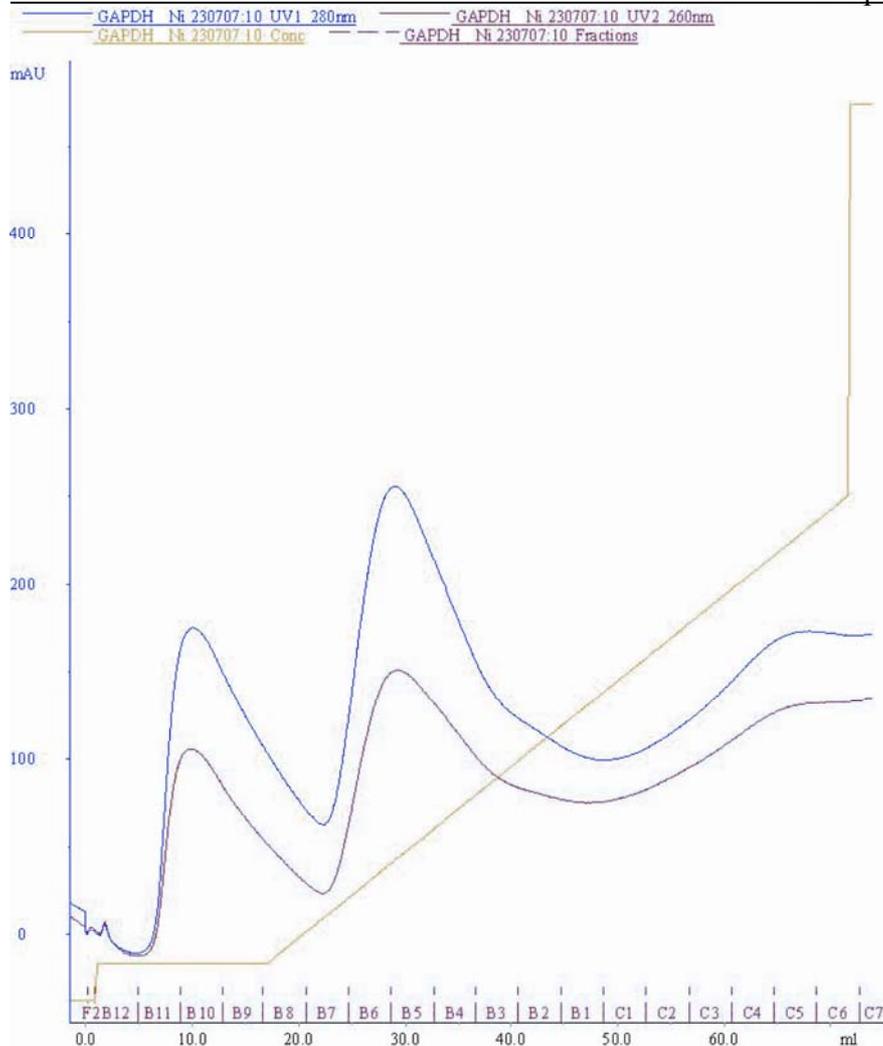


Figure 4. 1 Typical chromatograph for the purification of His₆-cGAPDH. The absorption at 280 and 260 nm (blue and red respectively) of protein is shown plotted against volume of elution with a gradient of buffer containing 500 mM imidazole (represented in green) (as described in Chapter 2.9.2). The corresponding fractions are also shown. Fractions B6-B3 contained His₆-cGAPDH and were pooled.

4.3. Enzymatic analysis of cGAPDH

Solution work confirmed the NADP^+ -dependence of cGAPDH by measuring the initial rate of reaction for cGAPDH at saturating NADP^+ concentrations and limiting glyceraldehyde-3-phosphate concentrations giving a k_{cat} and K_M for cGAPDH for glyceraldehyde-3-phosphate (g3p) of 7.55 s^{-1} and 0.41 mM respectively. There was also a significant turnover of g3p in the presence of NAD^+ as a cofactor, with a k_{cat} and K_M of 2.29 s^{-1} and 0.26 mM respectively, as summarised in Table 4.1 and Figure 4.2.

	g3p		
	K_M	k_{cat}	k_{cat}/K_M
NADP^+	0.41 ± 0.06	7.55 ± 0.44	18.4 ± 7.3
NAD^+	0.26 ± 0.04	2.29 ± 0.12	8.8 ± 3.0

Table 4. 1 Calculation of K_M and k_{cat} values for cGAPDH calculated over various concentrations of glyceraldehyde-3-phosphate at a fixed NAD(P)^+ concentration of 3 mM.

These values are consistent with that of GAPDHA, discussed in Chapter 3.3. The specificity for NAD^+ and NADP^+ was measured for cGAPDH as described in Chapter 3.3 and are summarised in Table 4.2.

	NAD^+			NADP^+		
	K_M	k_{cat}	k_{cat}/K_M	K_M	k_{cat}	k_{cat}/K_M
g3p	1.16 ± 0.30	12.57 ± 1.31	10.8 ± 4.4	0.77 ± 0.13	17.43 ± 0.96	22.6 ± 7.4

Table 4. 2 Enzymatic parameters of cGAPDH with varying concentrations of NAD^+ and NADP^+ at a fixed, saturating concentration of 3.0 mM g3p.

cGAPDH binds NADP^+ twice as strongly than it does NAD^+ , on the basis of their k_{cat}/K_M ratios in Table 4.2. This value is significantly lower than that for GAPDHA, which had a k_{cat}/K_M ratio between NADP^+ and NAD^+ 60 times. Comparisons with other NADP^+ -dependent GAPDHs within the literature also show a greater difference between specificity for the different nicotinamides and are summarised in Table 4.3. These results show that cGAPDH has a slight preference for NADP^+ binding over NAD^+ and may be able to utilise both depending on the energy status of the bacterium (discussed in Chapter 4.9).

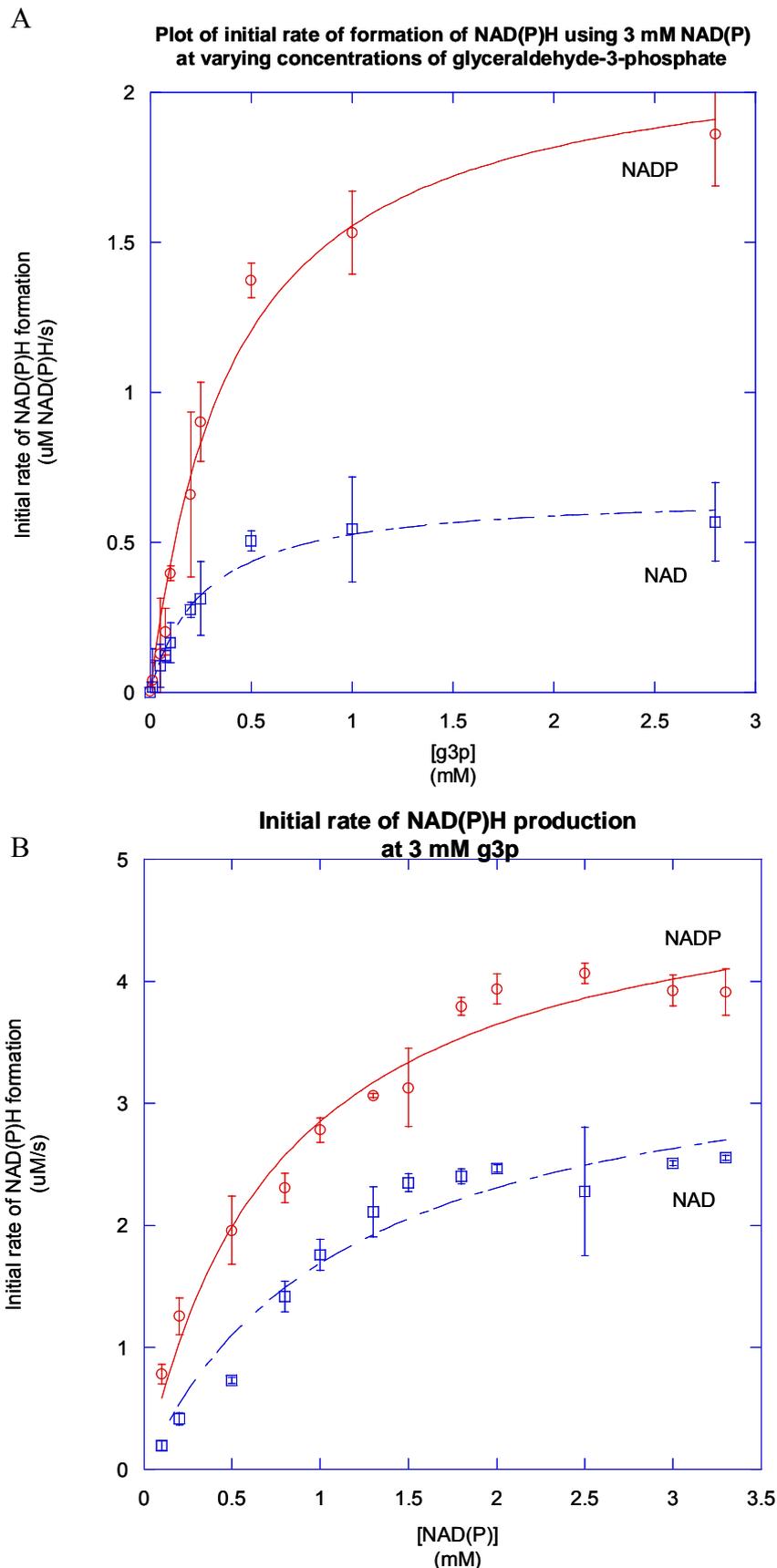


Figure 4. 2 Enzymatic analysis of cGAPDH. Panel A shows the plot of initial rates at a fixed concentration of 3 mM NAD⁺ or NADP⁺ and varying concentrations of g3p. Panel B shows the initial rate plots for a fixed g3p concentration at 3 mM and varying concentrations of NAD⁺ and NADP⁺.

GAPDH	NADP			NAD			NADP ⁺ / NAD ⁺	Reference
	K _M (mM)	k _{cat} (s ⁻¹)	k _{cat} /K _M (mM ⁻¹ s ⁻¹)	K _M (mM)	k _{cat} (s ⁻¹)	k _{cat} /K _M (mM ⁻¹ s ⁻¹)	k _{cat} /K _M ratio	
cGAPDH	0.77 ± 0.13	17.43 ± 0.96	22.6 ± 7.4	1.16 ± 0.30	12.57 ± 1.31	10.8 ± 4.4	2	Chapter 3.3
GAPDHA	0.07 ± 0.01	18.59 ± 0.75	265.6 ± 75.0	1.93 ± 0.29	8.06 ± 0.68	4.2 ± 2.3	63	
<i>A₄ Spinacia oleracea</i> *	0.03	61	2103	0.1	41	410	7	
GAPB <i>Bacillus subtilis</i>	0.9	8.0	9.3	5.7	1.0	0.2	46	(Fillinger <i>et al.</i> , 2000)

Table 4. 3 Comparison between different GAPDHs NADP⁺ and NAD⁺ specificity and that of *H. pylori* GAPDHA (Chapter 3) and *C. jejuni* GAPDH (current chapter).

***For *Spinacia oleracea* the reverse reaction was measured using NADPH and NADH. The difference in specificity is therefore reported in context.**

4.4. Crystallisation of cGAPDH-NADP⁺

Incomplete factorial screening using the commercial Wizard and Cryo I & II (Emerald BioSystems) was employed to determine crystallisation conditions for the cGAPDH in the presence of NADP⁺. Crystallisation trials for initial screening were set up with the Cartesian Honeybee 8 + 1 (Harvard Biosciences) in a humidified chamber, as has been described in Chapter 2.4.1. MRC plates containing two sitting drops were sealed and placed at 4 °C and 22 °C to equilibrate. Plates were monitored for crystal growth by the CrystalProHT (TriTek). Initial screening yielded thin plates as shown in Figure 4.2 in several conditions within the Wizard screen, typically containing 20% (w/v) PEG 1000 (1k) as the precipitant. Refinement of the crystallisation conditions by screening systematically around the pH and precipitant yielded further plates. Further optimisation using different salts from the Ozma 1 and 4k screens (Emerald BioSystems) gave larger crystals, typically out of conditions containing formate, acetate, and tartrate as anions, as shown in Figure 4.3.

Crystals of cGAPDH were prepared for cryocrystallography by transferring them to a cryo mixture of; 15% glycerol, 20% (w/v) PEG1/4k, 100mM acetate pH 4.5, or; 15% glycerol, 20% (w/v) PEG1/4k, 100mM Tris pH 6.5 and frozen in a stream of boiled off liquid nitrogen at 110 K prior to data collection.

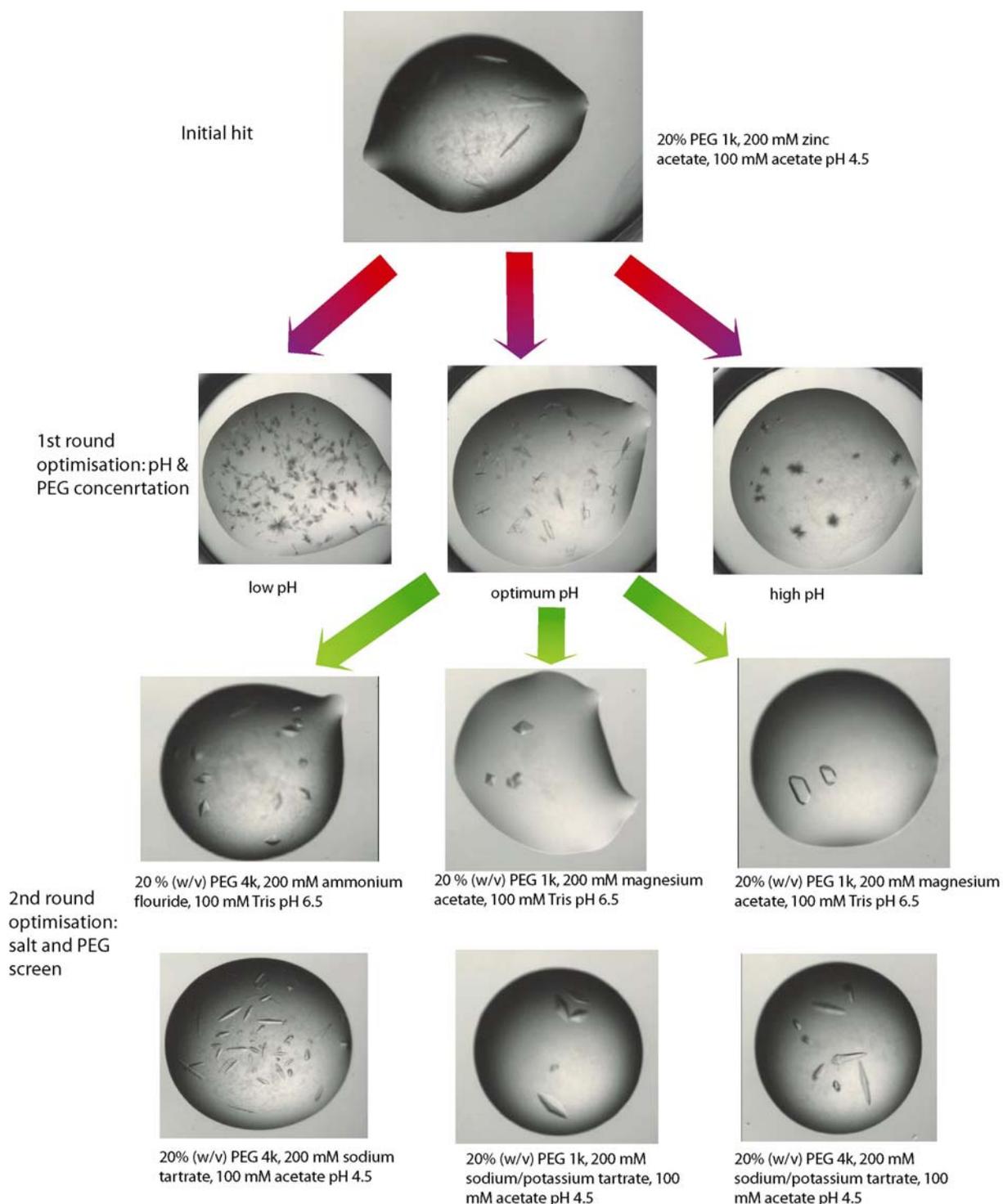


Figure 4. 3 Crystallisation strategy for cGAPDH grown in presence of NADP⁺. Initial spares matrix screening yielded several hits out of the Wizard I & II screens (Emerald Biosciences). Initial crystals were plates and were common in conditions containing PEG 1k. 1st round refinement of these conditions focusing on pH and PEG 1k concentration did not yield any further improvement of the plates, with multiple nucleation a problem at extreme pH ranges. 2nd round optimisation using the Ozma 1 & 4k screens (Emerald Biosciences) screening different salts at 20% (w/v) PEG 1/4k resulted in larger crystals out of conditions predominantly containing formate, acetate and tartrate as the anions.

4.5. Data collection and analysis of holo-cGAPDH

A complete data set was collected from a single cryo-cooled crystal (grown in the presence of NADP⁺) at ID14-3 at the ESRF, Grenoble, with a typical image shown in Figure 4.4. A fixed wavelength of 0.931 Å and an ADSC Q4R detector with a crystal to detector distance of 172.27 mm were used. In total 90 images were collected with a 1° oscillation width.

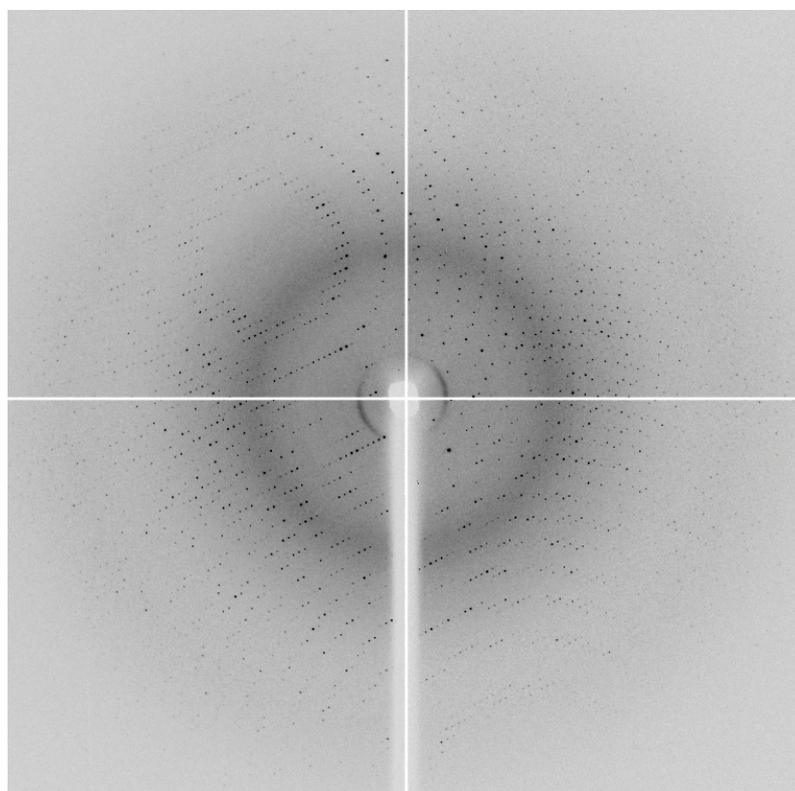


Figure 4. 4 Typical diffraction image of cGAPDH grown in the presence of NADP⁺. 90 images were collected from a single cryo-cooled crystal at 0.931 Å with a crystal to detector distance of 172.27 mm.

Analysis of the autoindexing routines of *MOSFLM* (Leslie, 1992) suggested an tetragonal I-centered cell. Satisfactory scaling was obtained in *I422* using *SCALA* (Evans, 2006).

The Matthews coefficient suggested the asymmetric unit consisted of one monomer of cGAPDH ($V_m = 2.79 \text{ \AA}^3 \text{Da}^{-1}$; $V_s = 55.53 \%$ solvent). Molecular replacement using Chain A of holo-GAPDHA, without NADP^+ and waters, gave a solution consistent with the space group being *I4₁22*. The data was rescaled in *I4₁22*, with an R_{merge} of 13% and an $\{I/(\sigma I)\}$ of 12.9 with the data collection statistics shown below in Table 4.4.

cGAPDH	
Space group	<i>I4₁22</i>
Unit cell parameters (a,b,c, Å)	90.75, 90.75, 225.48
($\alpha = \beta = \gamma$, °)	90.0
Resolution Limits	45.4-1.9 (2.0-1.9)
Number Observations	268069 (38889)
Number Unique Observations	37590 (5379)
Average Redundancy	7.1 (7.2)
Completeness	100.0 (100.0)
($I/\sigma(I)$)	12.9 (4.2)
$R_{\text{merge}}^{\dagger}$	0.13 (0.43)

Table 4. 4 Data collection stastics for cGAPDH. Values in parentheses are for the highest resolution shell. $R_{\text{merge}}^{\dagger} = \frac{\sum_{hkl} \sum_i |I_{hkl,i} - \langle I_{hkl} \rangle|}{\sum_{hkl,i} \langle I_{hkl} \rangle}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices h , k , and l and $\langle I_{hkl} \rangle$ is the mean intensity of that reflection

4.6. Structure determination of cGAPDH

Molecular replacement using Chain A of *H. pylori* GAPDHA from Chapter 3 was used as a search model in *PHASER* (Read, 2001) with waters and NADP⁺ omitted from the structure prior to molecular replacement. As there was only one monomer in the asymmetric unit due to the symmetry imposed by the *I4₁22* space group only one copy of Chain A was searched for with alternative space groups in *I422*. A solution from *PHASER* was found only for space group *I4122*. Manual rebuilding of the chain in *COOT* (Emsley and Cowtan, 2004) and subsequent refinement in *REFMAC5* (Murshudov *et al.*, 1997), refining with anisotropic temperature factors gave a map where electron density for NADP⁺ was clearly defined and could be built into after one cycle of refinement. Waters were added using *COOT* (Emsley and Cowtan, 2004) at positions where peaks were greater than 1.0 σ in the $m2F_o - F_c$ map and greater than 3.0 σ in the $mF_o - F_c$ difference map. As is shown in Figure 4.5, density corresponding to NADP⁺ is apparent after molecular replacement, in particular the phosphates, after several rounds of refining and rebuilding the NADP⁺ is clearly visible. The final R factor was 18% with an R_{free} of 22%, as shown in Table 4.5. From Table 4.5 it is clear that the B-factor for the NADP⁺ is lower than the main chain atoms. This low B-factor reflects how well NADP⁺ is bound in the crystal structure.

	holo-cGAPDH
Resolution range (Å)	35.7-1.9
$\%R$	18
R_{free}	22
RMSD bond lengths (Å)	0.012
RMSD bond angles (°)	1.46
No. protein atoms	2945
No. solvent atoms	391
Average B-factors (Å ²)	
Overall	22.0
Main chain atoms	10.4
Side chain atoms	12.0
Water molecules	25.2
NADP	9.9

Table 4. 5 Refinement statistics for holo-cGAPDH. $\%R = \frac{\sum_{\text{hkl}} ||F_o(\text{h, k, l}) - |F_c(\text{h, k, l})||}{\sum_{\text{hkl}} |F_o(\text{h, k, l})|}$. R_{free} is defined as R calculated on a random 5 % non refined data.

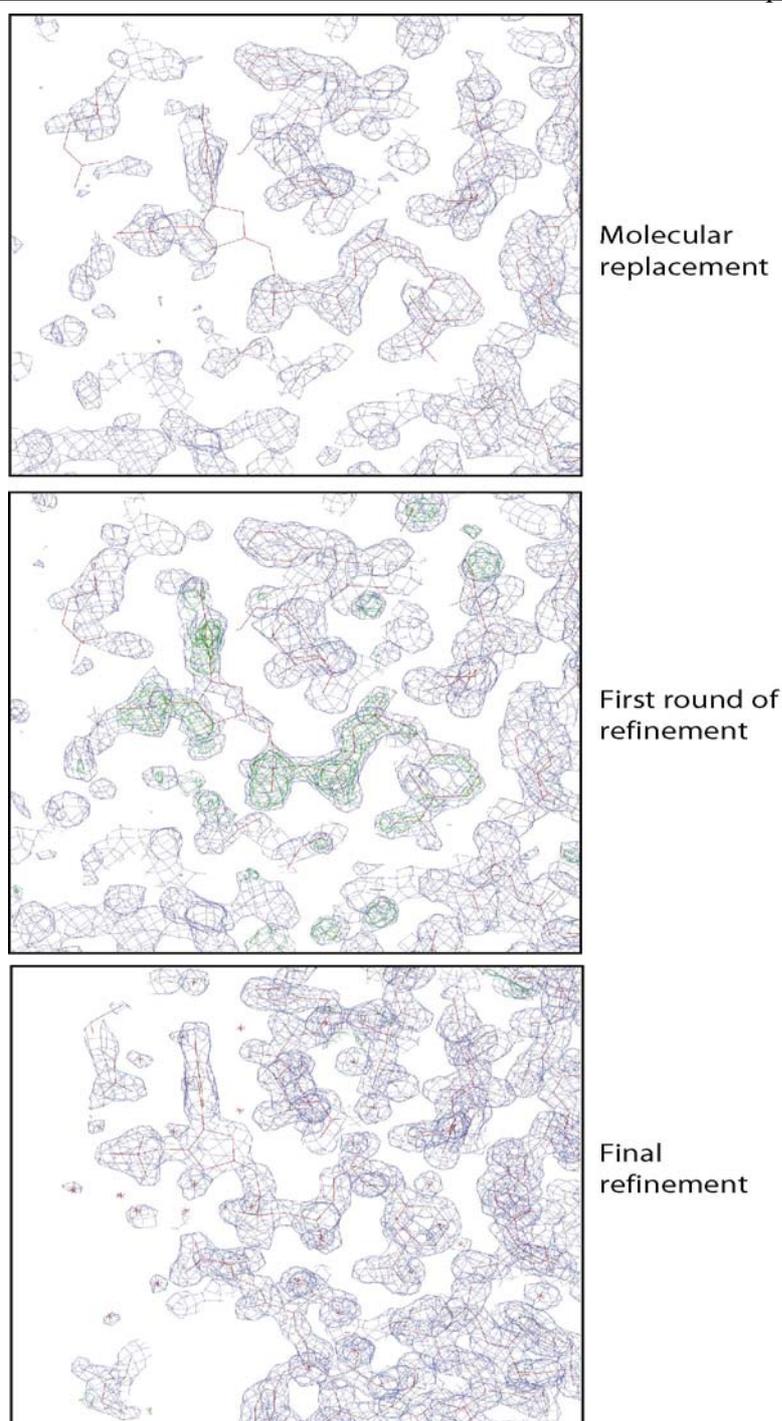


Figure 4. 5 Progress of refinement can be seen by following the development of the electron density for NADP^+ and the surrounding regions. After molecular replacement the phosphates for NADP^+ are apparent. The first round of refinement with five cycles in *REFMAC5* results in the nicotinamide moiety becoming distinguishable. Further refinement and model building in *COOT* results in clearly defined electron density for NADP^+ and the surrounding region, with waters now apparent. A $m2F_o-F_c$ map contoured at 1.2σ is shown in blue. In the first round and final refinement a mF_o-F_c difference map contoured at 3σ is shown in green. The protein is shown in red lines as is NADP^+ . NADP^+ is shown for clarity in the molecular replacement and first refinement map and was only built into the model during the latter stages of refinement.

4.7. Stereochemical analysis of cGAPDH model

Analysis of the stereochemical quality of the holo-cGAPDH model was carried out using *PROCHECK* (Laskowski *et al.*, 1993) and MolProbity (Davies *et al.*, 2007). Overall, the RMSD from standard bond lengths was 0.012 Å with an RMSD from standard bond angles of 1.46 °. The average B factor for the molecule is 22.0 Å² with 10.4 Å² for the main chain atoms, 12.0 Å² for the side chain atoms and 25.2 Å² for all 391 solvent molecules and 9.9 Å² for the NADP⁺, as summarised in Table 4.5. The Ramachandran plot shown in Figure 4.6 shows that 97.6 % of residues fall into the favoured region with 99.7 % of residues in the allowed region. Valine 237 which is in a disallowed region is shown in Figure 4.6 with clearly defined electron density.

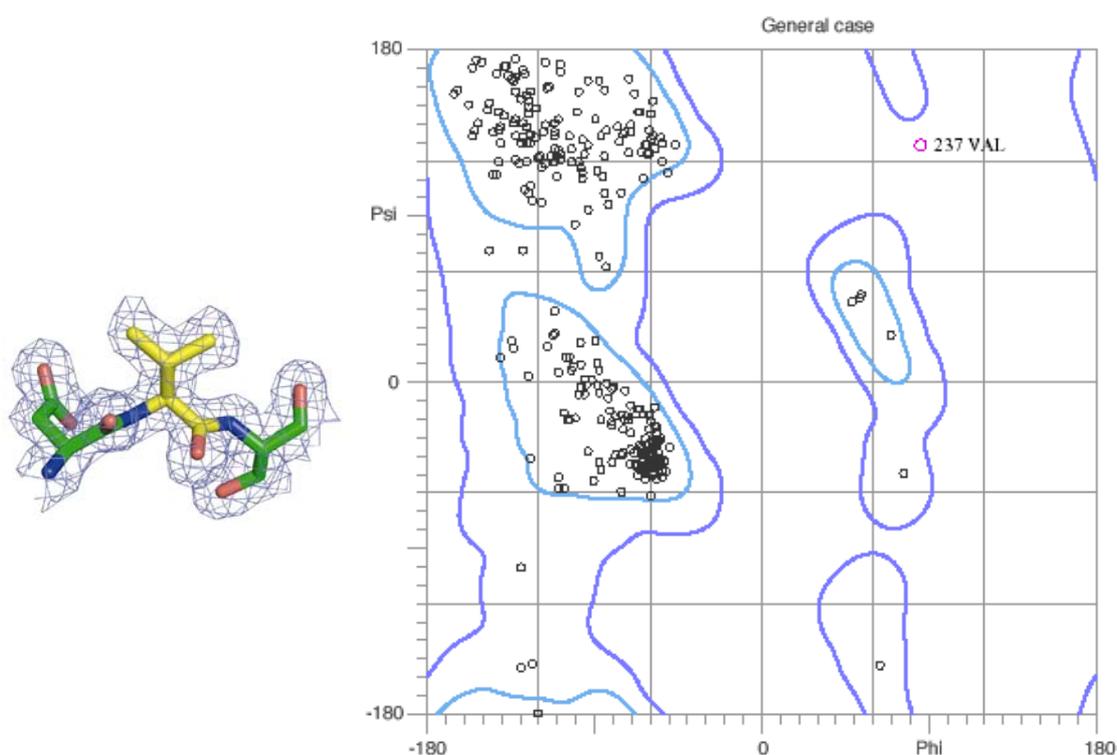


Figure 4. 6 Ramachandran plot for holo-cGAPDH generated from Molprobity (Davies *et al.*, 2007). The only residue as an outlier is Val237, shown to the left in yellow. The tripeptide is enclosed in a $2F_o-F_c$ map contoured at 1.0σ and shows unambiguous electron density about the peptide bond.

4.8. Structural analysis of cGAPDH

4.8.1. Overall Structure

cGAPDH is the first structure of a glyceraldehyde-3-phosphate dehydrogenase to be crystallised in an *I* point group. Owing to the cell dimensions only one monomer is present within the asymmetric unit. The monomeric structure of cGAPDH is shown in Figure 4.7 with the NADP⁺ binding domain and the catalytic domain depicted in red and green respectively and the NADP⁺ molecule shown as ball and stick, as for apo- and holo-GAPDHA in Chapter 3 and holo-GAPDHB in Chapter 5.

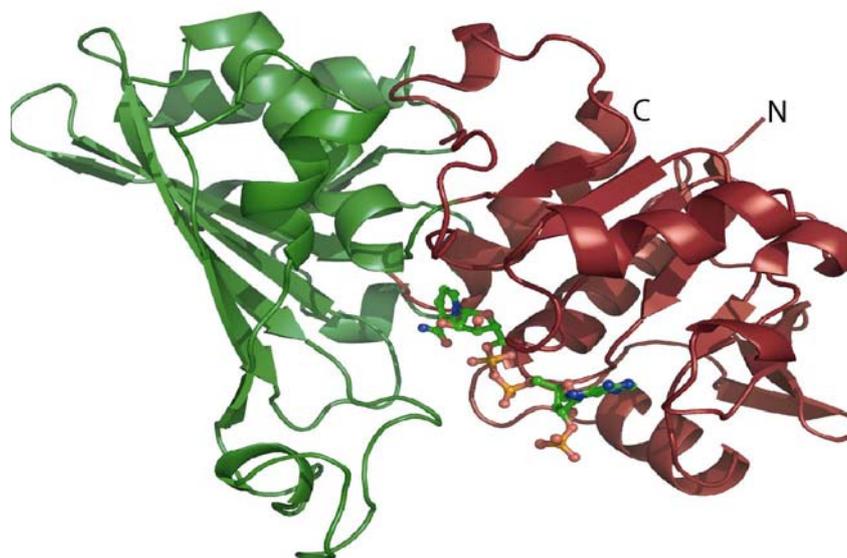


Figure 4. 7 Structure of the monomeric form of cGAPDH bound to NADP⁺. The NADP⁺ binding domain, corresponding to residues 1-147 and 312-328 is depicted in red and the catalytic domain, corresponding to residues 148-311 is depicted in green. As for apo and holo-GAPDHA in Chapter 3. NADP⁺ is shown as ball and stick.

A complete tetramer displaying 222 symmetry can be built from the symmetry related molecules within the lattice and this is shown in Figure 4.8. As is shown in Figure 4.8B the S-loop of the R-axis-related monomer coordinates the NADP⁺ molecule through Lys189 in the S-loop.

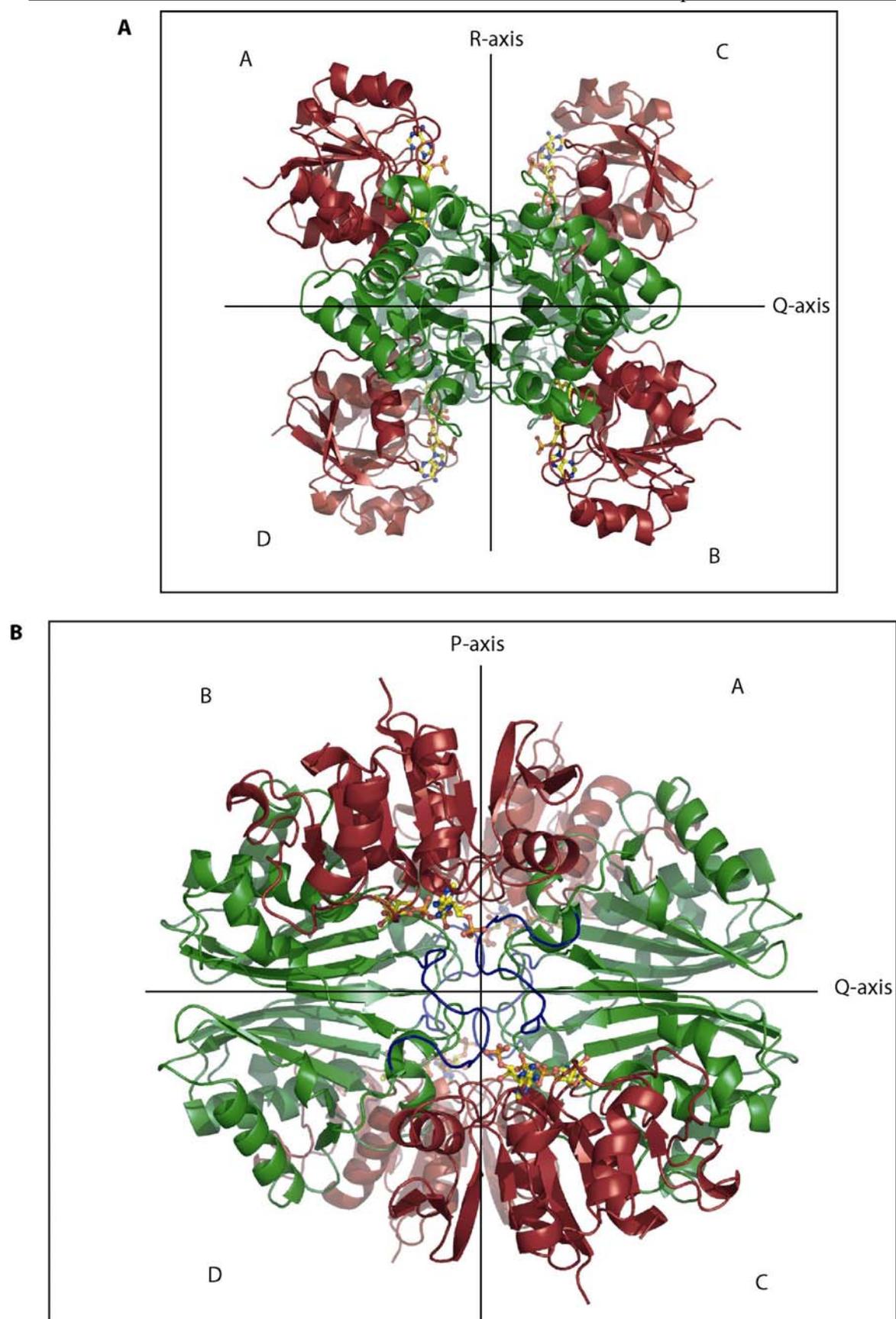


Figure 4. 8 Tetrameric arrangement of holo-cGAPDH. Representation of the catalytic and NADP⁺-binding domains is the same as Figure 4.2. Panel A shows the view of the tetramer along the P-axis and panel B shows the view along the R-axis. As in Figure 3.18 the S-loop of the R-axis related subunit, shown in blue, interacts with the NADP⁺.

4.8.2. Active site

The active site of cGAPDH is similar to that of other GAPDHs with regards to the presence of the catalytic Cys150 interacting with His177. Figure 4.9 shows the active site of cGAPDH with the nicotinamide being orientated in the *syn* conformation due to steric hindrance from Tyr317 and the internal hydrogen bond between the O1N and N7N of NADP⁺.

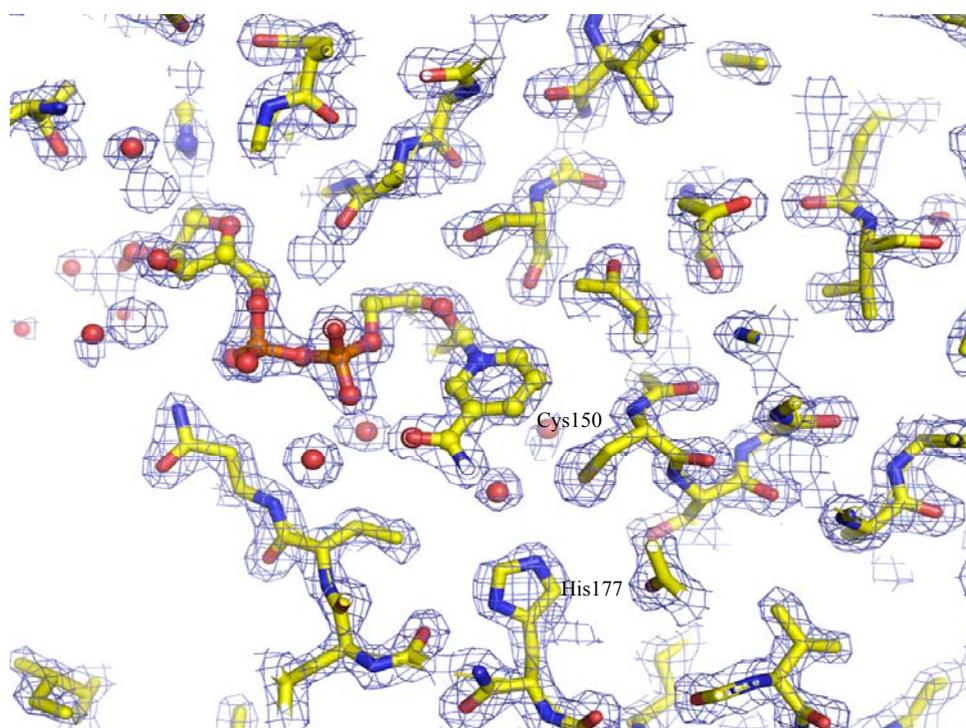


Figure 4. 9 Active site of holo cGAPDH showing the nicotinamide moiety of NADP⁺ and the active site Cys150 and His177 enclosed in an $2F_o-F_c$ map contoured at 1.4σ .

Figure 4.9 shows the active site waters and the coordination of the His177 which facilitates in increasing the acidity of Cys150 by proton extraction as discussed in Chapter 1.8.1.1.2. Several waters also constitute the active site with waters 580, 310 and 675 all being in a tetrahedral geometry approximately 3.2 Å from the Cys150 SG, which is the correct bond length for a sulphur-hydrogen bond, shown in Figure 4.10.

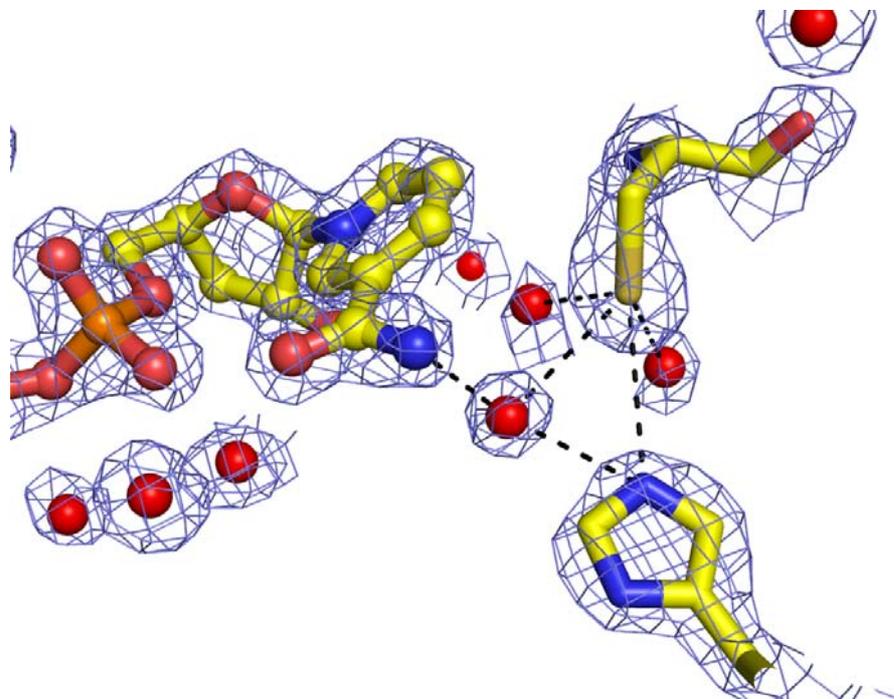


Figure 4. 10 Coordination of solvent atoms within the active site of cGAPDH. The nictinamide moiety and the active site Cys150, His177 and waters are enclosed in an $2F_o-F_c$ map contoured at 1.4σ . Interactions are only shown from the water and His177 to the cystine to highlight the tetrahedral arrangement.

Residue	Distance (Å)	NADP ⁺ atom
Arg78 NH2	2.7	O1X
Arg78 NH1	3.4	O3X
H ₂ O 695	2.7	O1x
Arg 78 NH2	3.3	N7A
H ₂ O 704	2.7	O2X
Arg 12 N	3.0	O2A
Ile 13 N	2.8	O1N
Asn 181 ND2	2.7	O1A
H ₂ O 352	2.7	O3B
Asn 313 OD1	2.8	N7N
H ₂ O 584	2.7	O2N

Table 4. 6 Distances of residues involved in coordinating NADP⁺ in holo cGAPDH.

4.8.3. NADP⁺ Binding site

The binding site for NADP⁺ in cGAPDH is similar to that of GAPDHA from *H. pylori* and other GAPDHs with the adenosine ring in the *anti* conformation and the N6 of the adenosine ring hydrogen bonding to the carbonyl of Arg78. In holo-GAPDHA Arg77 did not stack against the adenosine ring. However, it is clear that in holo-cGAPDH Arg78 stacks against the adenosine ring stabilising its *anti* conformation. Again the characteristic nucleotide binding motif; G-X₁-G-X₂-X₃-G forms a hydrogen bonding network from the backbone of the protein across to the NADP⁺. Figure 4.11A shows the main bonding network between the NADP⁺ molecule and cGAPDH with specific interactions shown in Table 4.6. Figure 4.10B shows the key residues responsible for interacting with the 2'-phosphate of NADP⁺. As is clear in Figure 4.11B Thr35 does not coordinate the 2'-phosphate of NADP⁺ with electron density for all the functional groups in this region being unambiguous.

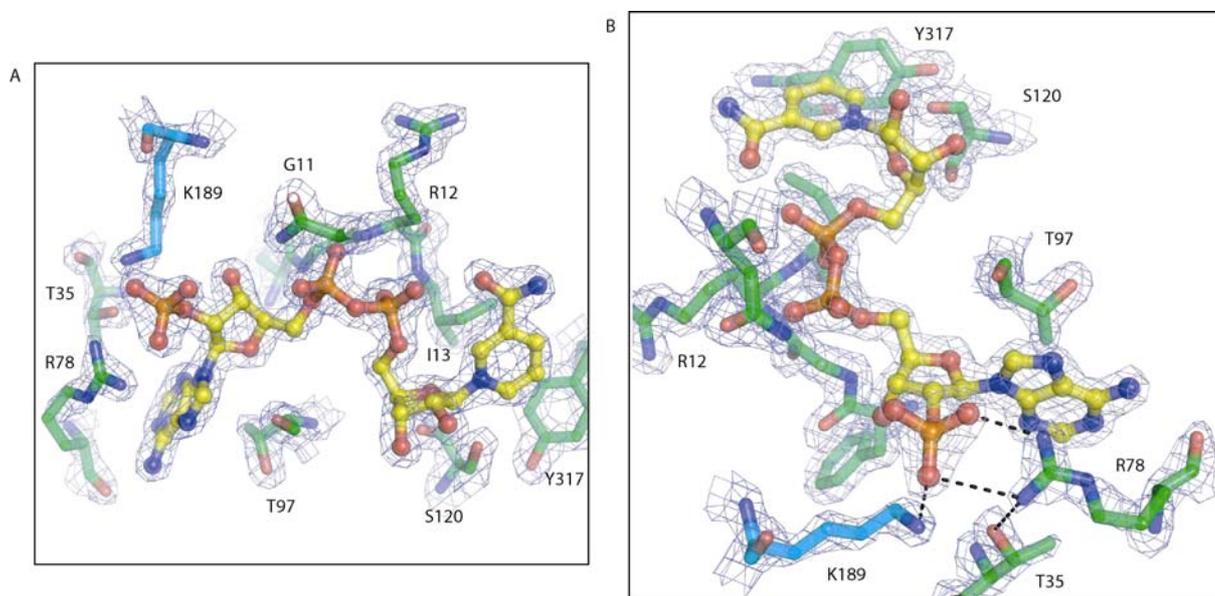


Figure 4. 11 Key residues binding to NADP⁺. Panel A shows the extended conformation of NADP⁺ highlighting key interactions of the nucleotide binding motif G₉-F₁₀-G₁₁-R₁₂-I₁₃-G₁₄. Panel B is the view from rotation about the P2B phosphate showing interactions from the R-axis related K189 (shown in blue) and from R78 to the 2'-phosphate. T35 although not interacting with the phosphate forms a hydrogen bond to R78 and is shown in comparison to holo-GAPDHA described in Chapter 3.9.3.

4.9. Analysis of holo-cGAPDH and GAPDHA

In the previous Chapter, (Chapter 3.12) sequence analysis of other GAPDHs within the *Campylobacterales* orders was used to identify other putative NADP⁺-dependent GAPDHs. *C. jejuni* was one such species identified and was of interest because of the organism's similar habitat and streamlined genome like *H. pylori*. Unlike *H. pylori* however, which has two GAPDH genes; *gapA* (examined in Chapter 3) and *gapB* (examined in Chapter 5), *C. jejuni* only has one gene encoding a GAPDH, which has been the subject of investigation for this Chapter. Enzymatic analysis in Chapter 4.3 showed that cGAPDH has a lower specificity for NADP compared to GAPDHA (Table 4.3). Analysis of the crystal structure of cGAPDH (Chapter 4.8.3) has shown that the "determinant" residue for specifying NADP⁺-dependence (identified in Chapter 3.10), Thr35, does not coordinate the 2'-phosphate of NADP⁺. As is shown in Figure 4.12, superimposition of the NADP⁺ from holo-GAPDHA and cGAPDH reveals that there is a 2.3 Å difference between the positions of the two threonines and this extra distance prevents Thr35 of cGAPDH coordinating the 2'-phosphate. Instead Thr35 forms a salt-bridge between Lys189 of the R-axis related subunit and Arg78. This could explain the reduced specificity between NADP⁺ and NAD⁺. As would be expected a strictly NAD⁺-dependent GAPDH would not be able to bind NADP⁺ due to the steric and charge clash between the 2'-phosphate on NADP⁺ (as demonstrated in Chapter 5.6.5). However, a NADP⁺-dependent enzyme would be able to accommodate NAD⁺, albeit at a lower affinity. A phosphate binding pocket would result in tighter binding of the 2'-phosphate from NADP⁺. The slight loss of enthalpy for NAD⁺ as a result of not occupying the phosphate binding pocket would result in its looser binding.

To determine the thermodynamics of binding other experiments would need to be considered, the easiest being isothermal titration calorimetry (ITC), which would give an enthalpy of binding of NADP⁺ and NAD⁺. This experiment which was not conducted during these studies owing to lack of sample, would also give an idea of cooperativity of cGAPDH with NADP⁺ and NAD⁺.

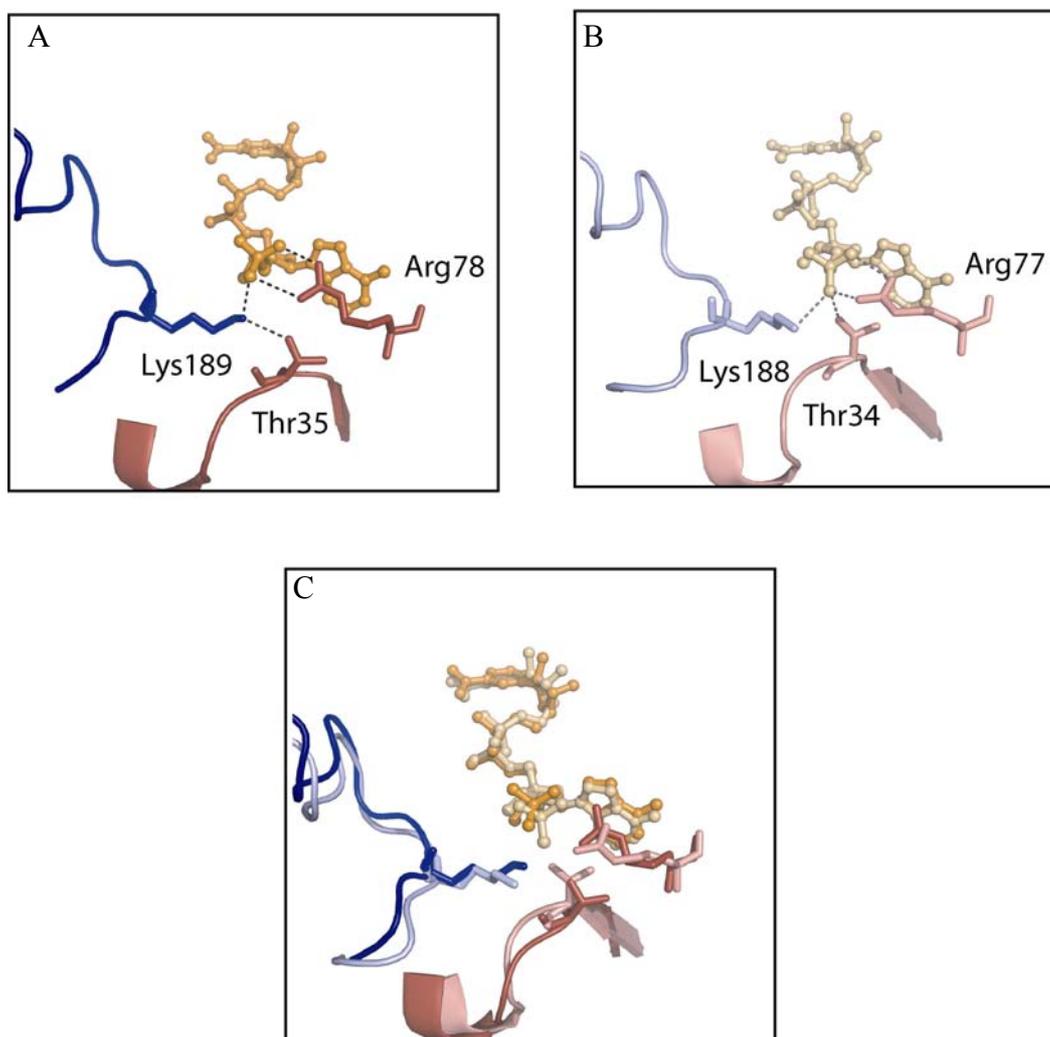


Figure 4. 12 Comparison of the NADP⁺ coordination by cGAPDH (red and blue) and holo-GAPDHA (salmon and light blue). Panel A shows the coordination of the 2' phosphate by cGAPDH, highlighting the lack of direct coordination by Thr35. For reference, panel B shows the corresponding residues of holo-GAPDHA coordinating the 2' phosphate of NADP⁺. Panel C is a least squares fit of the NADP⁺ of cGAPDH and GAPDHA by *LSQKAB* (Kabsch, 1976).

4.10. Conclusion

This part of the study aimed to ascertain the specificity for *C. jejuni* GAPDH based upon sequence analysis in Chapter 3 that suggested that there were other species within the *Campylobacterales* order that were NADP⁺-dependent. As discussed in Chapter 1. 3.2 and the start of this Chapter, *C. jejuni* is similar to *H. pylori* with respect of the organisms colonisation of gastrointestinal tracts. The metabolic pathways of *C. jejuni* and *H. pylori* appear similar, although there are a few subtle differences. Both genomes have undergone a process of reductive evolution, where genes not required for survival of the organism have been deleted (Andersson and Andersson, 1999). There is a complete gluconeogenic pathway but glycolysis is incomplete (Parkhill *et al.*, 2000) as in *H. pylori*, again confirming these two organisms lack of requirements for carbohydrates as the main energy source. The surprising feature is only one gene encoding a GAPDH, it is feasible that another enzyme may be encoded that has GAPDH activity but does not share any sequence similarity, this could only occur if it proceeded via a different mechanism to bacterial GAPDHs i.e. a ferredoxin mechanism as described for an *Archaeal* GAPOR (Van der Oost *et al.*, 1998). The dual specificity for cGAPDH may therefore be utilised by the bacterium for switching between the NADP⁺/NAD⁺ ratios. As described in Chapter 3.11, a NADP⁺-dependent GAPDH may function towards a gluconeogenic pathway, providing key five-carbon sugars for nucleotide biosynthesis. On the other hand NADPH could be the main electron donor, over NADH, into the electron transport chain of the bacterium. Therefore, any metabolism of glucose which would proceed through the Entner-Doudoroff pathway could generate the reductive power for ATP production or simply for biosynthesis, (although the latter is provided from the Pentose Phosphate pathway). It is plausible that NADPH is associated with the gluconeogenic pathway whilst NAD⁺ activity with the glycolytic pathway once amino acids have become depleted, forcing the metabolic flux in the opposite direction, hence glucose utilisation. As described in Chapter 7, the determination as to whether the *cgap* gene is expressed in either gluconeogenic or glycolytic conditions warrants further investigation. The remainder of this study will focus on the other GAPDH-encoding gene from *H. pylori*, *gapB*.

**Chapter 5. Structure of Glyceraldehyde-3-Phosphate
Dehydrogenase B**

5.1. Introduction

H. pylori has two genes encoding GAPDH, like other members of the *Campylobacterales* order, with the exception of *C. jejuni* discussed in the previous chapter. Other groups of bacteria tend to have two genes encoding for GAPDH. Studies on the enzymatic nature of both these enzymes in *E. coli* has met with conflicting results (Boschi-Muller *et al.*, 1997, Zhao *et al.*, 1995). In Chapter 3 it was shown that one of these genes within *H. pylori*, termed *gapA*, expresses a phosphorylating NADP⁺-dependent GAPDH and it was proposed that GAPDHA is involved in a gluconeogenic role within the bacterium, or is catalysing the oxidative reaction yielding NADPH which could be the main reductive coenzyme for the bacterium. The other gene, *gapB* has been the subject of investigation for this chapter. Putative sequence analysis of GAPDHB in Chapter 3.12, suggested that GAPDHB is likely to be NAD⁺-dependent. This chapter describes the cloning, expression and crystallisation of GAPDHB. The crystal structure of GAPDHB has been obtained in the presence of NAD⁺, confirming the predictions made in Chapter 3 regarding coenzyme specificity. Enzymatic analysis of GAPDHB shows it has activity towards glyceraldehyde-3-phosphate as a substrate and in addition, activity towards erythrose-4-phosphate. The enzymatic assay is consistent with GAPDHB displaying a non-phosphorylating erythrose-4-phosphate dehydrogenase activity. This type of activity has been implied for one of the *gap* genes identified in *E. coli* (Zhao *et al.*, 1995). This raises further interesting questions with regards to the metabolism within *H. pylori*. In addition, since the corresponding homologue is absent in *C. jejuni*, further questions about the function of GAPDHB are raised.

5.2. GAPDHB cloning and expression

The cloning of the *gapB* gene from *H. pylori* follows the same process as in GAPDHA and cGAPDH (Chapters 3.2 and 4.2 respectively) and is described in Chapter 2.7. Briefly, the gene encoding GAPDHB, *gapB* was cloned out of the genomic DNA of *H. pylori* into the pET151/D vector (Invitrogen) resulting in the addition of a hexa-histidine (His₆) affinity tag and a TEV cleavage site. His₆-GAPDHB was purified from the *E. coli* expressing strain *Rosetta DE3* transformed with the pET151/D-*gapB* vector. Purification of His₆-GAPDHB was achieved through nickel affinity chromatography as described in Chapter 2.7.2. A typical chromatogram is shown in Figure 5.1.

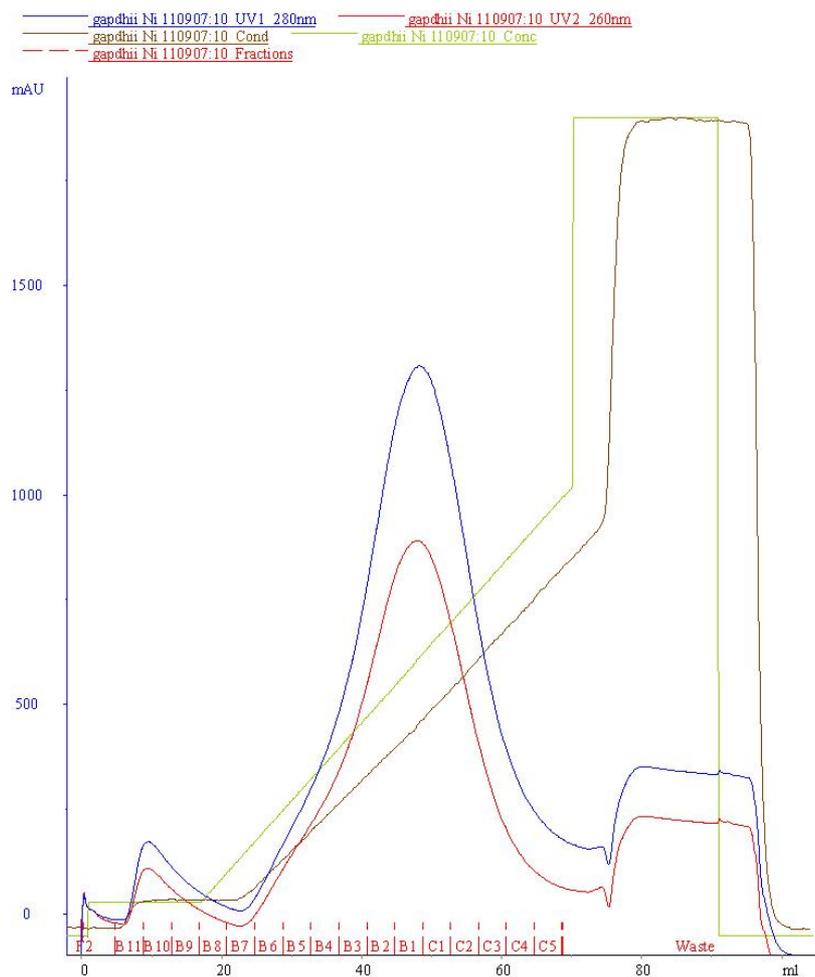


Figure 5. 1 Typical chromatogram for the purification of His₆-GAPDHB. The absorption at 280 and 260 nm (blue and red respectively) of protein is shown plotted against volume of elution with a gradient of buffer containing 500 mM imidazole (represented in green) (as described in Chapter 2.7.2). The corresponding fractions are also shown. Fractions B2-C3 contained His₆-GAPDHB were pooled.

5.3. Crystallisation of GAPDHB

Crystallisation trials were performed with the hexa-histidine tag still present on GAPDHB (as for holo-GAPDHA). Sparse matrix screening using the commercial screens Wizard and Cryo I & II screens (Emerald Biosciences) were setup as has been described in Chapters 3.4 and 4.4. Briefly, 100 nl drops were dispensed with a Genomics Solutions Cartesian Honeybee 8+1 (Harvard Bioscience) onto 96-well MRC plates (Innovadyne) with reservoirs containing 80µl of precipitant in a humidity chamber. Plates were sealed with transparent tape and monitored for crystal growth using CrystalProHT (TriTek) plate storage and imaging systems at 277 K and 293 K. Crystals appeared after four days and were predominately from the Cryo I and II screens with a preference towards lower molecular weight PEGs as precipitants. Crystallisation conditions were optimised using a Tecan 75 (Tecan) liquid-handling robot and 500 nl drops of protein and precipitant were dispensed from the Cartesian Honeybee (Harvard Bioscience). Suitable crystals of diffraction quality grew from a protein concentration of 8 mgml⁻¹ His₆-GAPDHB containing 1 mM NAD⁺ mixed with an equal volume of precipitant; 100 mM sodium acetate pH 4.0 and 38% (v/v) 2-methyl-2,4-pentanediol equilibrated at 293 K for five days. Typically two distinct crystal forms were found as shown in Figure 5.2, with crystal form A generally being diffraction quality crystals, whilst crystal form B, suffering from nucleation at the ends of the crystal did not diffract beyond 11 Å and no further information were obtained on these crystals.

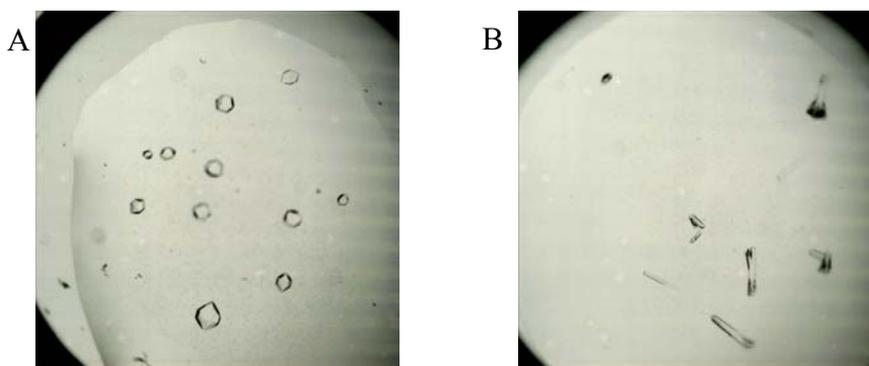


Figure 5. 2 Crystals of His₆-GAPDHB grown in the presence of 1 mM NAD⁺ obtained from screening around 100 mM sodium acetate pH 4.0, 38% (v/v) 2-methyl-2,4-pentanediol. Two distinct crystal forms were found, with crystal form A, generally yielding diffraction quality data beyond 4 Å, whilst crystal form B did not diffract beyond 11 Å.

5.4. Data collection and processing

Crystals were prepared for cryocrystallography by harvesting the crystals directly into a stream of boiled-off liquid nitrogen at 110 K with the precipitant acting as a suitable cryoprotectant.

Diffraction data were collected on the ID14-3 beamline at the ESRF, Grenoble at 0.931 Å on an ADSCQ4R CCD detector. A complete data set was collected from a single cryo-cooled crystal, with a crystal to detector distance of 269.6 mm with images collected at a 1° oscillation width over 120°, a typical diffraction image shown in Figure 5.3. Intensities were measured using *MOSFLM* (Leslie, 1992) with the autoindexing routines giving a solution consistent with a primitive hexagonal cell. The Matthews coefficient suggested the presence of a tetramer within the asymmetric unit ($V_m = 3.4 \text{ \AA}^3 \text{ Da}^{-1}$; $V_s = 63 \%$ solvent). The data were initially scaled using *SCALA* (Evans, 2006) consistent with the Laue group 622. Analysis of the distribution of intensities along the principle axes indicated the presence of either a 6_1 or a 6_5 screw. Molecular replacement using GAPDH from *Bacillus stearothermophilus*, 1GD1 (Skarzynski *et al.*, 1987) as a search model in *PHASER* (Read, 2001) gave a solution consistent only with space group $P6_522$. The data was reprocessed in this space group and then Scaled using *SCALA* giving the statistics shown in Table 5.1.

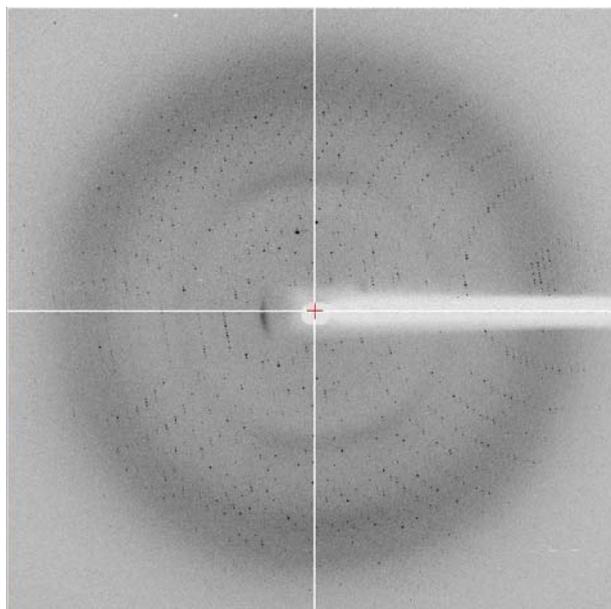


Figure 5.3 Typical diffraction image of His₆-GAPDHB grown in the presence of NAD⁺. A complete data set for 120° was collected at an oscillation width of 1° and a crystal to detector distance of 269.6 mm for a single cryo-cooled crystal at 0.931 Å.

Space group	P6 ₅ 22
Unit cell parameters (a,b,c, Å)	116.09,116.09, 253.11
($\alpha = \beta = \gamma$, °)	90.0, 90.0, 120.0
Resolution Limits	83.1-2.8 (2.95-2.80)
Number Observations	393489 (57245)
Number Unique Observations	50370 (7302)
Average Redundancy	7.8 (7.8)
Completeness	98.4 (99.6)
{I/ σ (I)}	17.0 (6.1)
[†] R _{merge}	0.094 (0.32)

Table 5. 1 Data collection stastics for His₆-GAPDHB grown in the presence of NAD. Values in parentheses are for the highest resolution shell. [†]R_{merge} = $\frac{\sum_{hkl} |\sum_i I_{hkl,i} - \langle I_{hkl} \rangle|}{\sum_{hkl,i} \langle I_{hkl} \rangle}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices h , k , and l and $\langle I_{hkl} \rangle$ is the mean intensity of that reflection

5.5. Structure determination of GAPDHB

Molecular replacement using GAPDH from *Bacillus stearothermophilus*, 1GD1 (Skarzynski *et al.*, 1987) with NAD⁺ and the waters omitted, as a search model in *PHASER* (Read, 2001) gave a solution with a LLG of 7621 as the top score. Refinement using *REFMAC5* (Murshudov *et al.*, 1997) was carried out with a random 5% of the reflections selected for the R_{Free}. After one cycle of refinement, the R factor dropped from 52 % to 38 %. Inspection of the electron density map showed the presence of NAD⁺ bound and NAD⁺ was built in and refined. Manual rebuilding of the sequence of the model was carried out using *COOT* (Emsley and Cowtan, 2004) with subsequent rounds of refinement in *REFMAC5* (Murshudov *et al.*, 1997), refining with overall B-factors due to the lower resolution of the data. After the sequence and geometry of the model had been corrected waters were added using *COOT* (Emsley and Cowtan, 2004) at positions where peaks were greater than 1.0 σ in the m2F_o-F_c map and greater than 3.0 σ in the mF_o-F_c difference map. Waters were manually inspected and removed if they were greater than 3.5 Å from hydrogen bond donors or acceptors. The final R factor for the model was 21% with a R_{Free} of 26%, the refinement stastics are summarised in Table 5.2.

	holo-GAPDHB
Resolution range (Å)	69.5-2.8
$^{\S}R$	21
R_{free}	26
RMSD bond lengths (Å)	0.011
RMSD bond angles (°)	1.523
No. protein atoms	10558
No. solvent atoms	169
Average B-factors (Å ²)	
Overall	33

Table 5. 2 Refinement statistics for holo-GAPDHB. $^{\S}R = \frac{\sum_{\text{hkl}} ||F_o(\mathbf{h}, \mathbf{k}, \mathbf{l})| - |F_c(\mathbf{h}, \mathbf{k}, \mathbf{l})||}{\sum_{\text{hkl}} |F_o(\mathbf{h}, \mathbf{k}, \mathbf{l})|}$. R_{free} is defined as R calculated on a random 5 % non refined data.

5.6. Structure analysis of holo-GAPDHB

5.6.1. Stereochemical analysis of GAPDHB

Analysis of the stereochemical quality of the holo-GAPDHB model was performed using *PROCHECK* (Laskowski *et al.*, 1993) and MolProbity (Davies *et al.*, 2007). The RMSD from standard bond lengths is 0.011 Å and 1.52 ° from the angles. The Ramachandran plot shows that 99.7% of residues are in the allowed region with 95% of residues in the favoured region, only one residue was an outlier, Valine 239, observed as an outlier in other GAPDHs, such as *C. jejuni* (Chapter 4.7). Figure 5.4A shows the Ramachandran plot for Chain A taken from MolProbity, with the only outlier, Val239 shown in Figure 5.4B, shown contoured with unambiguous electron density. In all there was one residue within each chain, which had no interpretable electron density, which was Serine 71 that forms part of a surface exposed loop in the NAD⁺ binding domain.

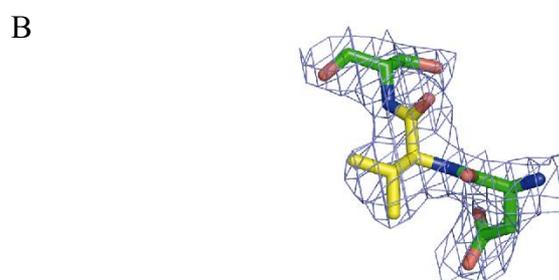
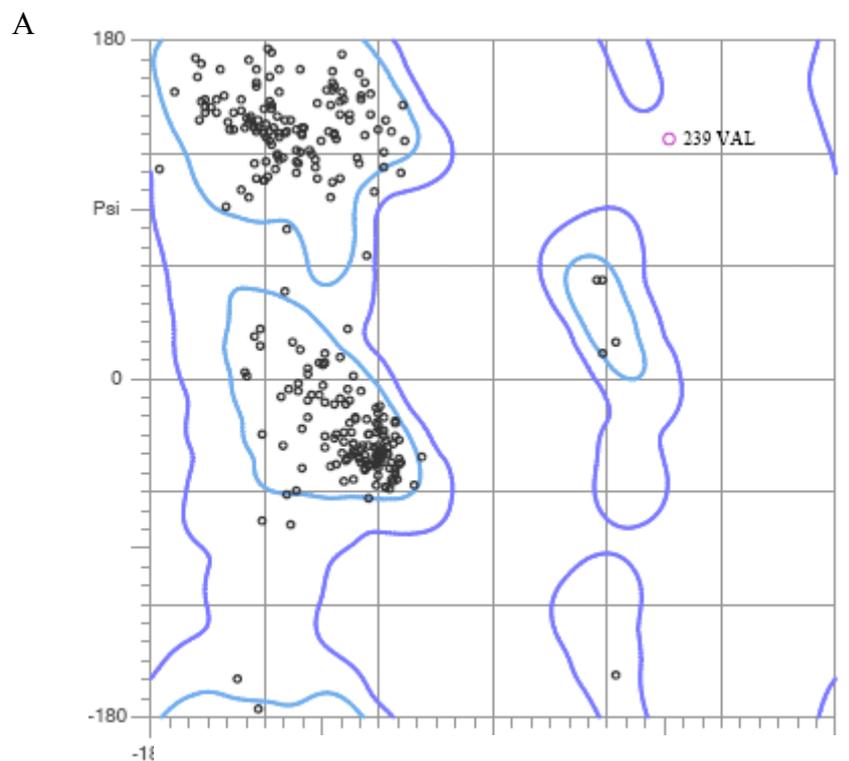


Figure 5. 4 A, Ramachandran plot for holo-GAPDHB, showing only Chain A for clarity. Val239 is also the only outlier in Chains B, C and D. B, The Ramachandran outlier, Val239 is shown in yellow within a $2F_0 - F_c$ map contoured at 1σ .

5.7. Overall Structure

The structure of holo-GAPDHB is a tetramer displaying 222 symmetry, as has been observed for GAPDHA and cGAPDH in Chapters 3 and 4 respectively. The structures of the monomer and tetramer are shown in Figure 5.5.

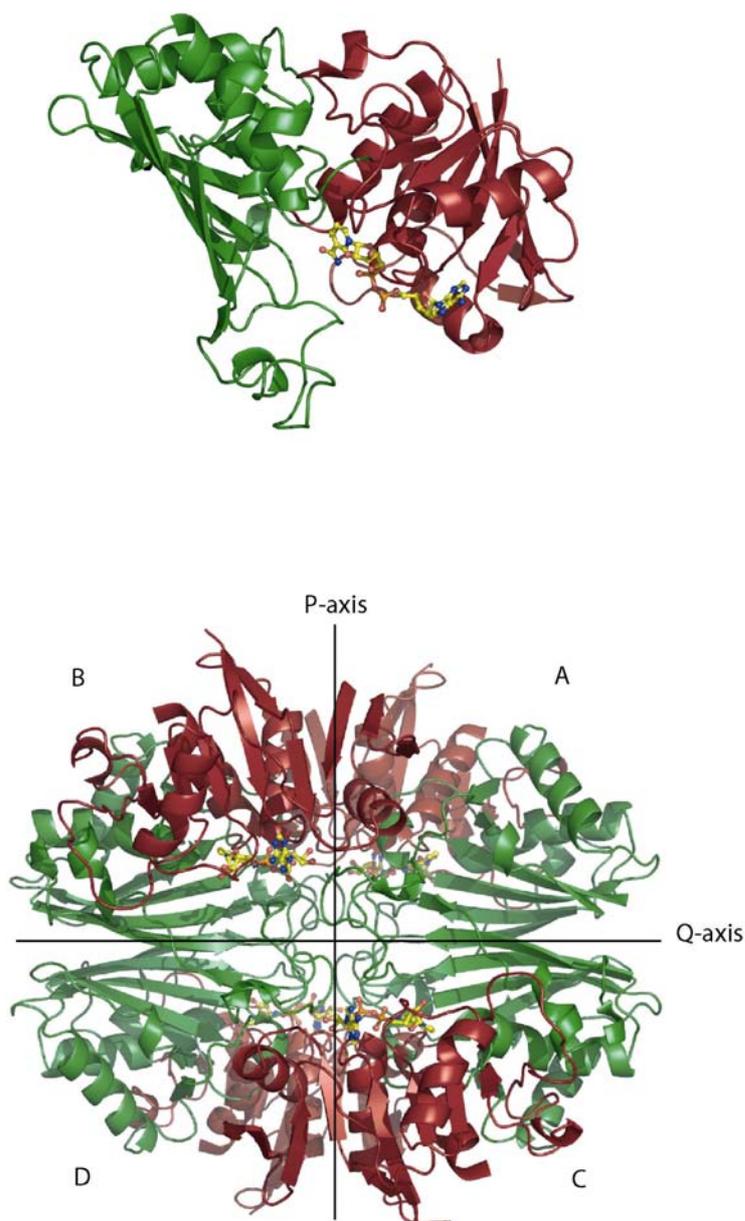


Figure 5. 5 Monomeric and tetrameric representation of holo-GAPDHB. A shows Chain A with the NAD⁺ domain, corresponding to residues 1-147 and 313-328 depicted in red and the catalytic domain, corresponding to residues 148-312 in green. B shows the view of the tetramer along the R-axis.

5.8. N-terminal hexa-histidine linker

Crystals of holo-GAPDHB were obtained with the hexa-histidine tag still present. As was seen for holo-GAPDHA in Chapter 3.9.1 residues encoding the TEV site are stabilised between the symmetry-related tetramer within the unit cell. Analysis of the electron density of the holo-GAPDHB model was consistent with 11 residues of the TEV linker protruding from the N-terminus of Chain D. No extra density was observed from Chains A, B and C. Figure 5.6 shows the secondary structure of the TEV linker, superimposed onto the TEV linker from Chain A of holo-GAPDHA. As can be seen in Figure 5.6, the secondary structure of the linker is the same for holo-GAPDHB as it is for holo-GAPDHA. Unlike holo-GAPDHA, only one Chain of holo-GAPDHB has a structured TEV site, and this is related to the greater disorder of the GAPDHB molecule.

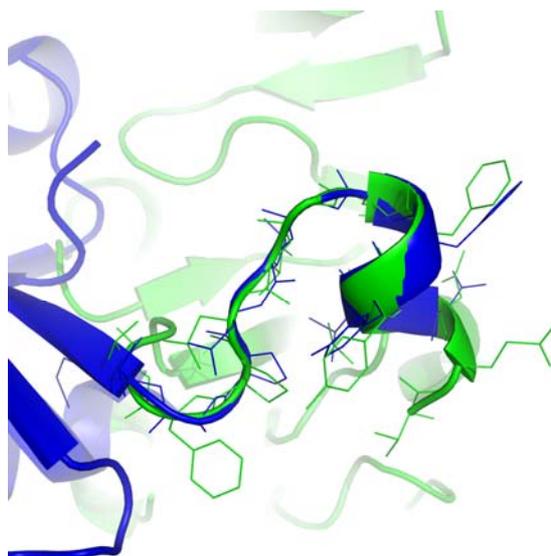


Figure 5. 6 Superimposition of the TEV linker from holo-GAPDHA (green) and holo-GAPDHB (blue). Residues corresponding to the TEV linker in Chain A of holo-GAPDHA and chain D of holo-GAPDHB were aligned using *LSQKAB* (Kabsch, 1976).

5.9. Unit cell analysis

Analysis of the packing of holo-GAPDHB shows that there are fewer contacts between the tetramer and the symmetry-related molecules, compared to holo-GAPDHA. Overall there are five different tetramers that are a minimum distance from holo-GAPDHB, compared to the seven that are found around holo-GAPDHA. This is as a result of the packing of the molecules in the crystallisation process. As there are no intramolecular contacts between the TEV site and the protein within the asymmetric unit, all the contacts occur between the TEV linker and the symmetry-related tetramer, as is shown in Figure 3.10 and Figure 5.6. The presence of the TEV linker forms a unique symmetry-related contact which favours this crystal form and space group. As only one contact is made per tetramer and because of the relatively loose packing within the unit cell, as is shown in Figure 5.7, the other TEV linkers and the histidines are disordered and protrude into the solvent channels.

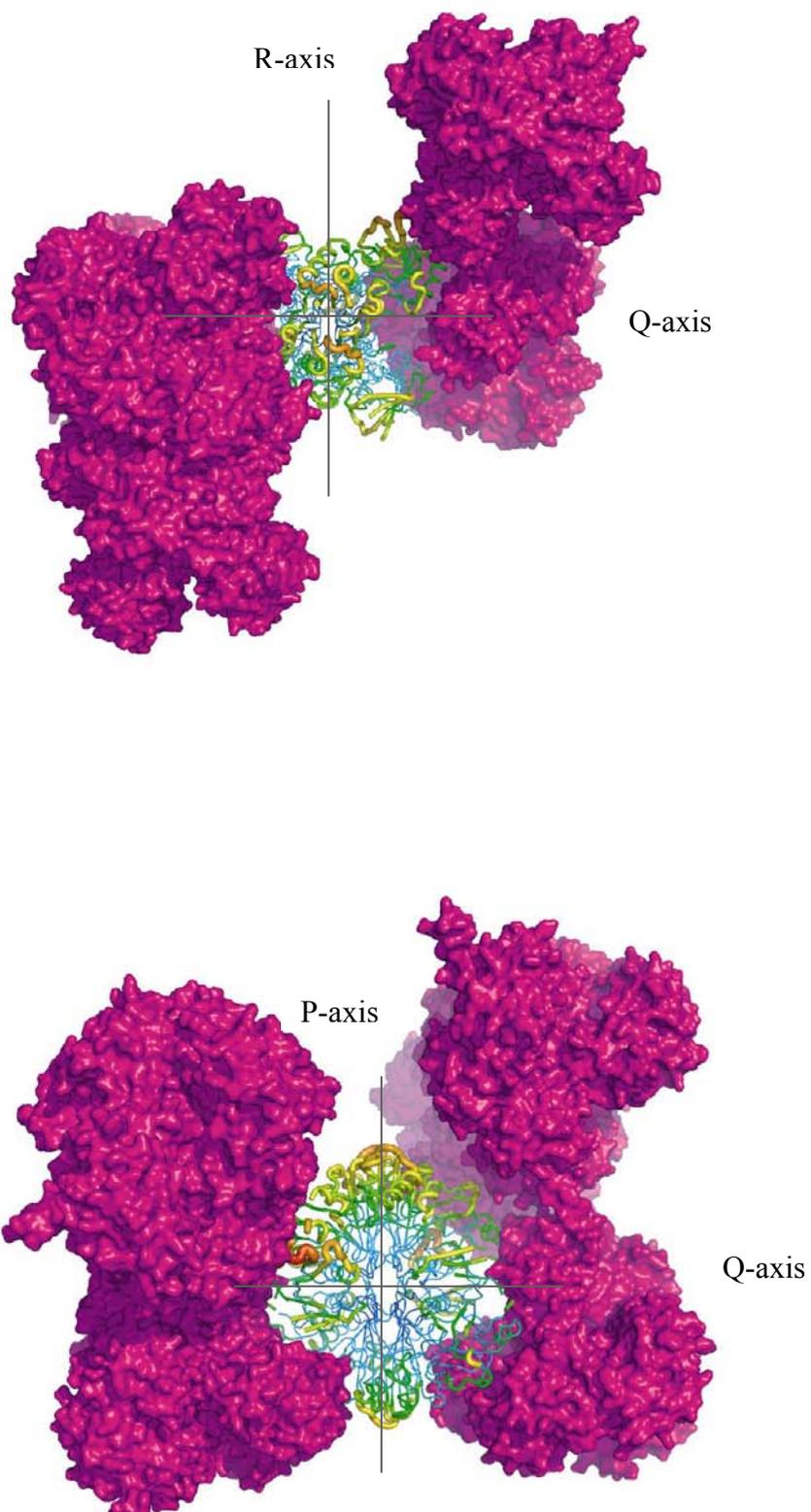


Figure 5. 7 Packing analysis of His₆-GAPDHB. The overall B-factor of the molecule is represented in putty format as described for Figure 3.11. Tetramers which are within a 5 Å distance from the molecule within the asymmetric unit are shown as surface representation.

5.10. NAD⁺ binding to GAPDHB

The analysis of NADP⁺ binding to GAPDHs has been extensively investigated in the previous two Chapters and NAD⁺ binding to GAPDHs and other dehydrogenases has also been rigorously investigated and discussed. For completeness differences between NAD⁺ and NADP⁺ binding will be considered in this chapter with respect to NAD⁺ binding in GAPDHB. Table 5.3 summarises all the residues that contact NAD⁺ and the residues are shown in Figure 5.8. Residues contacting NAD⁺ and for ensuring the stabilisation of its elongated conformation are, not surprisingly the same as those which coordinate NADP⁺, as den

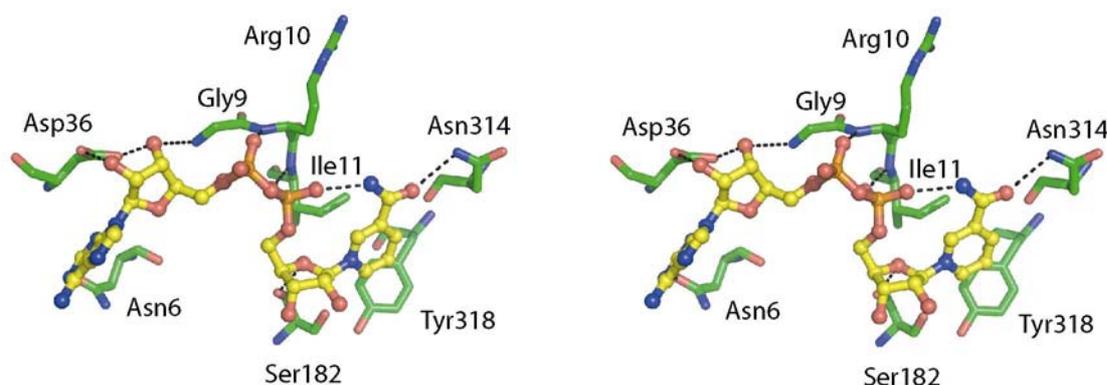


Figure 5. 8 Stereo view of the residues which coordinate NAD⁺ in holo-GAPDHB. An elongated form of NAD⁺ is shown, consistent with NADP⁺ in holo-GAPDHA and other GAPDHs. Residues that coordinate NAD⁺ are highlighted, with their distances shown in Table 5.3

Residue	Distance (Å)	NAD atom
Asp36 OD2	3.4	O2B
Gly9 N	3.6	O3B
Arg10N	3.0	O2A
Ile11N	2.9	O2N
Ser182 OG	2.5	O1A
Asn314 OG	3.3	O7N

Table 5. 3 Residues responsible for coordinating NAD⁺ in holo-GAPDHB are summarised with their distances shown. The nomenclature for NAD⁺ is analogous to NADP⁺ (Table 3.7). With the exception that there is simply a 2' hydroxyl group.

Chapter 3.10.1 demonstrated that a polar residue at the equivalent position of Thr35 in holo-GAPDHA is responsible for coordinating the 2'-phosphate of NADP^+ and was the case in several other NADP^+ -binding GAPDHs. *C. jejuni* GAPDH was the exception to this with respect to the lack of coordination from the threonine to the 2'-phosphate (Chapter 4.9). Superimposition of the NAD^+ and NADP^+ coenzymes from *Bacillus stearothermophilus* GAPDH (1GD1) and holo-GAPDHA respectively, showed that Thr35 was occupied by a leucine in 1GD1 and the residue preceding the leucine was an aspartate that coordinated the hydroxyl groups of the adenosine ribose (Chapter 3.10.1, Figure 3.15). Analysis of holo-GAPDHB shows that as in the case of 1GD1, an aspartate is found coordinating the hydroxyl groups of the adenosine ribose, and this would prevent binding of the 2'-phosphate of NADP^+ .

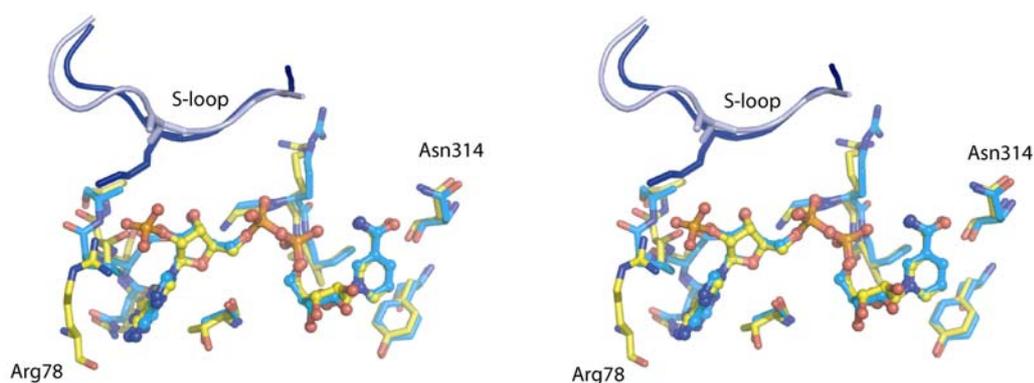


Figure 5. 9 Stereo view of the superimposition of the NAD^+ and NADP^+ coenzymes from holo-GAPDHB and holo-GAPDHA respectively. The carbon atoms for holo-GAPDHB are coloured light blue, with the carbon atoms for holo-GAPDHA coloured yellow. The S-loop from the R-axis related subunit is also shown (dark blue for holo-GAPDHA and light blue for holo-GAPDHB), with the residues at the equivalent position as Lys188 in GAPDHA, Ala190 in GAPDHB shown as sticks.

Figure 5.10 shows the coordination of the aspartate at position 36 by holo-GAPDHB, this coordination stabilises NAD^+ binding and the aspartate would provide electrostatic repulsion of the 2'-phosphate group of NADP^+ , hence preventing its binding. The S-loop of the R-axis related subunit in holo-GAPDHB does not coordinate the coenzyme as in the case of holo-GAPDHA. With the residue at the equivalent position of Lys188 from GAPDHA in

GAPDHB being an alanine. Regardless of the residue at this position the S-loop of GAPDHB would have to extend further to be able to contact the NAD^+ .

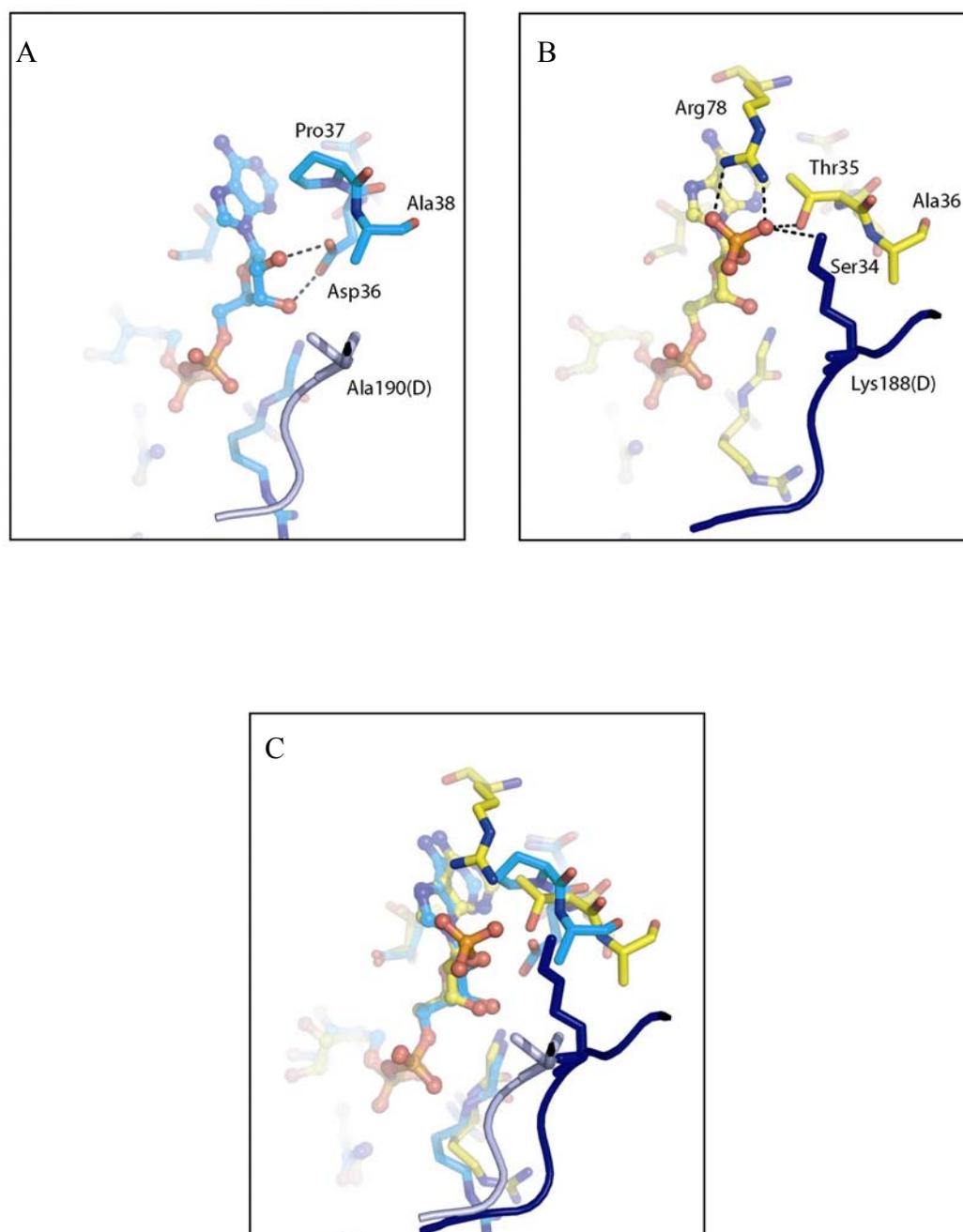


Figure 5. 10 Superimposition of the NAD^+ and NADP^+ coenzymes from holo-GAPDHB and holo-GAPDHA respectively using *LSQKAB* (Kabsch, 1976). Panel A shows the residues around the adenosine ribose, with Asp36 coordinating the two ribose hydroxyl groups and the invariant Thr35 found in NADP-dependent GAPDHs, as shown in Panel B, is replaced by Pro37. Panel B shows the residues from holo-GAPDHA coordinating the 2'-phosphate of NADP^+ . No equivalent residue is found for Arg78 in holo-GAPDHB, although it is present in other NAD^+ -dependent GAPDHs such as 1GD1. Panel C show the superimposition of holo-GAPDHA and holo-GAPDHB.

5.11. Enzymatic analysis of GAPDHB

Investigations into the presence of two *gap* genes in *E. coli* have suggested a duplication event of an ancestral eubacterial gene, leading to two similar (40% sequence identity) genes being present within *E. coli* (Nelson *et al.*, 1991). The function of these two genes in *E. coli* have been characterised with *gap1* displaying classical NAD⁺-dependent phosphorylating glyceraldehyde-3-phosphate activity and the other gene, *gap2* displaying non-phosphorylating erythrose-4-phosphate dehydrogenase activity, resulting in the production of 4-phosphoerythronate (Boschi-Muller *et al.*, 1997, Zhao *et al.*, 1995). Although it had been previously reported that GAPDH can catalyse the oxidation of erythrose-4-phosphate (Ishii *et al.*, 1964). However, (Ishii *et al.*, 1964) and (Zhao *et al.*, 1995), performed the reactions in the presence of phosphate and or arsenate. Therefore, the product produced could be envisaged as being 1,4-bisphosphoerythronate, although thin layer chromatography (TLC) had been carried out by Zhao *et al* to identify if the product was mono- or bi-phosphorylated, however they conceded that it was not always consistent to achieve sufficient resolution during TLC (Zhao *et al.*, 1995). Boschi-Muller *et al* were able to demonstrate non-phosphorylating erythrose-4-phosphate dehydrogenase activity in the absence of phosphate and show by ³¹P NMR that the product contains only one phosphate (Boschi-Muller *et al.*, 1997). To investigate *H. pylori*'s GAPDHs activities assays were performed with erythrose-4-phosphate (e4p) and glyceraldehyde-3-phosphate (g3p) as substrates. Previous analysis of *H. pylori* GAPDHA and *C. jejuni* GAPDH had demonstrated a very low residual activity for erythrose-4-phosphate which was only observed once phosphate/arsenate was present in the assay (Chapter 3.3 and Chapter 4.3 respectively). This low activity, which was not quantifiable, is thought to result from residual phosphorylated sugars present within the commercial erythrose-4-phosphate solution as a result of its chemical synthesis (Sigma provides e4p at 65% purity). Although Sigma are unable to confirm the manufacturing process of e4p it is thought to be a modification of the lead tetraacetate oxidation of glucose-6-phosphate which results in approximately 20% g3p (Simpson *et al.*, 1966).

Enzymatic analysis of His₆-GAPDHB performed as described in Chapter 2.5 of Materials and Methods are summarised below in Table 5.3 with the corresponding plots of substrate concentration versus initial rate shown in Figure 5.11.

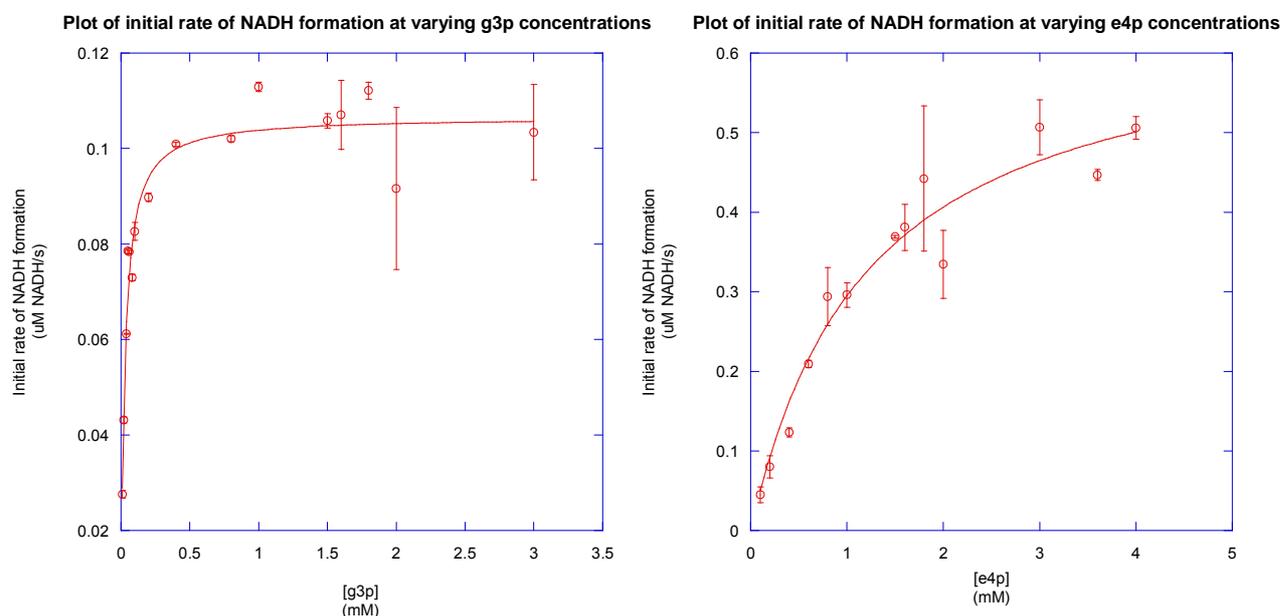


Figure 5. 11 Enzymatic analysis of GAPDHB with saturating concentration (3 mM) NAD^+ . Reactions were performed as described in Chapter 2.5. Initial rates were measured at varying concentrations of substrate, either, g3p (left graph) or e4p (right graph).

	g3p			e4p		
	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($mM^{-1}s^{-1}$)	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($mM^{-1}s^{-1}$)
+ 20 mM arsenate	0.027 ± 0.003	0.37 ± 0.008	13.7 ± 2.4	1.21 ± 0.26	2.24 ± 0.19	1.9 ± 0.7
no phosphate/arsenate	N.D.	N.D.	N.D.	6.2	0.5	0.07
after 20 mM arsenate addition	0.03 ± 0.002	0.7 ± 0.01	23.3 ± 5.0	4.03 ± 0.72	2.3 ± 0.13	0.6 ± 0.2
<i>B. stearotherophilus</i> GAPDH2	1.1	0.12	0.1	0.51	20.0	39.2
<i>B. stearotherophilus</i> GAPDH1	0.9	76.0	84.4	3.3	0.08	0.02

Table 5. 4 Summary of enzymatic parameters for GAPDHB. All reactions were calculated as described in Chapter 2.5. Values for *B. stearotherophilus* are at saturating NAD^+ concentrations in the presence of phosphate for g3p and absence of phosphate for e4p (Boschi-Muller *et al.*, 1997). No g3p activity could be detected in the absence of phosphate for GAPDHB.

These values do not show a clear preference for e4p over g3p by GAPDHB as has been shown for *B. stearrowthermophilus gap1*-encoded gene (Boschi-Muller *et al.*, 1997). Of interest is the very low detectable turnover of e4p in the absence of phosphate and the higher K_M when arsenate is present. As no activity is detected in the absence of phosphate for g3p then the non-phosphorylating reaction is taking place between e4p and GAPDHB, with water attacking the acyl-intermediate in place of phosphate, just like an aldehyde dehydrogenase (Chapter 1.8). No residual activity is detected for e4p in the absence of phosphate or arsenate in GAPDHA and cGAPDH, therefore implying that GAPDHB does possess some e4p-dehydrogenase activity.

An novel observation with GAPDHB is upon the addition of arsenate to the reaction, after it has progressed no longer than one minute and is still proceeding at a linear rate, it would be expected that the reaction rate would become that of when arsenate was present at the start of the reaction. This spike of arsenate gives initial rates with g3p comparable to when arsenate is present in the reaction from the start. However, from Table 5.3 the k_{cat}/K_M ratio is three fold lower for the arsenate spike when e4p is a substrate. One explanation for this is the non-phosphorylated product maybe inhibiting the phosphate-stimulated reaction. This conclusion, like the above values for k_{cat} and K_M values should be taken with great care, since the assay is only measuring initial rates. The rate determining step of the reaction (typically coenzyme exchange (Racker and Krimsky, 1952)). However, the rate determining step with the reaction of GAPDHB with e4p may not be coenzyme exchange; hence, pseudo activity is being measured. Secondly, the high levels of impurities in e4p (~30 %) may have other effects with respect to inhibition of GAPDHB. This second point could be clarified by HPLC analysis of the starting material and products. In addition, the e4p could be generated by using an erythrose-4-phosphate dehydrogenase to catalyse the reverse reaction giving e4p (assuming the same argument with regards substrate specificity does not exist as well).

5.12. Conclusion

This Chapter has shown that GAPDHB exhibits activity consistent with an erythrose-4-phosphate dehydrogenase. From Table 5.4 the activity appears to be phosphate stimulated, since with arsenate the reaction was linear with no product inhibition. However, product inhibition was found to occur with phosphate, this would imply phosphorylating e4p activity. Within metabolic processes there is no reference to the chemical 1,4-biserythronate being used in any pathways. Hence, there would be no need for *H. pylori* to generate the phosphorylated product of e4p unless there is an unidentified enzyme, which can utilise the product of the oxidative phosphorylation of e4p. Alternatively, another enzyme is able to utilise the product in a previously unreported mechanism. Both seem unlikely. Consistent with this phosphate-stimulated hypothesis for GAPDHB, the rate of deacylation of e4p in *B. stearotheophilus* GAPDHB is 20 s^{-1} compared to 10 s^{-1} (Boschi-Muller *et al.*, 1997). Boschi-Muller and co-workers argue that this difference in deacylation corresponds to an activation energy of $4.5 \text{ kcal mol}^{-1}$ required to break the thioacyl bond and release the 4-phosphoerythronate. This energy could be provided by a different conformation of the thioacyl bond with respect to the thioacyl bond of g3p-bound substrate.

To investigate if e4p is a true substrate of GAPDHB and the mechanisms, which dictate specificities between g3p and e4p. The ternary complex of NAD^+ -GAPDHB-g3p/e4p could provide an insight into substrate specificity, analogous to how the structures of holo-GAPDHA and holo-cGAPDH determined coenzyme specificity in the previous chapters. The determination of ternary complexes of holo-GAPDHA and holo-GAPDHB between g3p and e4p is the subject of investigation for Chapter 6.

Chapter 6
Ternary Complexes of *H. pylori* GAPDHA and GAPDHB mutants

6.1 Introduction

Chapter 5 demonstrated that GAPDHB is able to use glyceraldehyde-3-phosphate and erythrose-4-phosphate as substrates, although the kinetic data was not altogether decisive. To investigate this apparent dual substrate specificity further, a mutagenic approach was undertaken with the aim of obtaining ternary complexes between holo-GAPDHB and g3p and also holo-GAPDHB and e4p. A structural analysis of the complex between GAPDHB and e4p would be hoped to provide an understanding of specificity between the two substrates. As in the case for the holo-structures of GAPDHA, cGAPDH and GAPDHB (Chapters 3, 4 and 5 respectively) in providing an understanding for coenzyme specificity.

Several other groups have crystallised a ternary complex between a GAPDH, NAD^+ and substrate, in the case of Castilho and co-workers this was achieved through a non-reversible inhibitor of GAPDH (Castilho *et al.*, 2003) and for (Yun *et al.*, 2000) only the structure of the hemiacetal intermediate formed between the active site cysteine and g3p was obtained when crystals of the apo enzyme were soaked in g3p. The structure of the irreversible inhibitor glycidol phosphate has also been reported (Skarzynski *et al.*, 1987). In the case of Didierjean and co-workers, a mutagenic approach was undertaken to permit the co-crystallisation of g3p and NAD^+ (Didierjean *et al.*, 2003). A quaternary complex consisting of NAD^+ , g3p and phosphate has not been reported, possibly due to the dual conformations of g3p within the active site with the phosphate group of g3p being reported in the Ps site when it is bound in a non-Michaelis complex (i.e. inactive active site) (Didierjean *et al.*, 2003) and within the Pi site as a hemiacetal intermediate when soaked into apo-GAPDH crystals (Yun *et al.*, 2000). Analogues of g3p binding irreversibly to the active site cysteine have been shown to be present in both the Pi and Ps sites displaying dual conformation (Castilho *et al.*, 2003). Therefore, taken together this supports the flip-flop mechanism proposed by Skarzynski and colleagues (Skarzynski *et al.*, 1987). However, there is no definitive structural work that demonstrates the flip-flop mechanism with g3p occupying the Pi and Ps sites. Recently the structure of holo-GAPDH from *Bacillus stearothermophilus* with the acyl-enzyme intermediates phosphate occupying the Pi site with NAD^+ bound has provided further support for the Flip-flop mechanism (Moniot *et al.*, 2008).

A mutagenic approach was undertaken in the manner of Didierjean *et al* with the aim of obtaining a ternary complex between GAPDHB, NAD⁺ and g3p and GPADHB, NAD⁺ and e4p. The mutagenic approach would be the only approach which would allow the co-crystallisation of NAD⁺ and g3p/e4p with GAPDHB and also NADP⁺ and g3p/e4p with GAPDHA. It is plausible that during the crystallisation process g3p would be turned over in the presence of NAD⁺, even if phosphate was omitted, as water would be able to attack the hemiacetal intermediate over the time required for the crystals to grow. Soaking of holo-GAPDH wild type crystals with g3p would also be deemed unsuccessful due to the rapid turnover of g3p. The generation of NADH would cause conformational changes within the protein upon coenzyme exchange and possibly destroy the crystals. This conformational change upon coenzyme exchange has been shown by (Leslie and Wonacott, 1984), in addition, Chapter 3 showed the interdomain movement between apo and holo-GAPDHA.

H. pylori GAPDHA and GAPDHB were both successfully mutated at their active sites with the active site cysteine being replaced as a serine and an alanine. An alanine would render the enzyme completely inactive whilst a serine mutant may be able to attack the carbonyl of g3p or e4p, the reaction rate would be expected to be of several orders of magnitude lower than wild-type. Indeed, no activity was detected with either mutant, confirming that the active site cysteine does attack g3p and in the case of GAPDHB, e4p. The results of the co-crystallisation of mutant GAPDHA and GAPDHB in the presence of coenzymes and substrates is discussed within this chapter. Whilst the structures of the mutant complexes fail to provide an understanding of specificity between the two substrates and even unambiguously confirm the atomic structure of each substrate, analysis of its binding within the active site provides further evidence supporting the Flip-Flop mechanism.

6.2 Site-directed Mutagenesis

The active site cystines of GAPDHA and GAPDHB were mutated to either serine or alanine as described in Chapter 2.8.1. Briefly, the method chosen for the mutagenesis was chosen to drive the mutagenesis reaction towards completion. The typical method involving PCR amplification of the entire expression plasmid can sometimes result in few or no colonies and the effectiveness of the primers are not known until the resultant plasmid has been sequenced. Whereas in the method outlined in Chapter 2.2.6, although an additional cloning step is required, once the PCR has been carried out, it is highly probable that the mutation has been successful.

6.3 Protein expression and purification

GAPDHA and GAPDHB cystine mutants were expressed and purified in the same manner as for their respective wild-type forms with no difference in levels of expression or enzyme stability during purification being observed. Although prior to crystallisation the hexa-histidine affinity tags were removed from all mutant enzymes by the overnight incubation of the enzyme with TEV protease, as described in Chapter 2.8.2. The extra purification step of removing the hexa-histidine tag was as a result of analysis of the holo structures of GAPDHA and GAPDHB. In both instances, the histidines corresponding to the affinity tag were not visible during the electron density analysis and furthermore only the TEV protease site was interpretable in some instances, involved in the crystal packing of the tetramer (Chapter 3.9.2 and Chapter 5.6.4 for holo-GAPDHA and holo-GAPDHB respectively). Analysis of the overall structures of holo-GAPDHA and holo-GAPDHB had shown that although they had been crystallised in the presence of the hexa-histidine tag, only a short stretch of sequence corresponding to the TEV cleavage site was ordered with the actual tag protruding into the solvent channel. Therefore, the assumption was made that this disordered stretch of sequence may interfere with closer crystal contacts reducing the overall order of the molecule. Due to the need for a higher resolution structure of GAPDHB (the highest resolution was 2.8 Å after dozens of crystals had been screened) to be able to differentiate between e4p and g3p (the only difference between the two molecules is a CHOH group) the hexa-histidine tags were removed prior to crystallisation.

6.4 Crystallisation of GAPDHA and GAPDHB mutants

6.4.1 GAPDHA c149s

Owing to the time limitations, only the serine mutant of GAPDHA (c149s) was co-crystallised in the presence of g3p /e4p and NADP⁺. Crystallisation conditions were explored around the initial hit for apo-GAPDHA and focused on PEG-8000 as the precipitant at higher pH values (not surprisingly no conditions were found around the initial hit for holo-GAPDHA since the enzyme had the hexa-histidine tag present). Grid screening using the Ozma 8k crystallisation screen (Emerald Biosciences) with 100 mM Tris pH 8.0 as the buffer, gave multiple hits with crystals appearing overnight, with some appearing after less than an hour and growing to full size within 24 hours. The majority of crystals were needles, although some conditions, most notably (18% PEG-8000, 100 mM Tris pH 8.0, 200 mM NaI) did provide larger crystals, like those shown in Figure 6.1. Although the crystals appeared very quickly careful refinement of the crystallisation space may have yielded better quality crystals, which may have diffracted to a greater resolution. Unfortunately, time constraints prevented a thorough exploration of the conditions. The focus of crystallisation was applied to the GAPDHB mutants (Chapter 6.4.2) for the analysis of substrate specificity, which was not the aim for GAPDHA.

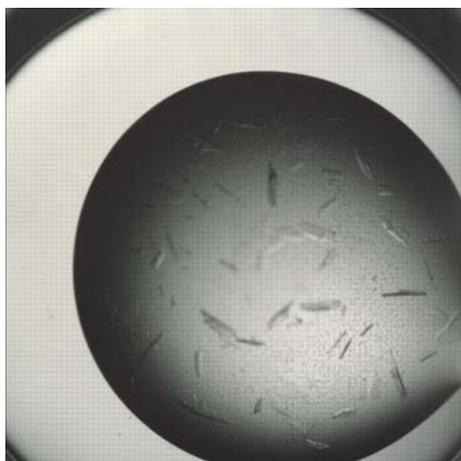


Figure 6. 1 Crystals of GAPDHA co-crystallised in the presence of NADP⁺ and g3p/e4p. Crystals appeared within several hours and would grow to full size within 24 hours.

6.4.2 GAPDHB c151a and c151s

The crystallisation conditions for GAPDHB mutants c151a and c151s were different from the conditions for His₆-GAPDHB, owing to the removal of the hexa-histidine affinity tag. To explore new crystallisation conditions, sparse matrix screening (Jancarik and Kim, 1991) using Cryo I and II screens (Emerald BioSystems) was carried, in the same manner as described in previous chapters (3.4, 4.4 and 5.3). Briefly, 12 mgml⁻¹ of mutant GAPDHB (c151a/s) buffered in 20 mM MES, 100 mM NaCl, 1 mM DTT pH 6.5, containing 2 mM NAD and 1 mM of either g3p or e4p, was mixed with an equal volume of precipitant from the Cryo screens (500 nl of precipitant mixed with 500 nl protein solution) and incubated at 293 K. Initial conditions focused on the lower molecular weight PEGs, at low pH values (as the case for His₆-GAPDHB). With 40% PEG-300, 100 mM phosphate-citrate pH 4.2, mixed with an equal volume of 12 mgml⁻¹ of c151s mutant GAPDHB containing 2 mM NAD⁺ and 1mM g3p/e4p, giving crystals of sufficient size and physical quality. Crystals grew within one week and were of suitable size and aesthetic quality from the initial screen and were also cryoprotected, so no further refinement of these conditions was carried out.

After data collection of the g3p and e4p co-crystallised GAPDHB c151s crystals (Chapter 6.5.2) initial electron density analysis suggested that citrate could be bound to the active site. Previous studies had shown that citrate indeed can displace the phosphate anions within GAPDH (Moras *et al.*, 1975), although no structure has been deposited in the PDB for this occurring within the GAPDH family. Previous structural analysis of the Dogfish M4 Apo-Lactate Dehydrogenase, also carried out by Rossmann, showed citrate occupying a position near the sulphate and NAD⁺ molecules (Abad-Zapatero *et al.*, 1987). In addition, although with the alanine mutant there would be no phosphate attack of the hemi-thioacetal intermediate, phosphate binding may compete with substrate binding. The strategy therefore, for crystallising GAPDHB c151a mutant involved omission of citrate and phosphate from the crystallisation conditions. A sodium acetate buffer was used instead of the phosphate-citrate buffer for screening for crystallisation conditions for c151a co-crystallised with NAD⁺ and either g3p or e4p. Suitably sized crystals, grown in a cryoprotecting mother liquor of 38% PEG-400, 100 mM sodium acetate pH 4.0, were obtained for the c151a mutant, crystals are shown in Figure 6.2B.

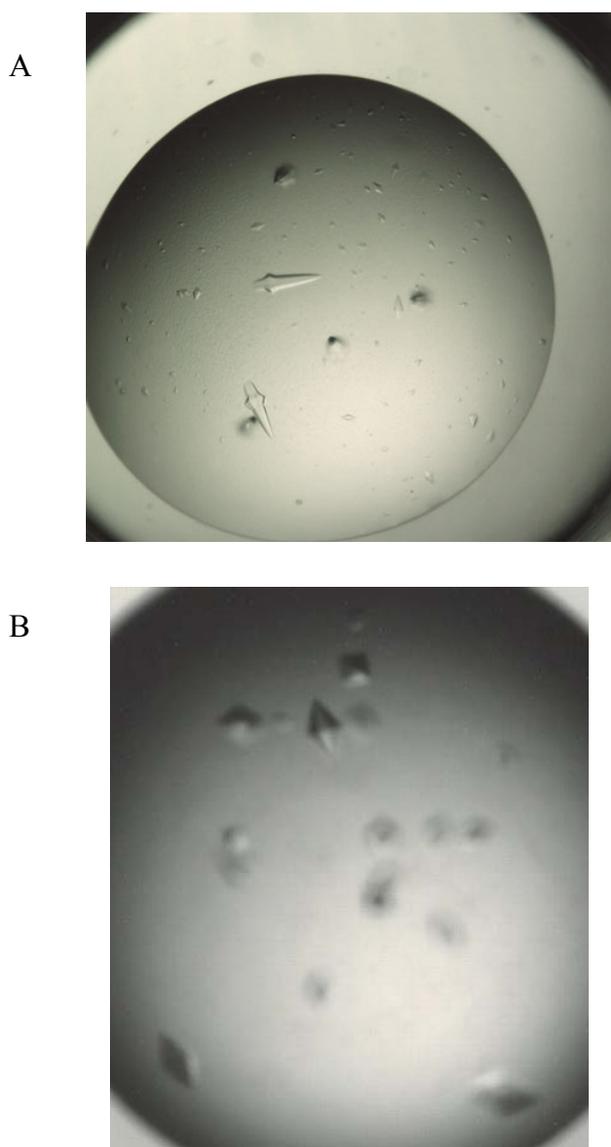


Figure 6. 2 A. Crystals of GAPDHB c151s grown in the presence of NAD^+ and either g3p or e4p, the smaller (more typical GAPDHB) crystals diffracted to a higher quality. **B.** Crystals of GAPDHB c151a grown in the presence of NAD^+ and either g3p or e4p,.

6.5 Data collection for GAPDH mutants

Availability of protein and time constraints meant that data collection for different mutants were on different beam lines at different dates. GAPDHA c149s mutant co-crystallised with NADP^+ and either g3p or e4p and GAPDHB c151s co-crystallised with NAD^+ and either g3p or e4p were collected at the ESRF on beam line ID14-1. GAPDHB c151a co-crystallised with NAD^+ and either g3p or e4p were collected at Diamond on beam line I03.

6.5.1 GAPDHA c149s

Crystals of the GAPDHA c149s mutant grown in the presence of g3p or e4p, were transferred through paratone-N oil as a cryoprotectant and frozen in a stream of boiled-off liquid nitrogen prior to data collection. Complete data sets were collected from single cryo-cooled crystals at beam line ID14-1 at the ESRF, Grenoble, using a fixed wavelength of 0.934 Å and an ADSCQ4R CCD detector with a crystal to detector distance of 288.9 mm for the g3p co-crystallised mutant and 265.5 mm for the e4p co-crystallised mutant. A typical diffraction image for g3p co-crystallised GAPDHA mutant is shown in Figure 6.3A, whilst Figure 6.4A shows a typical diffraction image for the e4p co-crystallised GAPDHA mutant. For each g3p and e4p co-crystallised GAPDHA mutant 180 images were collected at 1.0 ° oscillation width.

As is clear from the diffraction image of GAPDHA c149s crystallised in the presence of g3p, there is a second crystal lattice. This suggests the crystal may inherently be composed of two crystals, or the crystal cracked during the freezing process. The diffraction quality improves during oscillation around the phi axis as is clear in Figure 6.3B. However, as the crystal belongs to space group $P2_1$ 180° of data had to be collected. Therefore, during refinement of the cell parameters and integration of the diffraction intensities, the actual position of each spot becomes uncertain as the lattices begin to coincide. This is reflected in the R_{merge} , where a plot of R_{merge} against batch (image number) shows an improvement in data quality during the data collection process (a reduction in data quality during data acquisition can sometimes be attributed to radiation damage as is the case for GAPDHB crystals, Chapter 6.5.2). Typically, this was the best crystal of GAPDHA c149s co-crystallised in g3p, therefore an attempt was made to process the data.

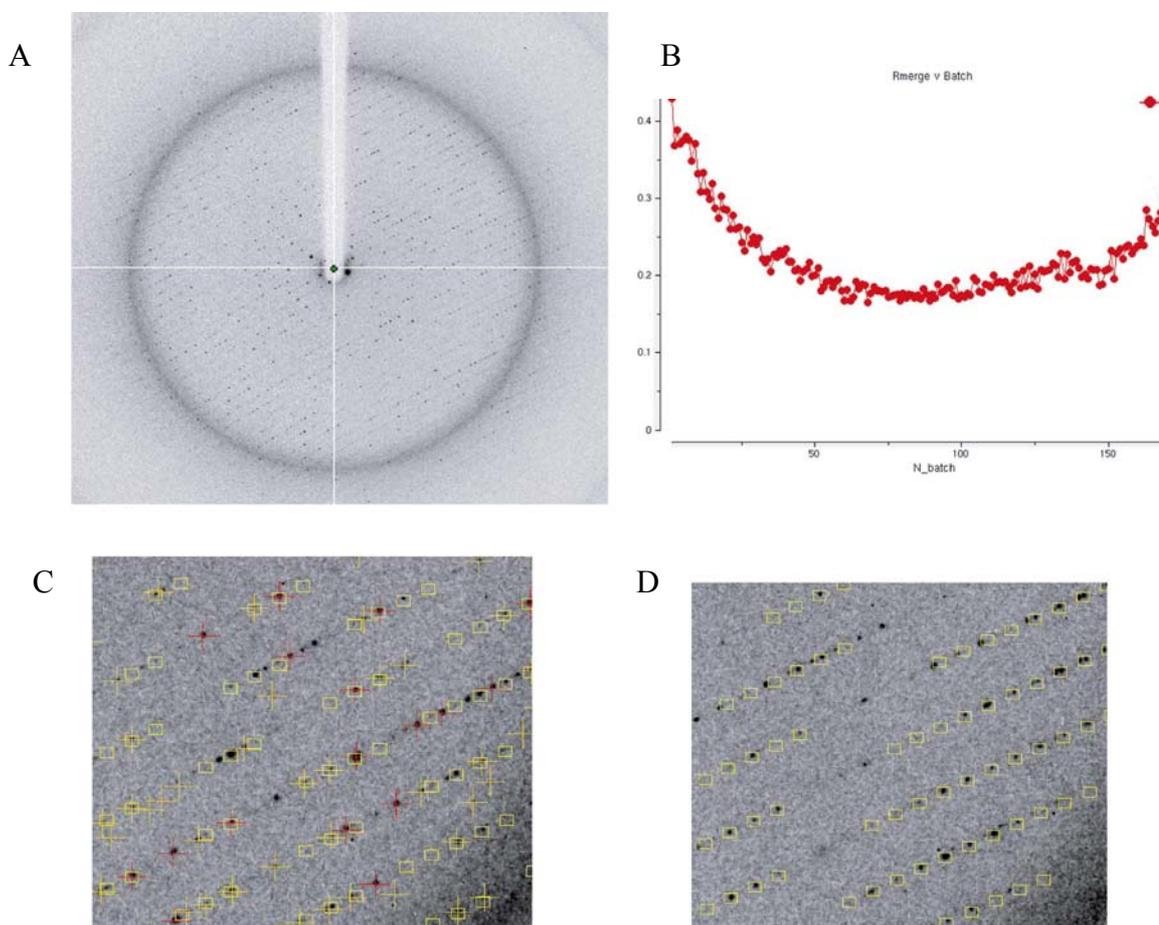


Figure 6. 3 Diffraction quality for GAPDHA c149s crystallised in the presence of g3p. Panel A shows the first diffraction image. C. Upon closer inspection, with the spot predictions shown after indexing the lattice of the crystal, extra spots belonging to a different crystal are present. Panel D is a magnified area after 15° of data collection; here the predominant crystal lattice is clearly visible. The appearance of the second crystal lattice during data collection is reflected in the R_{merge} . Panel B shows the plot of R_{merge} against batch. Note the high R_{merge} in the first few images where the two crystal lattices are almost coincident. Further, into data collection as the crystal is rotated nearly 180° the other crystal form appears on the diffraction images and is reflected in the increasing R_{merge} .

Data collection of the GAPDHA c149s crystal grown in the presence of e4p did not suffer from any apparent multiple crystal forms as is clear from Figure 6.4.

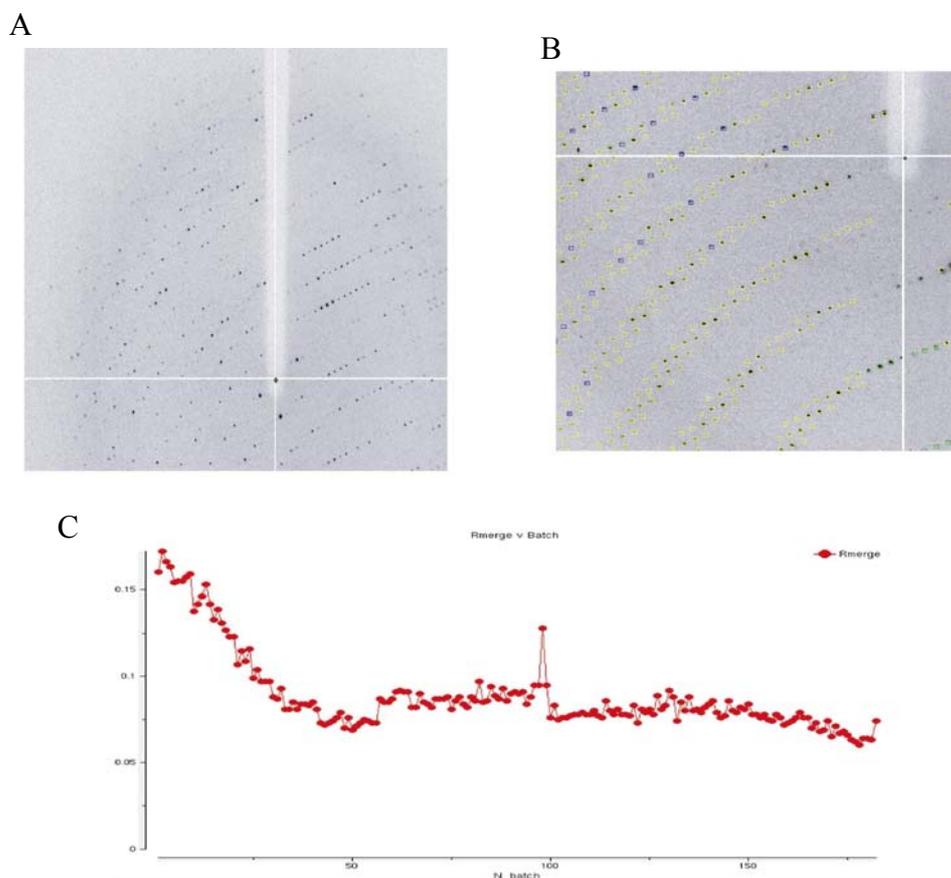


Figure 6. 4 Diffraction quality of GAPDHB c149s crystal grown in the presence of e4p. Panel A shows a typical diffraction image collected at ID14-1 (ESRF) on an ADSCQ4R CCD detector with a crystal to detector distance of 265.5 mm. Panel B shows the spot predictions after indexing of the lattice. Panel C shows the plot of R_{merge} versus image number, note the difference of scale between e4p and g3p co-crystallised crystals.

The merging R factor for GAPDHA grown in the presence of e4p is far more respectable than for the dataset from the g3p containing crystal, as is clear in Figure 6.4C. It should be stressed that the difference between substrate co-crystallised with holo-GAPDHA c149s has not influenced the quality of the data collected. Rather, only a few crystals out of those harvested diffracted suitably at the synchrotron, hence a workable data set could have been collected for g3p-containing crystals if more crystals had been screened.

6.5.2 GAPDHB c151s

Crystals of the GAPDHB c151s mutant grown in the presence of g3p or e4p were mounted in the appropriately sized litho loop and with the mother liquor being a suitable cryo protectant, frozen directly into liquid nitrogen prior to data collection. A complete data set was collected from a single cryo-cooled crystal of g3p and a crystal of e4p co-crystallised GAPDHB c151s at beam line ID14-1 at the ESRF, Grenoble. In all thirty crystals co-crystallised in the presence of g3p or e4p were screened for their diffraction quality, with resolutions typically being in the order of 3.5-4.0 Å, with the rare few crystals diffracting to a higher resolution. Crystals co-crystallised in either g3p or e4p were then remounted prior to a complete data set being collected. A fixed wavelength of 0.934 Å was used with diffraction data collected on an ADSCQ4R CCD detector with a crystal to detector distance of 324.0 mm for the g3p co-crystallised crystal and 300.5 mm for the e4p co-crystallised crystal. For both data sets, 90 images were collected with an oscillation width of 1°. Typical diffraction images are shown in Figure 6.4.

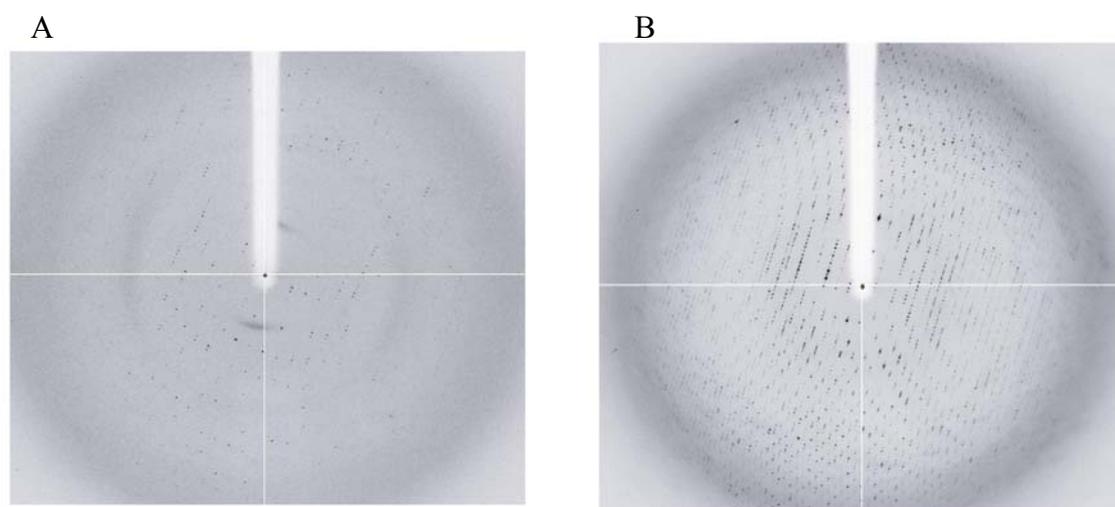


Figure 6. 5 Typical diffraction images for GAPDHB c151s crystals grown in the presence of; A, g3p and B, e4p. Both data sets were collected at ID14-1 (ESRF) on an ADSCQ4R CCD detector with a crystal to detector distance of 324.0 mm for g3p containing crystal and 300.5 mm for e4p containing crystal. For each data set 90 images were collected with an oscillation range of 1 °.

6.5.3 GAPDHB c151a

Crystals of GAPDHB c151a grown in the presence of either g3p or e4p were prepared for data collection in the same way as for the GAPDHB c151s crystals described previously. Again, several crystals had to be screened before a crystal of diffraction quality high enough for the interpretation of any possible substrates within the active site could be collected. A complete data set was collected for crystals grown in the presence of g3p or e4p at beam line I03 at the Diamond Light Source, Oxford. An ADSCQ315 CCD detector was used with a crystal to detector distance of 400 mm for the g3p and e4p co-crystallised crystals. A fixed wavelength was used of 0.92 Å with 120 images being collected for each crystal at 1.0° oscillation width. Again typical diffraction images are shown in Figure 6.5 for GAPDHB c151a g3p and e4p crystals.

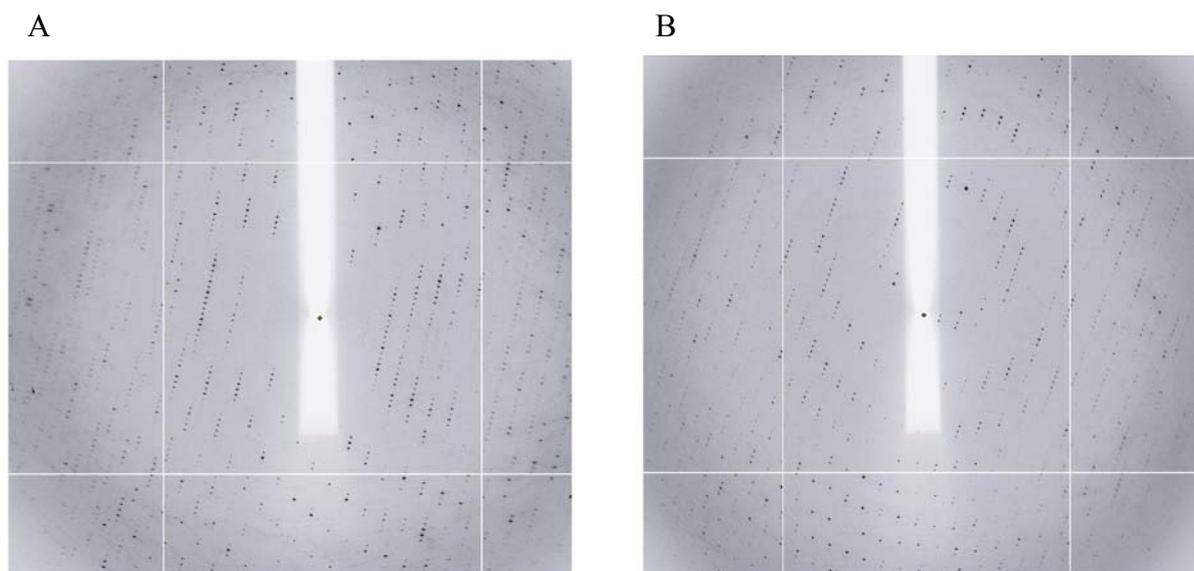


Figure 6. 6 Typical diffraction images for GAPDHB c151a mutant grown in the presence of A g3p and B e4p. Both data sets were collected at beam line I03 Diamond, Oxford at a fixed wavelength of 0.92 Å, with a crystal to detector distance of 400 mm for each data set. In all for each data set 120 images were collected at a 1° oscillation width.

6.6 Data processing

Both GAPDHB mutants and GAPDHA mutants were grown in the same conditions for g3p and e4p co-crystallisations. As each of these crystals has a similar cell dimension then only the overall strategy for the data processing of each mutant is given.

6.6.1 GAPDHA c149s

Intensities for the g3p and e4p containing crystals were scaled using *MOSFLM* (Leslie, 1992) with the autoindexing routines giving a solution with a primitive monoclinic cell. The data were scaled in *SCALA* (Evans, 2006) in $P2_1$ with the statistics shown in Table 6.1. Examination of the principal axes showed intensity distribution consistent with the space group $P2_1$, consistent with apo, untagged-GAPDHA described in Chapter 3.6. Due to the removal of the hexa-histidine tag and with crystallisation conditions of GAPDHA c149s being similar to apo-GAPDHA, it is no surprise that the two crystal forms are of the same space group. Solvent content analysis using four molecules of GAPDHA c149s within the asymmetric unit indicated a 54.3 % solvent content with a Matthews coefficient of $2.71 \text{ \AA}^3 \text{ Da}^{-1}$ (Matthews, 1968). As is clear from Table 6.1 the R_{merge} for GAPDHA c149s crystallised in the presence of g3p is particularly high, as discussed in Chapter 6.4.1.

6.6.2 GAPDHB mutants

Results from the autoindexing routines of *MOSFLM* (Leslie, 1992) suggested a primitive trigonal cell. The data were scaled using *SCALA* (Evans, 2006), initially in the highest Laue group, *P3*. Examination of intensities along the principal axis indicated either a 3_1 or 3_2 screw axis. The Matthews coefficient was consistent with a tetramer being within the asymmetric unit, with the typical asymmetric unit consisting of a $V_m = 4.7 \text{ \AA}^3 \text{ Da}^{-1}$ and $V_s = 74 \%$ solvent.

Due to the different cell dimensions of the mutant GAPDHBs, molecular replacement using wild type GAPDHB was used as a search model in *PHASER* (Read, 2001), sampling for four copies of GAPDHB in the asymmetric unit, giving a solution only in the space group *P3₂*, of which all the subsequent GAPDHB mutants were rescaled into. A summary of the data collection statistics for the GAPDHB mutants is shown in Table 6.1. The high R_{merge} for the GAPDHB mutants is a result of radiation damage, consistent with that observed for wild type holo-GAPDHB in Chapter 5.4.

Of particular significance from table 6.1 is the quality of the data. The R_{merge} is very high for the outer bin, taken with the low $\{I/\sigma(I)\}$ in the outer bin suggests that the data is very poor at the higher resolution reported. Rather than cutting the data off at a lower resolution and giving better merging statistics, it is better to include the data for the molecular replacement and refinement. Assuming there are no systematic errors in the higher resolution shell, then the data will still have a significant contribution to the electron density calculation.

	c149s + g3p	c149s + e4p	c151s + g3p	c151s + e4p	c151a + g3p	c151a + e4p
Beam line	ID14-1	ID14-1	ID14-1	ID14-1	I03	I03
Space group	$P2_1$	$P2_1$	$P3_2$	$P3_2$	$P3_2$	$P3_2$
Unit cell parameters (a,b,c, Å)	70.2, 146.1, 144.2	70.5, 108.2, 108.9	111.6, 111.6, 198.9	111.8, 111.8, 199.2	115.7, 115.7, 199.6	112.6, 112.6, 199.4
(α , β , γ , °)	90.0, 89.9, 90.0	90.0, 102.0, 90.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0
Resolution Limits	51.3-2.7 (2.8-2.7)	48.2-2.5 (2.6-2.5)	55.8-3.0 (3.1-3.0)	55.9-2.8 (2.9-2.8)	69.3-2.5 (2.6-2.5)	62.7-2.5 (2.6-2.5)
Number Observations	14220 (21552)	210547 (29618)	212793 (29996)	204402 (28870)	331352 (45129)	371428 (53524)
Number Unique Observations	73729 (11023)	56165 (8128)	56566 (8282)	69497 (10071)	96497 (13847)	100141 (14722)
Average Redundancy	1.93 (1.96)	3.78 (3.65)	3.76 (3.62)	2.94 (2.87)	3.43 (3.26)	3.70 (3.64)
Completeness (I/ σ (I))	91.0 (93.1) 4.2 (1.5)	99.9 (99.5) 12.7 (5.1)	100.0 (99.9) 6.7 (1.4)	99.4 (98.3) 6.0 (1.2)	98.1 (96.6) 9.7 (1.3)	99.9 (100.0) 8.3 (1.1)
$\dagger R_{\text{merge}}$	0.22 (0.39)	0.09 (0.31)	0.25 (0.98)	0.17 (0.95)	0.11 (0.93)	0.14 (1.27)

Table 6. 1 Data collection statistics for all GAPDH mutant structures. Only the mutation is shown together with whether it was crystallised in the presence of g3p or e4p. Values in parentheses are for the highest resolution shell.

$\dagger R_{\text{merge}} = \frac{\sum_{hkl} [\sum_i |I_{hkl,i} - \langle I_{hkl} \rangle|]}{\sum_{hkl} \langle I_{hkl} \rangle}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices h , k , and l and $\langle I_{hkl} \rangle$ is the mean intensity of that reflection

6.7 Structure Determination

6.7.1 GAPDHA c149s

Even though the GAPDHA c149s mutant crystals grown in either g3p or e4p belonged to the same space group as apo-GAPDHA, their cell dimensions were different by up to 50 Å and 8° (apo-GAPDHA a, b, c, 75.2, 100.6, 97.8 Å respectively, compared to GAPDHA c149s g3p of a, b, c, 70.2, 146.1, 144.2 Å respectively). Therefore, molecular replacement was required for structure determination. Molecular replacement was carried out using *PHASER* (Read, 2001) using Chain A of holo-GAPDHA with NADP⁺ and the waters removed, searching for four copies within the asymmetric unit searching in the space group *P2*₁. For the g3p and e4p co-crystallised mutant GAPDHAs, a tetramer displaying the 222 symmetry consistent with an intact GAPDH enzyme was found. The least log likelihood gain (LLG) for the solution found for the g3p structure was 1095, whilst a LLG for the e4p structure was 7342. Upon refinement in *REFMAC5* (Murshudov *et al.*, 1997) the R factor dropped from 36 % to 30 % for the GAPDHA mutant grown in the presence of e4p. However, the GAPDHA mutant grown in the presence of g3p did not refine with the R factor remaining at 44 %. This lack of refinement and the low LLG is indicative of the molecular replacement solution of the g3p structure was not correct and analysis of the electron density confirmed this observation.

Manual rebuilding of the mutant GAPDHA e4p model was carried out in *COOT* (Emsley and Cowtan, 2004) with subsequent cycles of refinement in *REFMAC5* (Murshudov *et al.*, 1997). Once the R-factor had reached 26 % and the geometry of the model was acceptable, waters were added using *COOT* (Emsley and Cowtan, 2004) and manually inspected as described previously (Chapter 3.7.1). Within the e4p model there was a clear difference between the active site within each monomer. In particular, Chains B, C and D only contained waters, whereas Chain A had electron density consistent with the e4p phosphate occupying the Ps site (Figure 6.10, Chapter 6.9.2). Restraints for e4p were generated from *LIBCHECK*, added to the GAPDHA c151s model in *COOT* (Emsley and Cowtan, 2004), and refined in *REFMAC5* (Murshudov *et al.*, 1997). The overall B-factor for e4p occupying Chain A was 36 Å². A summary of the refinement statistics for mutant GAPDHA containing e4p and the GAPDHB mutant complexes is summarised in Table 6.2.

6.7.2 GAPDHB mutants

Molecular replacement was carried out with the c151s mutant grown in the presence of g3p as described in Chapter 6.6.2 using wild type GAPDHB as a search model, with *PHASER* (Read, 2001) giving a LLG of 2179 as the highest solution. Upon refinement in *REFMAC5* (Murshudov *et al.*, 1997) the R-factor dropped from 34 % to 30 %, taken together with the LLG of 2179 for the search model, suggests the molecular replacement solution was valid. A summary of the iterative model building and refinement process for the GAPDHB mutants is shown on the following page in Figure 6.7.

The refined GAPDHB c151s model was used as a starting model to refine the other mutant structures, owing to all the cell dimensions of the GAPDHB mutants being very similar (as is clear from Table 6.1). Once each mutant structure had been refined in the absence of ligand, the appropriate ligand was fitted into the electron density through *COOT* (Emsley and Cowtan, 2004). The ligand was then refined in *REFMAC5* (Murshudov *et al.*, 1997) with further modifications within *COOT* (Emsley and Cowtan, 2004). In the case of the c151s crystal grown in the presence of e4p the initial refined map did not have convincing electron density for the phosphate group of e4p. In this case a composite omit map was generated (as described in Chapter 2.4.3.1) and used in *COOT* (Emsley and Cowtan, 2004) to fit the molecule. Once all ligands had been fitted and refined, omit maps were calculated with the ligand removed. This process was important in demonstrating the presence of bound ligands within the active sites, especially when dual conformation of the ligands in question often made it difficult to assign the position for the molecules within the active sites with confidence.

Figure 6. 7 Flow chart representation of the iterative cycles of refinement and model building for the GAPDHB mutants, with the incorporation of either g3p or e4p ligands as dictated by the data set collected. Only one molecular replacement calculation needed to be ran and this was for GAPDHB c151s mutant grown in the presence of g3p. After this model was rebuilt and satisfied the geometric and stereochemical parameters it was used to refine the other mutant data sets. Once each model satisfied the geometric and stereochemical parameters the appropriate ligand was added. Ligand restraint files (.cif) were generated from *LIBCHECK* and used in subsequent model building/refinement processes. Once the ligand was incorporated and electron density for the ligand was consistent after an omit map had been calculated then waters were added and inspected through *COOT* (Emsley and Cowtan, 2004).

6.8 Stereochemical analysis

All GAPDH mutant structures display the characteristic 222 symmetry. As crystals of c149s grown in the presence of g3p did not give a valid molecular replacement solution, it will not be discussed further. Analysis of the stereochemical quality of the model was performed using *PROCHECK* (Laskowski *et al.*, 1993) and MolProbity (Davies *et al.*, 2007). The refinement statistics for GAPDHA c149s grown in the presence of e4p is shown, along with the other GAPDHA mutants in Table 6.2. The Ramachandran plot of Chain A for the GAPDHA c149s mutant with e4p is shown in Figure 6.8 A and is consistent with that for holo-GAPDHA (Chapter 3.8.2).

The GAPDHB mutants all have the same residues in the disallowed regions as in holo-GAPDHB, again this is no surprise since there is little conformational change of the tetramer on binding substrate, as has been observed for other GAPDH-substrate complexes (Didierjean *et al.*, 2003). The Ramachandran plot for GAPDHB c15s mutant grown in the presence of e4p is shown in Figure 6.8 B.

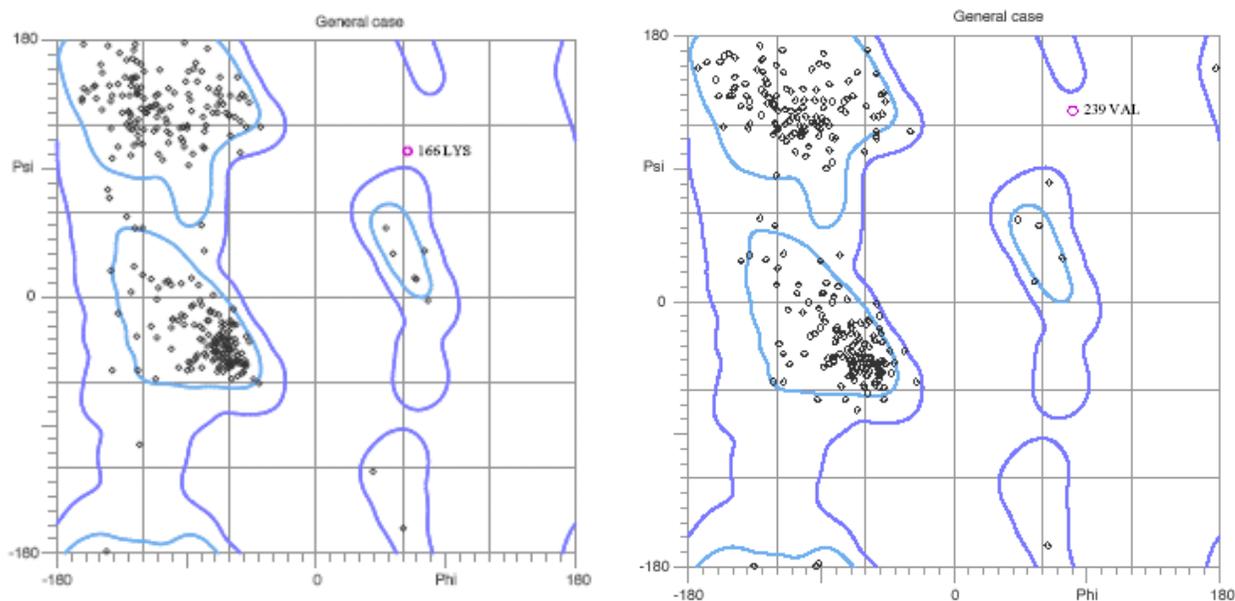


Figure 6. 8 Ramachandran plots of Chain A for; A, GAPDHA c149s grown in e4p and B, GAPDHB c151s grown in e4p.

	GAPDHA c149s e4p	GAPDHB c151s g3p	GAPDHB c151s e4p	GAPDHB c151a g3p	GAPDHB c151a e4p
Resolution range (Å)	42.0-2.7	55.8-3.0	53.8-2.8	55.8-2.5	66.5-2.5
No. reflections	5613	56497	69369	96471	99759
Completeness (%)	99.0	99.8	99.2	98.0	99.6
Matthews Coefficient	2.71	4.73	4.84	4.71	4.65
Solvent content (%)	54.3	73.8	74.4	73.7	73.3
R-factor	19	22	22	22	21
R _{free} -factor	25	27	26	26	26
RMSD bond lengths (Å)	0.013	0.018	0.015	0.010	0.025
bond angles (°)	1.66	2.13	1.85	1.65	2.37
No. solvent atoms	458	123	143	125	495
B-factor analysis (Å ²)					
Overall	22	58	52	50	55
g3p/e4p Chain A	36	20	20	44	57
B		20	21	61	68
C		19	21	46	52
D		19	21	60	84

Table 6. 2 Refinement statistics GAPDHA and GAPDHB mutants. [§] $R = \sum_{hkl} ||F_o(h, k, l) - |F_c(h, k, l)|| / \sum_{hkl} |F_o(h, k, l)|$.

R_{free} is defined as R calculated on a random 5 % non refined data.

6.9 Structural analysis of mutant complexes

6.9.1 GAPDHA c149s + e4p

Although no suitable model could be made for GAPDHA c149s mutant grown in the presence of g3p, an interpretable model was found for the GAPDHA c149s mutant grown in the presence of e4p. Interpretation of the electron density after molecular replacement revealed unambiguous electron density for NADP^+ . Superimposition of the holo-GAPDHA structure from Chapter 3 with that of the c149s mutant using *LSQKAB* (Kabsch, 1976) aligning the whole tetramer reveals little conformational change within the molecule with an RMSD of 0.4 Å for the $\text{C}\alpha$ atoms. The nicotinamide moiety also does not undergo any conformational changes, implying that hydride transfer has not taken place. During refinement of the c149s crystals grown in the presence of e4p analysis of the active site of Chain A revealed, electron density occupying the Ps site. E4p could be modelled into this site with full occupancy, as is shown in Figure 6.9. Residues responsible for contacting e4p in Chain A of c149s are shown in Table 6.3

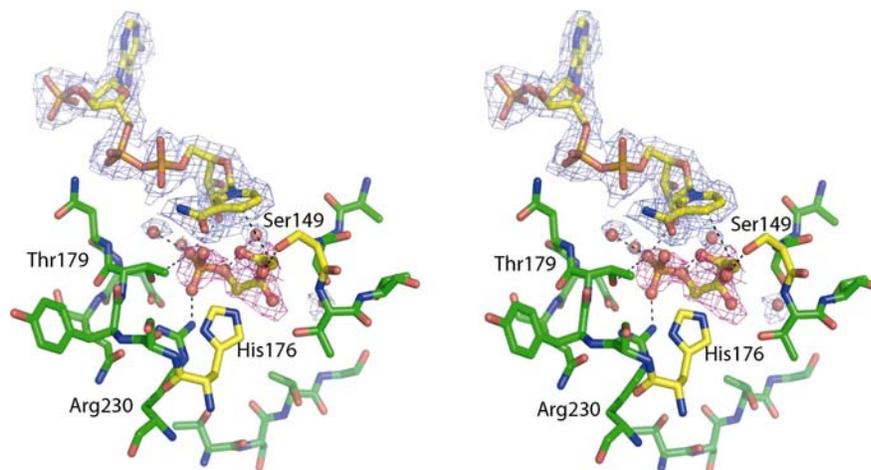
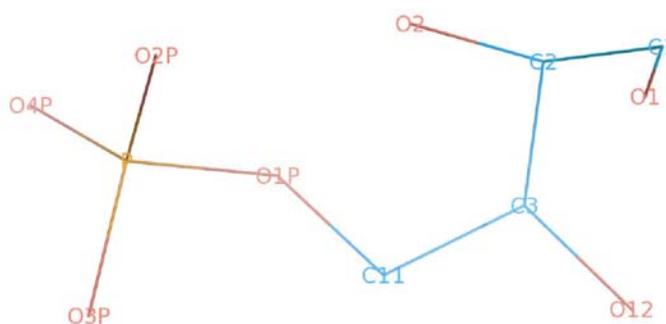


Figure 6.9 Analysis of the electron density within the active site of GAPDHA c149s grown in the presence of e4p. Stereo representation of Chain A is shown, with e4p shown as ball and stick enclosed by an omit map at 1.0σ shown in magenta, calculated as described in Chapter 2.4.13.1 with e4p being omitted from the model. For clarity, the $2F_o - F_c$ map contoured at 1.0σ is also shown, in blue, enclosing NADP^+ and two water molecules within the active site.

Residue	Distance (Å)	e4p
OG Ser149	2.9	O1 e4p
NH2 Arg230	2.2	O3P e4p
H ₂ O5	3.9	O4P e4p
OG1 Thr179	2.9	O3P e4p
H ₂ O287	2.8	O4P e4p

Table 6. 3 Distances for contacts to e4p from Chain A of GAPDHA c149s. Below is the schematic representation for e4p, showing the PDB atom nomenclature.



Chains B, C and D did not contain any interpretable electron density that could be attributed to e4p and therefore, waters were modelled into the active site as shown in Figure 6.10. The greater disorder of Chains B and D could account for why e4p could not be modelled into the active site (packing was analogous to holo-GAPDHA in Chapter 3.9.2). Half of sites reactivity could explain the lack of interpretable electron density for e4p in Chain C, although, half of sites reactivity has been observed only for inhibitors of GAPDHs (Chapter 1.8.1.1.2) and the enzymatic analysis of GAPDHA with e4p suggested it did not act as an inhibitor to g3p turnover. Comparison of Ser149 in Chain B, shows that the hydroxyl group is rotated some 45° away from the active site (Figure 6.10), in a non-attacking position. This is in contrast to Ser149 in Chain A, which is 3 Å from the carbonyl oxygen O1 on e4p (Figure 6.9). The phosphate group of e4p occupying the Ps site formed by Thr179 and Arg230 does not exhibit a high electron density as would be expected by a phosphate group, for comparison the phosphate moiety's of NADP⁺ are very well defined. This suggests that the occupancy is not 100% and this is reflected firstly by the absence of any interpretable e4p electron density in Chains B, C and D. Secondly, a high B-factor for e4p within the model compared to the main chain and NADP⁺, is consistent with lower

occupancy of e4p. Though it should be stressed that B-factors at this resolution can not be used to interpret occupancy.

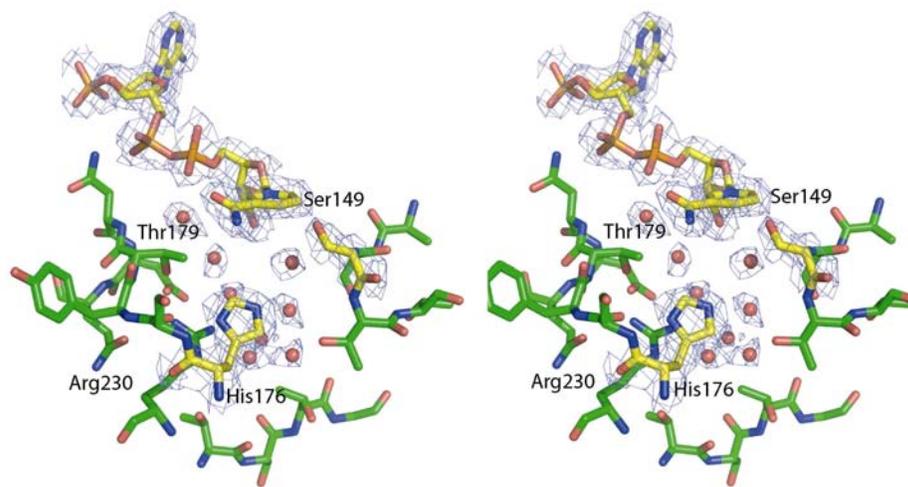


Figure 6. 10 Stereo view of the active site of GAPDHA c149s grown in the presence of e4p. Residues comprising the active site of Chain B are shown, with the Ser149, His176 and NADP⁺ enclosed by a 2F₀-F_C map contoured at 1 σ .

6.9.2 GAPDHB c151s + g3p

Analysis of the electron density after molecular replacement within the active site of GAPDHB c151s mutant grown in the presence of g3p revealed electron density around Ser151 extending into the Ps and Pi sites. G3p was refined into this density and could be built within two different conformations for each chain, with the C3 phosphate of g3p occupying the Ps and the Pi sites. A composite omit map was calculated (as described in Chapter 2.4.13.1) and showed extra un-modelled electron density consistent with a small molecule (g3p) bound in two conformations, shown in Figure 6.11.

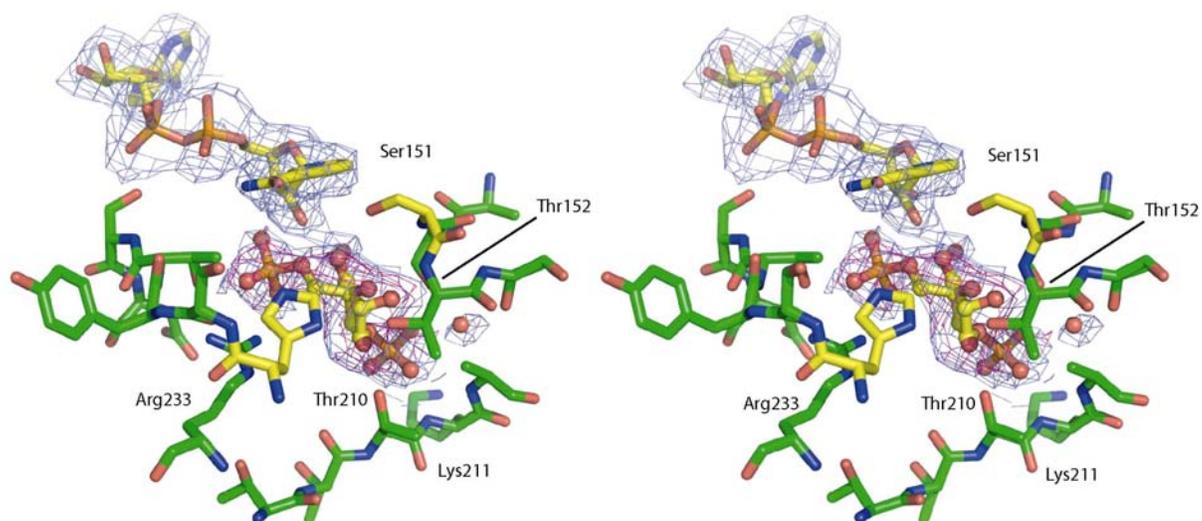
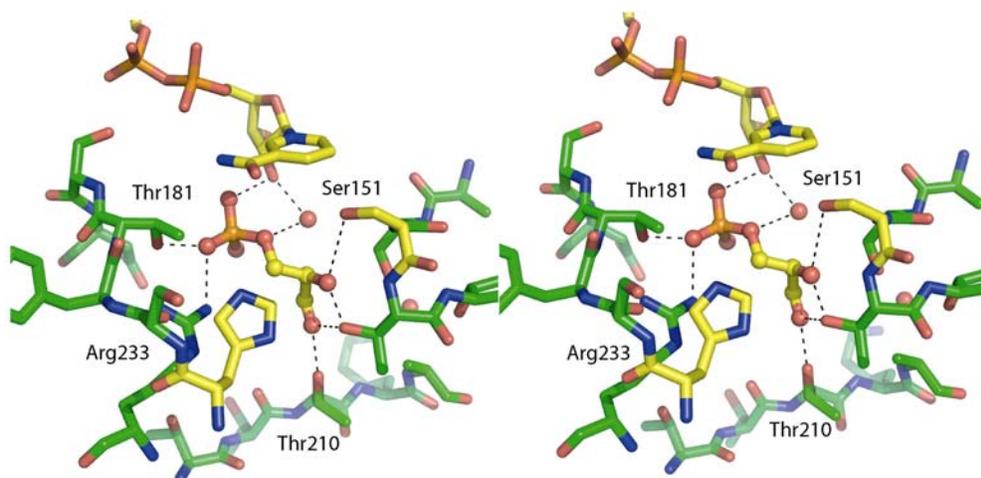


Figure 6. 11 Stereo view of the active site of the GAPDHB c151s mutant grown in the presence of g3p. G3p in two alternative conformations occupying the Ps and Pi sites and is enclosed by an omit map calculated as described in Chapter 2.4.13.1, shown in magenta at 0.8σ . NAD⁺ is shown enclosed by a $2F_0-F_C$ map at 1.0σ shown in blue, which also encloses g3p.

Closer inspection of the g3p molecule occupying the Ps site reveals the hydroxyl groups of Thr152 and Thr210 coordinate the carbonyl oxygen of g3p. This is in contrast as to how e4p is coordinated within the Ps site of the c151s mutant (Chapter 6.9.4) and in the structure form Didierjean and co-workers (Didierjean *et al.*, 2003). The coordination of g3p in both the Ps and Pi sites is shown in Figure 6.12 A and B respectively, with the corresponding distances in Table 6.4.

A



B

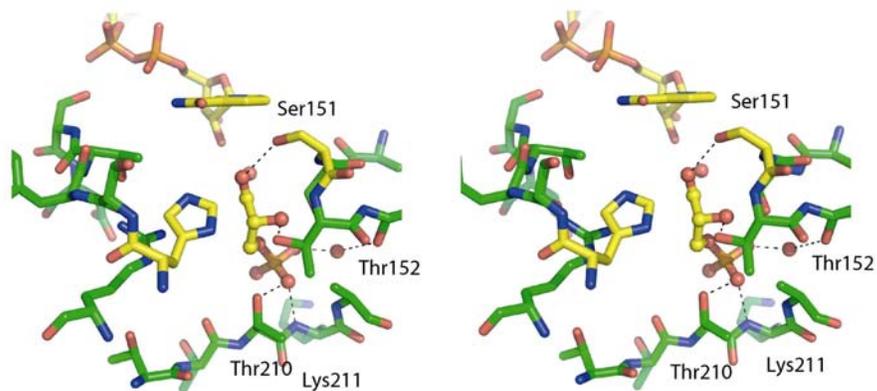
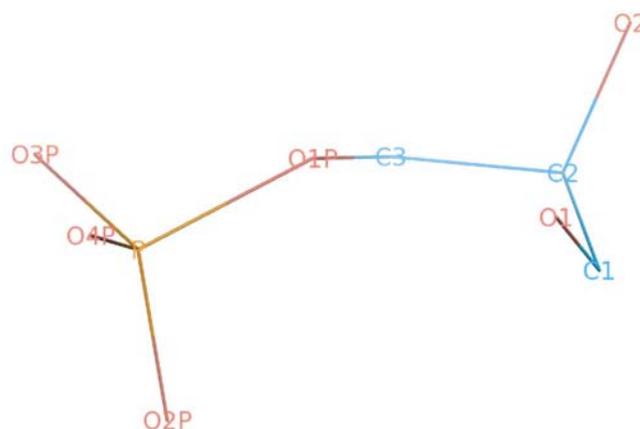


Figure 6. 12 Stereo view of the active site of GAPDHB c151s displaying dual conformations of g3p within the active site. Residues from Chain A are shown with the active site His178 and Ser151 shown as sticks with yellow carbon atoms. NAD⁺ is also shown with yellow carbon atoms as is g3p, which is also shown in, ball and stick representation. Panel A shows g3p occupying the Ps site, where Panel B shows g3p occupying the Pi site.

Residue	Distance (Å)	g3p
Ps site		
OGThr181	2.3	O4P
NH2 Arg233	2.8	O4P
O2D NAD	2.9	O3P
OG Ser151	3.6	O2A
OG1 Thr152	2.2	O2A
OG1 Thr152	2.8	O1A
OG1 Thr210	2.4	O1A
Pi site		
N Lys211	3.0	O3P
OG1 Thr210	2.3	O3P
H2O 73	2.4	O4P
OG1 Thr152	2.0	O2
OG1 Ser151	3.1	O1

Table 6. 4 Distances between GAPDHB c151s Chain A and g3p in the active site, the Pi and Ps sites are defined based upon the discussion in Chapter 1.8.1.1.2. Below is a representation of g3p displaying the PDB atom nomenclature.



6.9.3 GAPDHB c151s + e4p

Analysis of the electron density in the active site for GAPDHB c151s grown in the presence of e4p showed strong electron density around Ser151, with pronounced density for the position where phosphate would occupy the Ps and Pi sites. The same pattern of strong density around the active site and two extra regions corresponding to the phosphate in the Ps and Pi site was noticed in the omit map, calculated as described in Chapter 2.4.13.1, Figure 6.13. This suggested that e4p was bound within the active site of GAPDHB. E4p was refined in dual conformation as was g3p (Chapter 6.9.3). The final refinement values for GAPDHB c151s + e4p are shown in Table 6.2. Superposition of all four chains of the c151s + e4p tetramer showed that the e4p C4 phosphate occupied two positions within the Pi site, in addition the C1 and C2 positions of e4p also occupied a different position. The e4p molecule occupied the same position in all four chains, as shown in Figure 6.14.

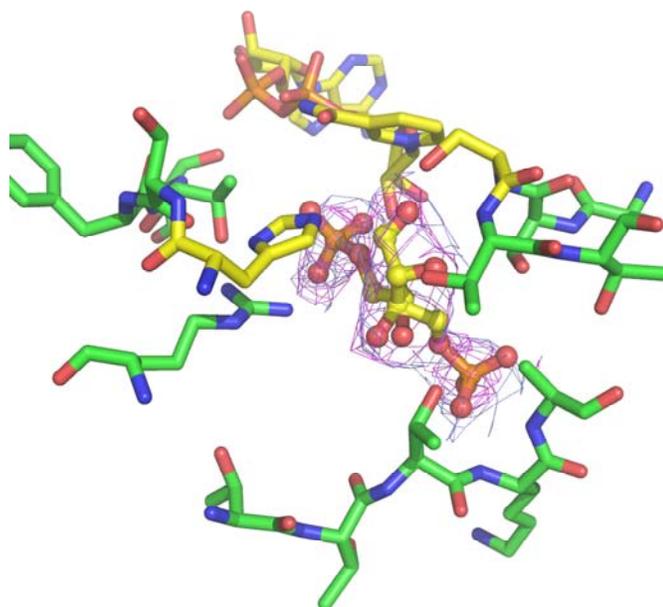


Figure 6. 13 View of the active site of Chain A of the c151s mutant grown in the presence of e4p. E4p is shown in dual conformations enclosed by an omit map (magenta) calculated as described in Chapter 2.4.13.1 and contoured at 1.0σ and a $2F_0 - F_C$ map contoured at 0.8σ .

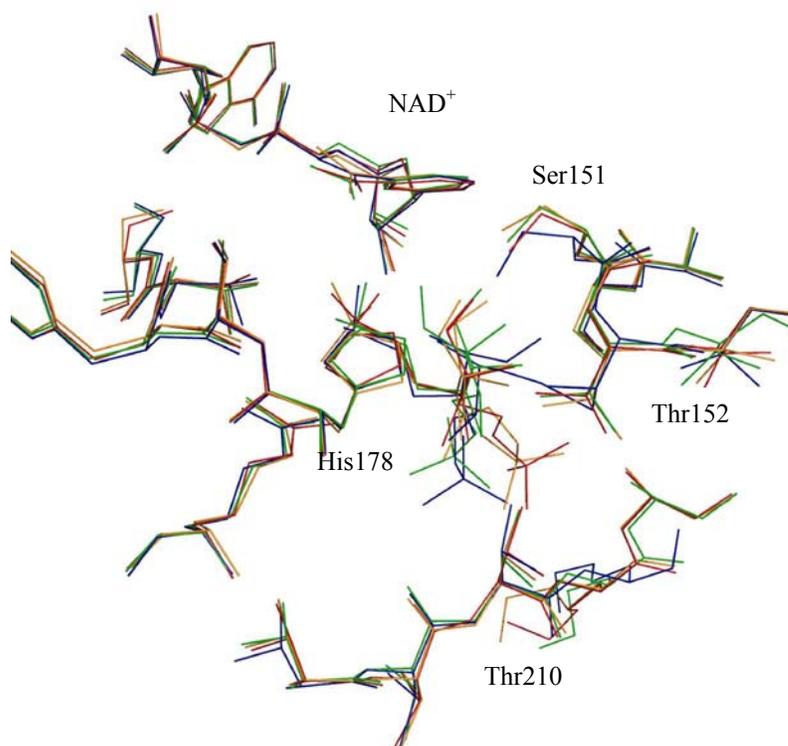


Figure 6. 14 Alignment of Chains B, C, and D onto Chain A of GAPDHB c151s mutant bound to e4p using *LSQKAB* (Kabsch, 1976). Chain A is coloured red, B orange, C, green and Chain D blue.

The residues involved in coordinating the e4p are the same as those that coordinate g3p within the Ps site, although Thr210 (typically forming the Pi site also coordinates the second hydroxyl group of e4p). The Pi site has a greater number of residues that coordinate e4p compared to g3p, as is shown in Table 6.5 and Figure 6.15B. Predominantly the amide groups from Lys211 and Arg212 are involved in the coordination of the C4 phosphate group of e4p. Arg233, involved in the binding of the C4 phosphate within the Ps site provides additional stability to the second hydroxyl group of e4p once it is bound within the Pi site. A thorough discussion of the binding of e4p and g3p is provided in Chapter 6.10.

Ps site		
OGThr181	3.8	O3P
NH2 Arg233	3.6	O3P
O2D NAD	2.2	O4P
OG Ser151	2.4	O1
OG1 Thr152	2.6	O2
OG1 Thr210	2.6	O12
OG Ser150	3.1	O2
Pi site		
OG Ser151	2.4	O1
OG Ser150	3.1	O2
OG1 Thr152	2.6	O2
NH2 Arg233	3.4	O12
OG1 Thr210	3.0	O12
N Lys211	2.4	O3P
OG1 Thr210	2.5	O3P
N Arg212	3.4	O3P

Table 6. 5 Residues involved in contacting e4p whilst it occupies the Ps or Pi site in the GAPDHB c151s mutant. The e4p atom nomenclature is the same as the schematic representation shown in Table 6.3.

From analysis of g3p and e4p binding to the c151s mutant, e4p has a greater number of contacts between the protein once its C4 phosphate occupies Pi site, whilst g3p has a greater number of contacts to the protein whilst the C3 occupies the Ps site. Although this can not prove anything about binding affinities of the substrate to the protein it could imply that the reaction mechanism is different when g3p and e4p are used as substrates and offer an explanation for the ambiguous enzymatic analysis of wild-type GAPDHB in Chapter 5.7. However, to confirm or disprove this hypothesis appropriate pre-steady state kinetics would be required along with further investigation, discussed in Chapter 6.10 and Chapter 7.

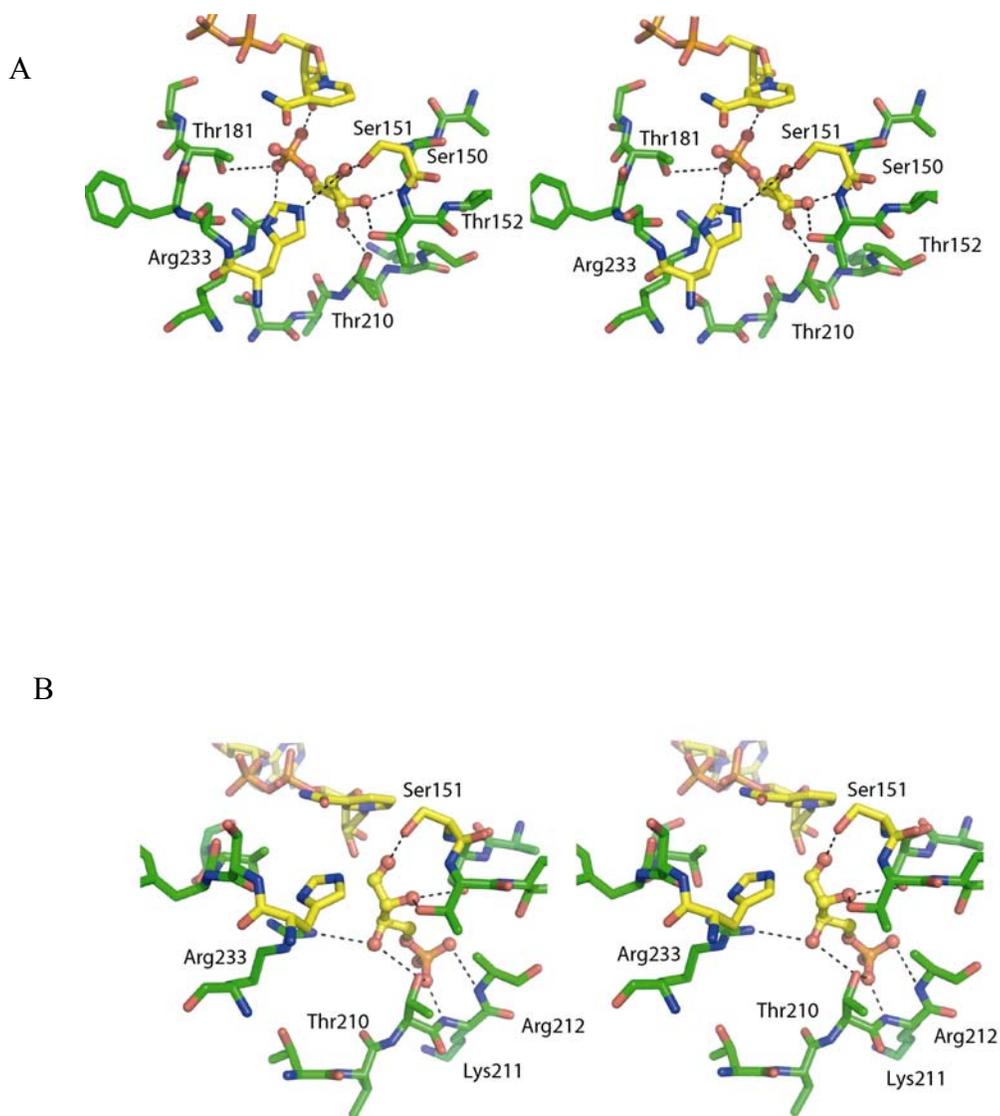


Figure 6. 15 Stereo pairs of the active site of GAPDHB c151s mutant complexed with e4p. Panel A shows the residues coordinating the Ps site, whilst panel B show residues coordinating e4p whilst it occupies the Pi site. For clarity the active site Ser151 and His178 are shown with yellow carbon atoms, as is also NAD⁺ and e4p.

6.9.4 GAPDHB c151a + g3p

Inspection of the c151a mutant in complex with g3p revealed a significant proportion of g3p was bound within the Ps site, with a very small amount bound within the Pi site. G3p would not refine fully whilst the Ps site was 100% occupied and due to the lower resolution, it was not possible to determine the real occupancies between the two sites. Therefore, g3p was refined with equal occupancy in both sites. Figure 6.16 shows g3p bound within the active site of Chain A

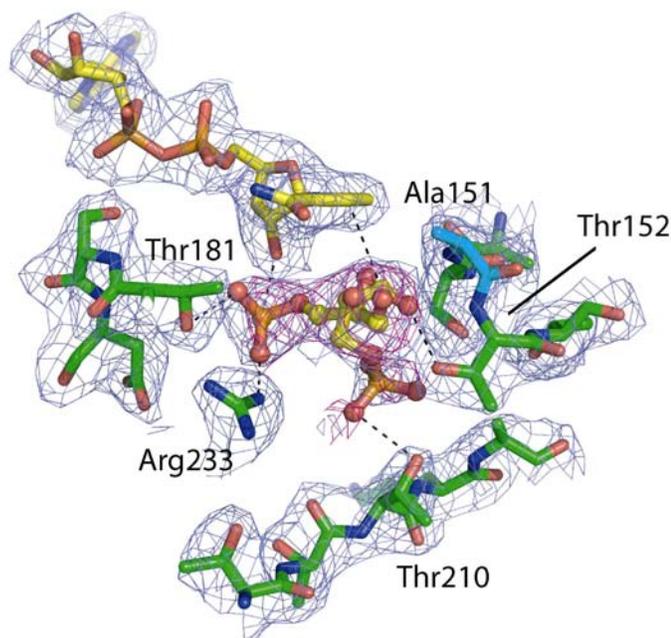


Figure 6. 16 View of the active site of the GAPDHB c151a mutant in complex with g3p. G3p is shown in dual conformations occupying the Ps and Pi sites. The entire model is enclosed by an $2F_0-F_C$ map contoured at 1.0σ shown in blue. An omit map calculated for g3p (as described in Chapter 2.4.13.1) is shown in magenta enclosing g3p and is contoured at 0.8σ .

The g3p occupies the Ps site in a different conformation to g3p bound to the c151s mutant shown in Figure 6.17 and Figure 6.12 respectively. Although the conformation g3p occupies in the c151a mutant is the same as the c149a mutant from *Bacillus stearothermophilus* (1nqo), discussed in Chapter 6.10. There is an additional difference in conformation between the g3p in the Pi site with respect to the c151s mutant.

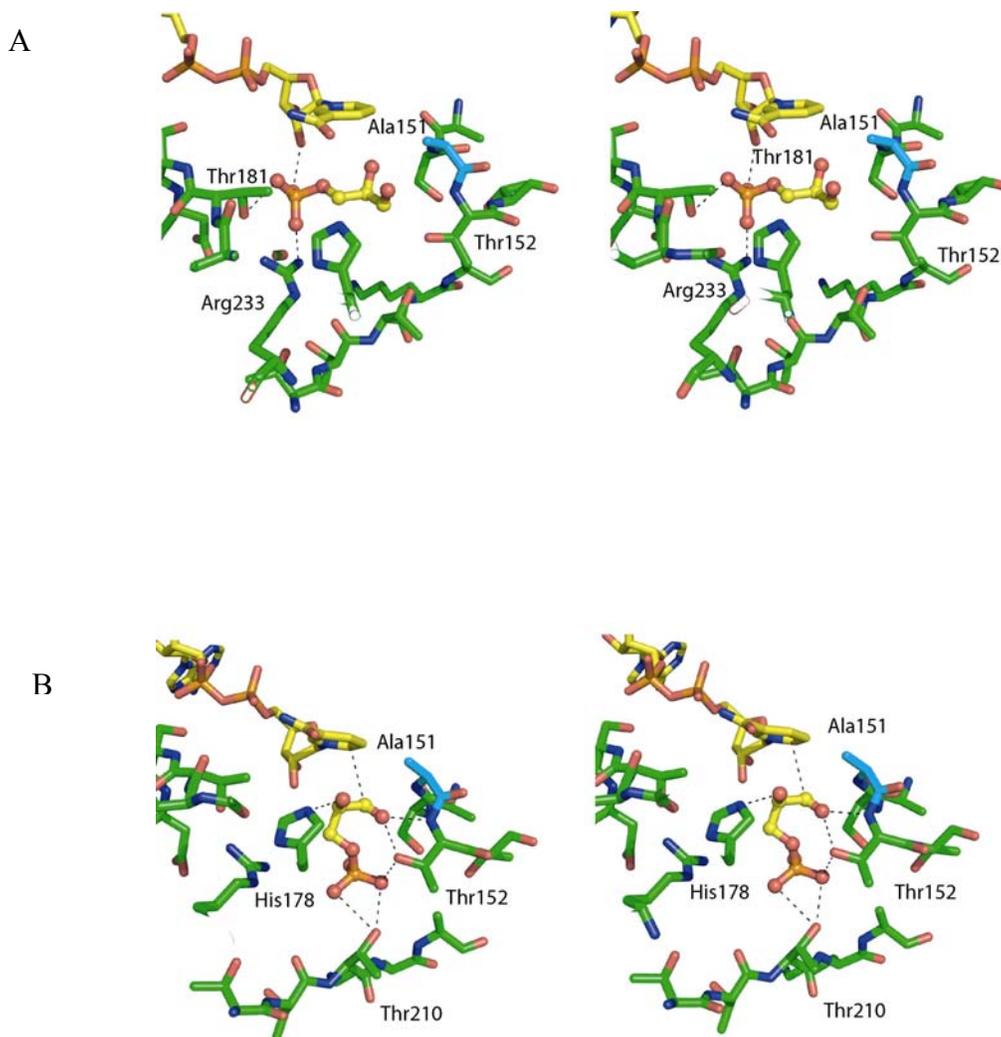


Figure 6. 17 Stereo view of the active site of c151a Chain A. Panel A shows the g3p occupying the Ps site and panel B g3p occupying the Pi site. Alanine 151 is shown with light blue carbon atoms, whilst NAD⁺ and g3p are shown with yellow carbon atoms.

Residue	Distance (Å)	g3p
Ps site		
OG Thr181	2.6	O3P
NH2 Arg233	2.8	O2P
O2D NAD	3.1	O1P
Pi site		
OG1 Thr210	3.0	O3P
OG1 Thr210	2.8	O2P
OG1 Thr152	2.7	O2P
N Thr152	2.9	O1
NE2 His178	3.2	O2
C4N NAD	5.2	C1

Table 6. 6 Residues interacting with g3p table. PDB nomenclature for g3p is shown in Table 6.4.

6.9.5 GAPDHB c151a + e4p

Whereas analysis of the c151a mutant with g3p indicated that most of the g3p was bound within the Ps site, analysis of the electron density of the c151a mutant with e4p indicated a significant amount of e4p occupied the Pi site within Chain A and the entire Pi site in Chain B, as is shown in Figure 6.18.

Residue	Distance (Å)	e4p
Ps site		
OG Thr181	2.7	O4P
NH2 Arg233	2.7	O4P
O2D NAD	2.7	O3P
NE2 His178	3.3	O12
OG1 Thr152	2.5	O12
OG1 Thr152	3.0	O2
H2O 16	2.4	O1
Pi site		
H2O 183	2.9	O3P
OG1 Thr210	3.1	O3P
OG1 Thr210	3.3	O2P
OG1 Thr152	2.5	O4P
OG1 Thr152	2.3	O12
NE2 His178	3.0	O12
H2O 16	2.4	O1

Table 6. 7 Residues involved in contacting e4p whilst it occupies the Ps or Pi site in the GAPDHB c151A mutant.

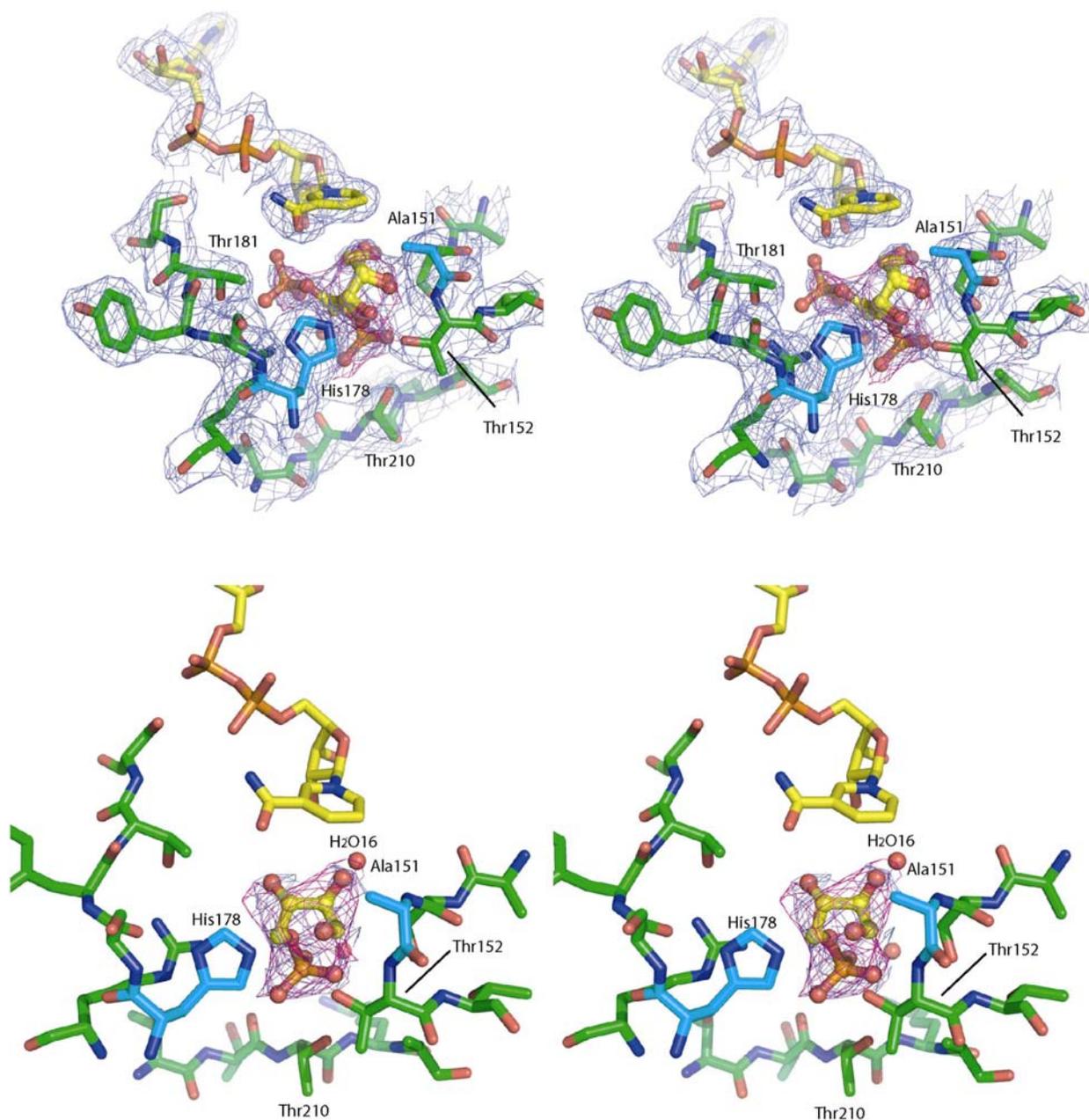


Figure 6. 18 Stereo pair of the active site of the GAPDHB c151a mutant in complex with e4p. Panel A shows Chain A of c151a with e4p occupying the Ps and Pi sites. The active site is enclosed by an $2F_0-F_C$ map shown in blue contoured at 1.0σ . Furthermore e4p is enclosed by an omit map shown in magenta at 0.5σ . Panel B shows the active site of Chain B in the c151a mutant, e4p is again enclosed by an omit map (calculated as described in Chapter 2.4.13.1) contoured at 0.8σ .

Figure 6.19 shows the conformation of e4p in the Ps and Pi sites of c151a Chain A, whilst the contacting residues are summarised in Table 6.7.

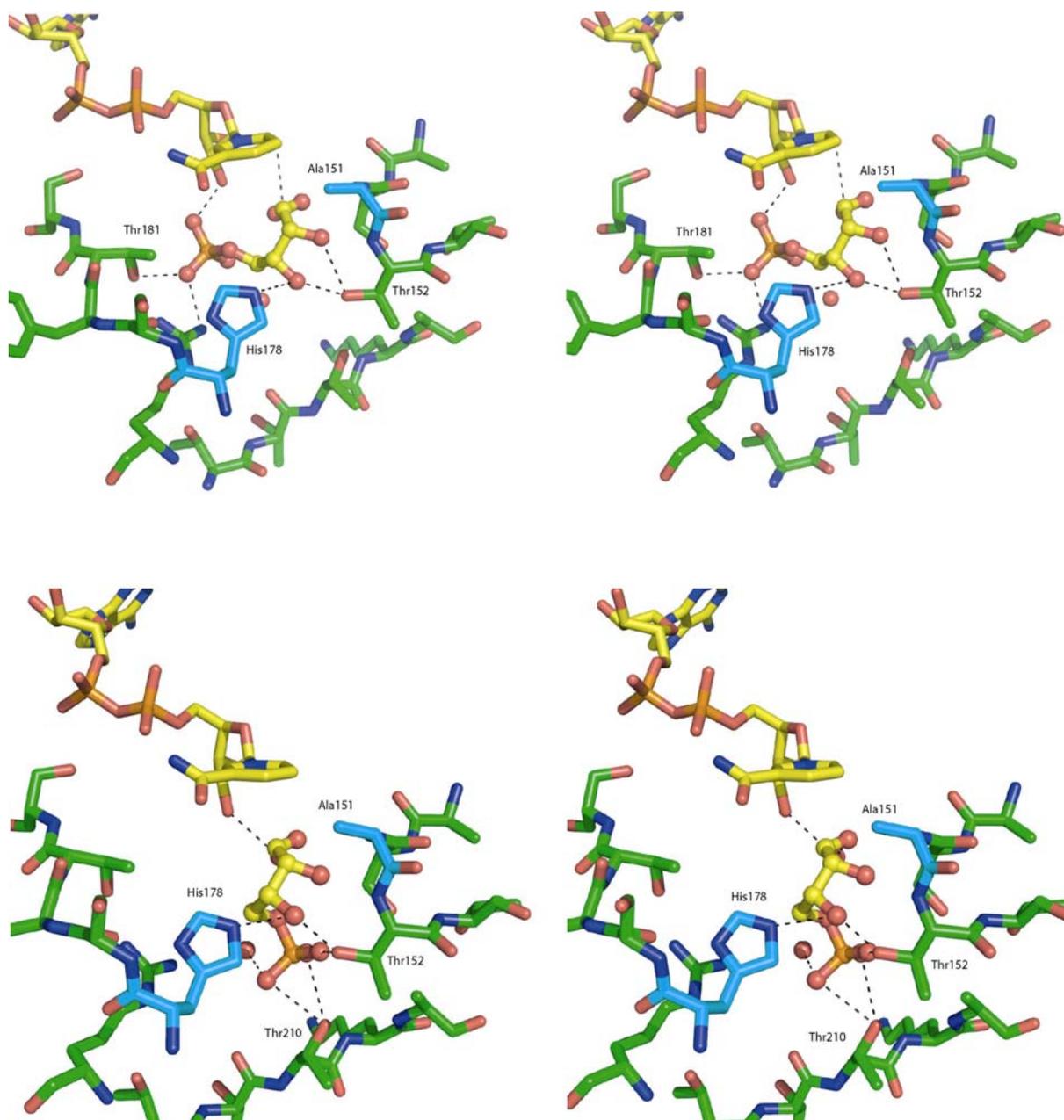


Figure 6. 19 Stereo pairs of the active site of Chain A in the c151a GAPDHB mutant. Panel A shows e4p in the Ps site, whilst panel B shows the conformation of e4p occupying the Pi site. The active site alanine 151 and histidine 178 are shown with blue carbon atoms, whilst NAD⁺ and e4p are depicted with yellow carbons.

6.10 Conclusions

Owing to the low resolution of the mutant structures, it was not possible to discriminate between the binding of g3p or e4p. However, as would be expected, the serine mutation provided coordination of the O1 of g3p and e4p when the substrate phosphate occupied the Pi site. Whilst when the substrate phosphate occupied the Ps site, the serine coordinated the O2 of g3p and e4p. Therefore, the serine mutant can not be regarded as a true Michaelis-Menten complex. Neither can the alanine mutant be regarded as a true complex.

6.10.1 GAPDHB substrate specificity

The poor resolution of the complexes and the additional problem of dual conformations of the substrates imply that at this resolution the specificity of GAPDHB for g3p and e4p can not be determined. Figure 6.20 demonstrates how the two molecules, g3p and e4p appear identical in the mutant structures. It is conceivable that the e4p containing structure actually contains g3p bound from the contamination with the e4p preparation (Chapter 2.1). In addition, even if e4p is bound in the active site it does not show that it is a true substrate for GAPDHB. This is confirmed by the observation of electron density consistent with e4p in the active site of the c149s GAPDHA mutant (Chapter 6.9.2). It could be argued that the addition of a great excess of a compound of similar structure may allow it to bind to the active site. However, a counter argument would be the observation of a racemic mixture of g3p bound within the active site, whereas only the L-enantiomer is observed.

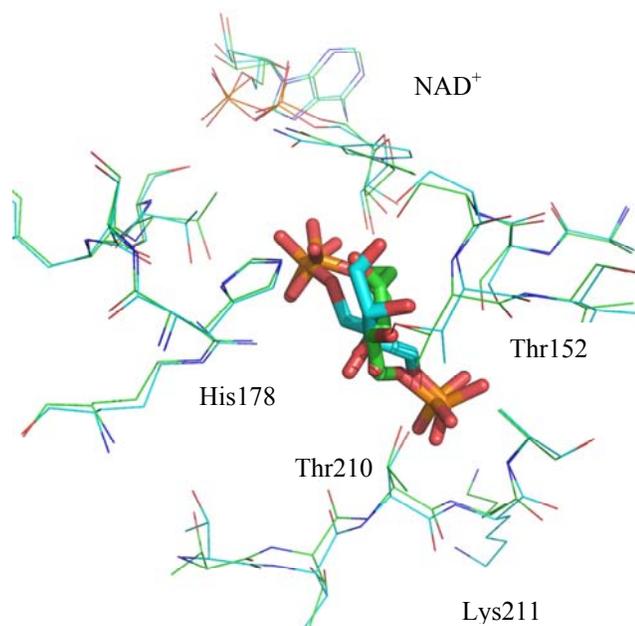


Figure 6. 20 Analysis of the similarities in binding of g3p and e4p to the c151s GAPDHB mutant. g3p is shown as stick with green carbons, whilst e4p shown as stick with cyan carbons.

6.10.2 Analysis with other GAPDH complexes

To date no single crystallographic study has been able to show the Flip-Flop mechanism, rather different structures have been consistent with the Flip-Flop mechanism occurring and this study has been no exception. The structure of g3p complexed with the alanine mutant of *Bacillus stearothermophilus* GAPDH showed that the g3p C3 phosphate occupied the Ps site (Didierjean *et al.*, 2003). A least squares fit of the holo structure by Didierjean *et al* with the alanine GAPDHB mutant shows the two molecules of g3p occupying the Ps site align almost perfectly (Figure 6.21). This helps confirm that the electron density is consistent with g3p at this position.

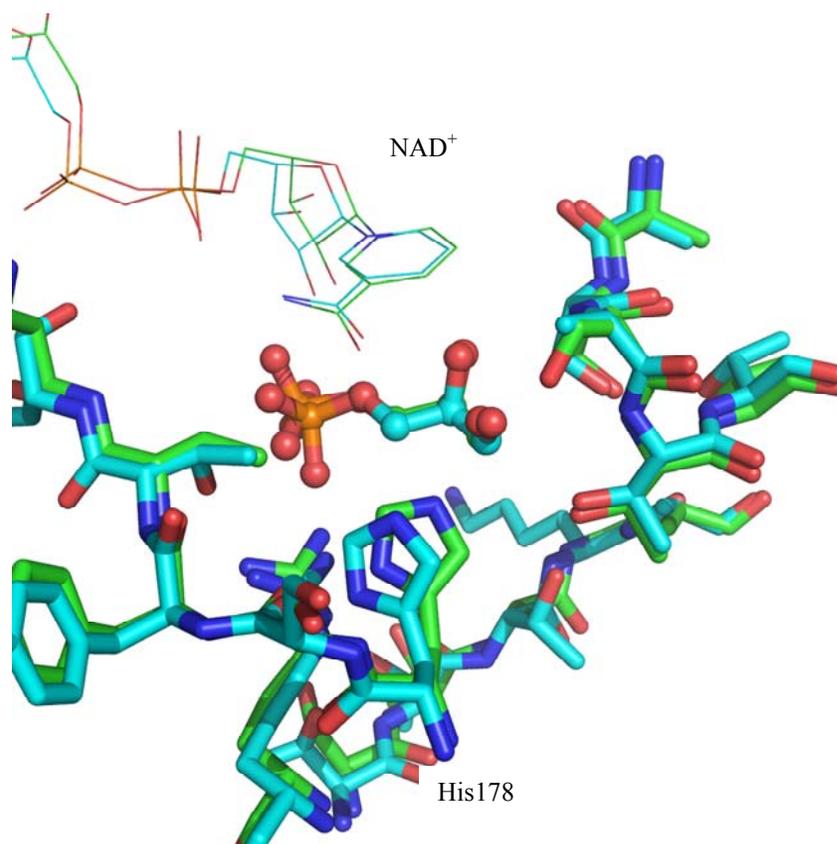


Figure 6. 21 Superimposition of the holo GAPDHB c151a mutant (cyan) with the serine mutant of *Bacillus stearothermophilus* GAPDH determined by (Didierjean *et al.*, 2003) (green). For clarity only the g3p whose C3 phosphate occupies the Ps site in GAPDHB is shown. *LSQKAB* (Kabsch, 1976) was used to superimpose only the catalytic residues of the two enzymes.

Likewise, a superposition of mutant GAPDHB with the *E. coli* GAPDH determined by Yun *et al* in complex with glycerol-3-phosphate (Yun *et al.*, 2000). This shows that the C3 phosphate of g3p occupying the Pi site in the c151a and c151s structures is consistent with the position of the covalent bound *E. coli* apo GAPDH observed in the Pi site (Figure 6.22). In addition, the recent structure of *Bacillus stearothermophilus* holo GAPDH with the thioacyl g3p bound, occupying the Pi site further supports the observation that g3p is bound within the Pi site, in addition to the Ps site.

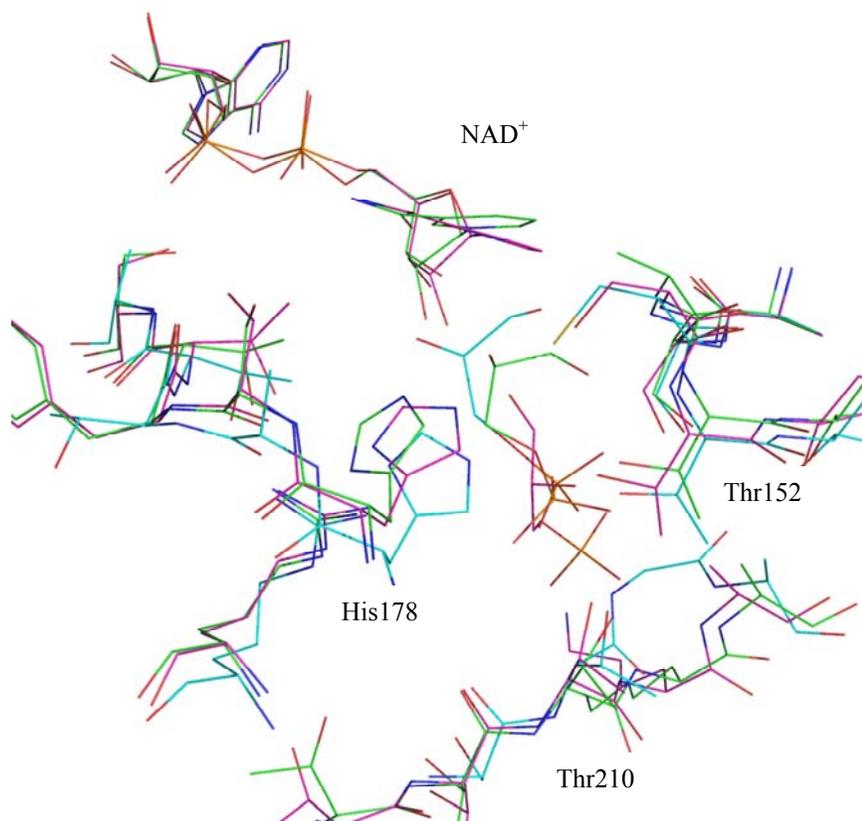


Figure 6. 22 Superimposition of the c151a (green carbons) and c151s (magenta carbons) GAPDHB mutants of *H. pylori* onto the *E. coli* GAPDH containing the covalently bound glycerol-3-phosphate (Yun *et al.*, 2000) (cyan carbons). For clarity only the g3p occupying the Pi site is shown for the GAPDHB mutants.

Analysis of Figure 6.22 demonstrates that the alanine mutant complex resembles more closely the covalent complex observed by (Yun *et al.*, 2000) rather than the serine mutant, highlighting the discrepancies between the two mutants discussed in Chapter 6.10. Of interest is the varied position of His177 within the structures and the apparent orientation of the nicotinamide ring for the c151s mutant GAPDHB. Although the resolution is poor there is the possibility that the c151s mutants are covalently bound to the substrate and an hydroxylacyl intermediate is being observed. If a charge transfer between the nicotinamide ring and Ser151 existed and could be measured within a similar range to the Racker band (Chapter 1.8.1.1.2)(Racker and Krimsky, 1952) then it would be possible to detect whether g3p is bound through the use of a crystal spectrophotometer at the beamline prior to data collection. This would confirm that a covalently bound enzyme existed within the crystals.

6.10.3 *H. pylori* GAPDH and the Flip-Flop reaction mechanism

The complexes of the GAPDHB mutants have not been able to determine mechanisms underlying specificity between g3p and e4p. Rather, the complexes are the first to be observed for a GAPDH where both the Pi and Ps sites are occupied by a substrate phosphate within the same structure. This observation that g3p is able to occupy the Pi and the Ps site whilst NAD⁺ is bound supports the Flip-Flop mechanism described in Chapter 1.8.1.1.2. This demonstrates that during the reaction mechanism the C3 phosphate of the thioacyl intermediate is able to occupy the Pi site allowing for hydride transfer to take place between C1 and the nicotinamide ring. Coenzyme exchange would allow the g3p C3 phosphate to occupy the Ps site, stabilised by new NAD⁺ binding, permitting phosphate binding to the Pi site and subsequent attack of the C1 resulting in collapse of the thioacyl bond and subsequent product release.

Chapter 7 Discussion

7.1 Introduction

Based upon sequence homology two genes encoding for GAPDHs (*gapA* and *gapB*) within *H. pylori* have been annotated as NAD⁺-dependent GAPDHs (Tomb *et al.*, 1997, Alm *et al.*, 1999). This is in contrast to *E. coli*, which contains an annotated gene encoding an NAD⁺-dependent GAPDH and a gene encoding for a non-phosphorylating erythrose 4-phosphate dehydrogenase (Boschi-Muller *et al.*, 1997, Zhao *et al.*, 1995). In addition, *C. jejuni* only has one GAPDH encoding gene (based upon sequence homology) and is also annotated as NAD⁺-dependent. This study has shown that the annotations for the *gap* genes for *H. pylori* and *C. jejuni* are incorrect, highlighting the care that has to be taken with using genomic information to infer an organisms metabolism.

The following sections will summarise the findings of this study and demonstrate how neither genomic analysis nor structural analysis alone can be used to predict functions for identified genes. Instead a combination of the two methods, in addition to other techniques can provide the important questions to ascertain a function and global understanding of an organism's metabolism. A global metabolic understanding and a structural model can provide new avenues for the development of drug-based targets, to treat these enteric pathogens.

7.2 NADP⁺-dependent GAPDHs

7.2.1 Cooperativity

The first *gap* gene to be investigated was *gapA* from *H. pylori*, as described in Chapter 3. Genomic analysis had inferred that *gapA* encoded for an NAD⁺-dependent GAPDH. Enzymatic analysis showed that GAPDHA was NADP⁺-dependent and this was confirmed by the crystal structure of GAPDHA co-crystallised with NADP⁺ (Chapter 3.10.3). The crystal structure had also been obtained for apo GAPDHA from previously grown crystals. Inspection of the nicotinamide binding and catalytic domains of holo and apo GAPDHA revealed a rotation of the nicotinamide binding domain with respect to the catalytic domain upon binding to NADP⁺. This rotation is consistent with the conformational change associated with binding of nucleotide as reported in the *Bacillus stearothermophilus* GAPDH (Leslie and Wonacott, 1984). Furthermore, residues were observed to contact NADP⁺ from the R-axis related subunit, supporting the motion for cooperativity. Analyses of several GAPDHs have shown that negative cooperativity exists

with respect to coenzyme binding (Peczon and Spivey, 1972), although positive cooperativity has been reported for the yeast GAPDH (Gennis, 1976). This study did not attempt to analyse whether positive or negative cooperativity existed for GAPDHA owing to limiting amount of protein and time constraints. However, the use of the calorimetric method, isothermal titration calorimetry (ITC) would show whether GAPDHA displays positive or negative cooperativity.

7.2.2 Determinants of coenzyme specificity

The structural analysis of holo GAPDHA allowed the examination of the residues involved in coordinating NADP^+ , in particular, the examination of residues that provided specificity for NADP^+ over NAD^+ by forming contacts with the 2' phosphate. Chapter 3.10 showed that a hydroxylated residue at position 35 is consistently found in the Rossmann folds of other enzymes, which bind NADP^+ and in GAPDHA this residue is a threonine. The only other GAPDHs that bind NADP^+ are found in the chloroplasts of plants and in *Archaea*. The residues involved in binding can additionally come from the S-loop of the R-axis related subunit. Whereby, lysine 188 in GAPDHA coordinates the 2' phosphate of NADP^+ . In *Spinacia oleracea* this position is occupied by a serine, whilst in *Methanothermus fervidus* the s-loop does not make contact with the NADP^+ (Chapter 3.10.1). Mutational analysis by Branlant and co-workers (Duee *et al.*, 1996) (Didierjean *et al.*, 1997) had identified the hydroxylated position of residue 34 as being important in binding NADP^+ but had wrongly identified residues at position 77 as being an additional determinant for NADP^+ binding (discussed in Chapter 3.10.1). This error was evident from the comparison of GAPDHA with *Bacillus stearothermophilus* GAPDH (Figure 3.14) and GAPDHB (Figure 5.10). Using the information obtained about the specificity for GAPDHA towards NADP^+ , sequence analyses were undertaken to identify further possibly NADP^+ dependent GAPDHs within the *Campylobacterales* order. The phylogenic representation of gap nicotinamide binding domains of species within the *Campylobacterales* yielded two results. Firstly all species within the *Campylobacterales* order had two annotated gap-encoding genes, with the exception of *C. jejuni*. Secondly there appeared to be one NAD^+ dependent and one NADP^+ dependent GAPDH for each species, whilst the only GAPDH for *C. jejuni* was identified as NADP dependent. Enzymatic analysis of purified *C. jejuni* GAPDH showed that it was NADP dependent, although it also exhibited significant activity in the presence of NAD (Chapter 4.3). *C. jejuni* GAPDH only exhibited three fold specificity for NADP over NAD, whilst GAPDHA had twenty fold specificity between the two cofactors (summarised in Table 4.3).

The difference between the absolute specificity of GAPDHA for NADP⁺ and the preference of cGAPDH for NADP⁺ and NAD⁺ could not have been predicted by sequence analysis, although it did suggest that NADP⁺ would be the coenzyme selected. The enzymatic analysis was unable to explain the discrepancies in cofactor specificity. Instead the crystal structure of cGAPDH, cocrystallised with NADP⁺ showed that the NADP⁺ was not coordinated in the same way as NADP⁺ in GAPDHA or *Spinacia oleracea*. The difference in coordination of NADP⁺ in cGAPDH was as a result of the invariant threonine at position 35 not making direct contact to the 2' phosphate group of NADP⁺, as shown in Chapter 4.8.3. In this instance a structural analysis was able to speculate as to the discrepancies noted in the enzymatic analysis. From the enzymatic analysis, coupled with the structural analysis of cGAPDH it can be hypothesised that cGAPDH displays dual cofactor specificity. The global context of this in *C. jejuni* metabolism is discussed in a latter section. Further biophysical analysis using ITC would give a quantitative analysis of the thermodynamic prosperities of cGAPDH binding to NADP⁺ and to NAD⁺.

7.3 Substrate specificity

Studies of two gene products in *E. coli* that code for GAPDH has demonstrated that one gene encoded an NAD⁺-dependent phosphorylating GAPDH, with the other gene (previously annotated as a GAPDH-encoding gene) encoded a non-phosphorylating erythrose-4-phosphate dehydrogenase (Zhao *et al.*, 1995, Boschi-Muller *et al.*, 1997). To test if this was the case for *H. pylori* both GAPDHA and GAPDHB were tested for activity with glyceraldehydes-3-phosphate (g3p) and erythrose-4-phosphate (e4p). GAPDHA displayed a negligible activity with e4p as a substrate and this low level of activity was attributed to the impurity of the e4p purchased from the manufacture (reported at 65-70 % purity, as tested enzymatically by Sigma). Similarly cGAPDH was also tested for activity with e4p using either NAD⁺ or NADP⁺ as cofactors, but for neither cofactor no activity with e4p could be measured. However, e4p activity was observed with GAPDHB. To determine whether or not the reaction was phosphate (or phosphate analogue) dependent, arsenate was omitted from the assay. Table 5.3 shows that there is a detectable activity for GAPDHB with e4p as substrate once arsenate is omitted, in a similar test no activity can be measured with g3p as substrate when arsenate is omitted. When the assay was performed

in the absence of arsenate (and phosphate) the addition of arsenate to the reaction would be expected to cause the reaction to proceed at a rate equivalent to that if arsenate had been present before the reaction commenced. This was the case for when g3p was the substrate for GAPDHB and had additionally been observed when measuring the activities of GAPDHA and cGAPDHB. However, with e4p as a substrate for GAPDHB some activity remained in the absence of arsenate, but the reaction proceeded at a significantly lower reaction rate upon the addition of arsenate compared to if the reaction had not been allowed to proceed in the absence of arsenate. The only explanation which can be attributed to this is slight inhibition of GAPDHB through another reaction catalysed in a phosphate-independent reaction (Chapter 5.7). This enforces the notion of the limitation of the kinetic assay in determining the enzymatic parameters of GAPDHs. Since there is no alternative assay available for measuring the catabolic reaction and with the reverse reaction (catalysing the anabolic reaction) being difficult to measure accurately (as it has to be coupled through other enzymes in the glycolytic pathway due to the instability of the 1,3-bis phosphoglycerate (Chapter 2.5)) then only the context of the enzymatic analysis can be used.

The e4p dependence of GAPDHB would require further investigation. One such method to resolve this would be product analysis from the reaction between holo GAPDHB and e4p in the presence and absence of phosphate. Classically, the products from enzymatic reaction of GAPDHs with substrates were resolved by thin layer chromatography (Ishii *et al.*, 1964). The use of HPLC using an anion exchange column would be the most effective way to analyse the products of the reaction. This analysis would allow the discrimination between mono- and bis-phosphorylated sugars and resolve the extra alcohol group in the four carbon sugar. Further kinetic analysis of the enzyme, focusing on determining if the rate limiting step is still the NADH exchange would be carried out using stopped-flow kinetics. Inspection of the enzymatic parameters shown in Chapter 5.7 suggests that there are two different reaction mechanisms for g3p and e4p as substrates, depending when phosphate is present or not. An understanding of classical biochemical pathways would dictate that the activity for GAPDHB with e4p should be non-phosphorylating, since there are no reports of 1,4-bisphosphoerythronate within the literature whereas the 4-phosphoerythronic acid is reported and is an intermediate in pyridoxol 5'-phosphate biosynthesis furthermore this function has been attributed to GAPDHB in *E. coli* (Zhao *et al.*, 1995). However, there was no indication that GAPDHA and cGAPDH were phosphorylating,

NADP⁺-dependent GAPDHs, thus care must be taken with the assumption that GAPDHB also has to display non-phosphorylating erythrose 4-phosphate dehydrogenase activity.

Since the enzymatic analysis of GAPDHB was unable to provide an understanding of the physiological role of its reactions with e4p and whether or not it is truly phosphorylating or merely activated by phosphate, a mutagenic approach was undertaken to allow for the cocrystallisation of GAPDHB with NAD⁺ and either g3p or e4p. It was hoped that the crystal structure of e4p within the active site would unambiguously show if it was a true substrate of GAPDHB. The binding of e4p within the active site of GAPDHB was to be compared to the binding of g3p in the active site. To allow for the ternary complex analogue formation the active site cysteine was mutated to both alanine and serine. Chapter 6 focused on the mutagenic approach of the GAPDHA and GAPDHB from *H. pylori*. Owing to time constraints this mutagenic approach was not undertaken with cGAPDH and even though GAPDHA, like cGAPDH did not use e4p as a suitable substrate, the structures of GAPDHA cocrystals would be a negative control, if e4p was not observed within the structure. However, the structures of the GAPDHB mutants in complex with either g3p or e4p were unable to demonstrate substrate specificity owing to the low resolution and dual conformers of the substrates in the active site. Instead of addressing substrate specificity the structures of the mutant complexes of GAPDHB was able to provide structural evidence for the Flip-Flop reaction mechanism of GAPDHs, which has been observed in other GAPDH structures.

7.4 Analysis of GAPDH reaction mechanism

As mentioned in Chapter 7.3 the mutagenic approach undertaken in Chapter 6 was to address the issue of substrate specificity between g3p and e4p by GAPDHB. Owing to the low resolution of the structures it was not possible to differentiate between g3p and e4p. The structures of the GAPDHB c151a and c151s mutants co-crystallised with NAD⁺ and either g3p or e4p instead revealed electron density within the active site consistent with either g3p or e4p occupying a dual conformation. Other structures of GAPDHs with substrates have focused along two main methods. They are, soaking or co-crystallisation of substrate or inhibitor, or a

mutagenic approach permitting the co-crystallisation of substrates. The mutagenic approach was used by Didierjean *et al* whereby g3p was observed bound with the C3 phosphate occupying the Ps site in the holo GAPDH from *Bacillus stearothermophilus*. It should be noted that the electron density for the C3 phosphate was not as intense as the phosphates from NAD⁺ and this may reflect that there may have been some dual conformation not modelled by (Didierjean *et al.*, 2003). The soaking of substrates into wild-type crystals has met with limited success. For the *E. coli* GAPDH some dual conformation was noted when g3p was soaked into apo GAPDH crystals (Yun *et al.*, 2000). Whilst for the *Trypanosoma cruzi* GAPDH the co-crystallisation of inhibitors revealed dual conformation further supporting the Flip-Flop mechanism (Castilho *et al.*, 2003). More recently, Branlant and co-workers soaked g3p into wild-type *Bacillus stearothermophilus* holo GAPDH crystals and observed the thioacyl intermediate with its C3 phosphate occupying the Pi site (Moniot *et al.*, 2008). The work in Chapter 6 clearly shows that dual conformations of substrate exist in the mutated GAPDHB, further supporting the Flip-Flop mechanism.

7.5 Metabolic inferences within the Campylobacterales order

This study has inferred that using genomic analysis alone is not sufficient to understand a global metabolic role within an organism. Throughout this study the identification of NADP⁺-dependent GAPDHs and also a putative GAPDH with e4p activity has permitted the consideration of the role of that particular GAPDH within the bacteria and its significance for the bacteria's metabolism. Further experiments whereby the genes for these GAPDHs are "knocked-out" and the viability on various media and micro-array testing of the gap genes expression during different stages of metabolism would help to answer these questions. Experiments investigating two GAPDHs from the *Thermoproteus tenax* have been able to demonstrate the expression of one GAPDH in autotrophic conditions (Brunner *et al.*, 2001). Northern and Western blot analysis of cyanobacter *Synechocystis* GAPDHs have been able to show which GAPDH is expressed in chemoheterotrophic conditions (Valverde *et al.*, 1997). A similar approach to the two outlined would be able to identify if GAPDHA is expressed in gluconeogenic conditions i.e. medium consisting of amino acids and no sugars and whether GAPDHB is expressed under glycolytic conditions of glucose rich medium. Applying a similar argument for the study of *cgap* in *C. jejuni* would only infer if the organism used a predominantly glycolytic or gluconeogenic role.

7.6 Design of drug targets for GAPDHs

Once the viability of the bacteria after knock-out or mutational studies of the *gap* genes had been determined the possibility of GAPDH as a viable drug target for the treatment of *C. jejuni* or *H. pylori* would be investigated. A structure based drug design strategy would be employed to lead to the design of inhibitors against GAPDH. Of particular interest GAPDHA would be an ideal drug target with it being NADP⁺-dependent and the human and gut *E. coli* contain strictly NAD⁺-dependent GAPDHs. Likewise, cGAPDH could offer an interesting drug target. The design of inhibitors against cGAPDH or GAPDHA targeting the coenzyme binding. A similar approach has been undertaken with the *Trypanosoma cruzi* glycosomal GAPDH whereby the inhibitor chalepin inhibits the enzyme through preventing NADP⁺ binding (Pavao et al., 2002).

Therefore, the atomic understanding of a structure can provide insight into its molecular function which can lead to the design of further experiments to complement existing biochemical and cell-based investigations. The underlying principle is that even though a family of structures may be well characterised, a thorough understanding of a structure and mechanism is enough to permit speculation and investigation into its function within the organism and lead to the rational for drug-based design, once complemented by other investigations.

Publications during thesis

Acta Crystallographica Section F

**Structural Biology
and Crystallization
Communications**

ISSN 1744-3091

**Paul R. Elliott, Daniel Evans,
Jacqueline A. Greenwood and
Peter C. E. Moody***Henry Wellcome Laboratories for Structural
Biology, Department of Biochemistry, University
of Leicester, Leicester LE1 9HN, England

Correspondence e-mail: pce1@leicester.ac.uk

Received 24 February 2008

Accepted 2 July 2008

Expression, purification, crystallization and preliminary X-ray analysis of an NADP-dependent glyceraldehyde-3-phosphate dehydrogenase from *Helicobacter pylori*

The classical glycolytic pathway contains an NAD-dependent glyceraldehyde-3-phosphate dehydrogenase, with NADP-dependent forms reserved for photosynthetic organisms and archaea. Here, the cloning, expression, purification, crystallization and preliminary X-ray analysis of an NADP-dependent glyceraldehyde-3-phosphate dehydrogenase from *Helicobacter pylori* is reported; crystals of the protein were grown both in the presence and the absence of NADP.

1. Introduction

Helicobacter pylori, a micro-aerophilic Gram-negative rod-shaped bacterium, is a dangerous human pathogen that colonizes the upper gastrointestinal tract, resulting in gastritis (Marshall & Warren, 1983). *H. pylori* infection has been implicated in a variety of diseases ranging from peptic ulcers to gastric cancer and mucosal-associated lymphoma (Kuipers *et al.*, 1995; Eidt *et al.*, 1994). Patients infected with *H. pylori* have a 20% lifetime risk of developing peptic ulcer disease and a 2% risk of developing gastric cancer (Ernst, 2000; Kuipers, 1999). *H. pylori* infection is more common in developing countries than in developed countries, with 80% of the population in developing countries carrying the bacteria compared with 25–50% of the population in developed countries (Dunn *et al.*, 1997). Current treatment of *H. pylori* consists of antibiotic treatment coupled with proton-pump inhibitors. Since these treatments have a variety of side effects, and with the onset of antibiotic resistance, new treatments are required (Mendz *et al.*, 1995).

Published analysis of the complete genome of *H. pylori* suggested that its only carbohydrate source is glucose. Furthermore, this analysis suggested that glucose is the main source for substrate-level phosphorylation (Tomb *et al.*, 1997). The *in vitro* growth requirements of the bacteria show that no other carbohydrates are required for growth (Mendz & Hazell, 1993) and the only carbohydrate permease present is glucose-specific (Alm *et al.*, 1999). However, the genes for the key enzymes in glycolysis appear to be absent (Alm *et al.*, 1999; Tomb *et al.*, 1997). The Entner–Doudoroff pathway is an alternative route for glucose catabolism; this pathway is constitutively induced in *H. pylori* (Mendz *et al.*, 1994; Wanken *et al.*, 2003) and has been shown to be the main pathway for the oxidation of glucose (Chalk *et al.*, 1994). Genes for the other carbohydrate metabolic pathways such as the pentose-phosphate shunt and the majority of the gluconeogenic enzymes appear to be present; furthermore, activity has been detected *in vivo* for the pentose-phosphate pathway (Mendz *et al.*, 1993). One common link between all the carbohydrate pathways is glyceraldehyde-3-phosphate. Glyceraldehyde-3-phosphate can feed into the pentose-phosphate shunt for the generation of key five-carbon sugars that are required for DNA biosynthesis. The pentose-phosphate shunt also produces NADPH, which is required for biosynthetic processes and the prevention of oxidative stress. Glyceraldehyde-3-phosphate also acts as a link from the Entner–Doudoroff pathway into glycolysis/gluconeogenesis, whereby glyceraldehyde-3-phosphate undergoes reversible oxidative phosphorylation by glyceraldehyde-3-phosphate dehydrogenase (GAPDH) to



© 2008 International Union of Crystallography
All rights reserved

crystallization communications

give NADH and 1,3-bisphosphoglycerate. Therefore, an understanding of the pathogen's GAPDH enzyme is crucial to understanding the core metabolism of this organism.

This investigation studied the product of one of the two putative GAPDH genes identified in *H. pylori*, *gapA* (gi:6626253; NCBI, NIH); the gene product from *gapA* is a tetrameric (4×36 kDa) enzyme called GAPDHA. Biochemical analysis, which will be presented elsewhere, is consistent with the dependence of GAPDHA on NADP and raises interesting questions about the function of glucose metabolism in *H. pylori*. GAPDHA has been cloned, expressed, purified and crystallized with and without NADP. Preliminary electron-density analysis was consistent with NADP bound at the active site only when NADP was cocrystallized with GAPDHA.

2. Materials and methods

2.1. Cloning and overexpression

The full nucleotide sequence encoding GAPDHA was cloned into pET151/D (Invitrogen) containing an N-terminal His₆ tag linked by a TEV protease site. The forward primer was CACCATGCCAAT-TAGAAT and the reverse primer was TATATAGCACAAAATTAA. Dideoxy nucleotide sequencing confirmed the presence of the full-length *gapA* sequence in the vectors.

Escherichia coli strain Rosetta DE3 transformed with pET151/D-*gapA* was grown at 303 K in the rich medium 2YT supplemented with $100 \mu\text{g ml}^{-1}$ ampicillin and $35 \mu\text{g ml}^{-1}$ chloramphenicol. Upon reaching an OD₆₀₀ of 0.6, the culture was cooled to 291 K and incubated overnight at this temperature; isopropyl β -D-1-thiogalactopyranoside induction was not found to be necessary for sufficient overexpression in this case. The cells were harvested the following morning and frozen prior to purification.

2.2. Purification

The purification of His₆-tagged GAPDHA took place as follows. Cell pellets were thawed in buffer A (20 mM Na₂HPO₄, 500 mM NaCl, 50 mM imidazole pH 7.4) supplemented with protease-inhibitor cocktail VII (Calbiochem). The suspension was sonicated and cell debris was removed by centrifugation. The supernatant was loaded onto a 5 ml Hi-Trap Nickel Sepharose column (Amersham Biosciences) equilibrated in buffer A and eluted against a linear gradient of buffer B (20 mM Na₂HPO₄, 500 mM NaCl, 500 mM

imidazole pH 7.4). Fractions with enzymatic activity were pooled and dialysed extensively against buffer C (20 mM Na₂HPO₄, 50 mM NaCl, 1 mM DTT pH 7.2). To facilitate cleavage of the hexahistidine tag, approximately half of the sample was incubated overnight with TEV protease as stipulated in the manufacturer's instructions (Invitrogen). Samples of cleaved and uncleaved GAPDHA were separately loaded onto a 5 ml Hi-Trap Sulfoethyl Sepharose cation-exchange column equilibrated with buffer C and eluted against a linear gradient of buffer D (20 mM Na₂HPO₄, 1 M NaCl, 1 mM DTT pH 7.2). A single peak was collected at 550 mM NaCl for both the His-tagged and untagged GAPDHA and these were judged to be ~95% pure by SDS-PAGE analysis. The final yield of protein was approximately 10 mg per litre of culture.

Untagged GAPDHA and His₆-GAPDHA were concentrated to 12 mg ml^{-1} using Amicon Ultra-15 centrifugal filter units (10 kDa molecular-weight cutoff; Millipore) and buffer-exchanged into 20 mM MES, 100 mM NaCl, 1 mM DTT pH 6.5 prior to crystallization.

2.3. Crystallization and data collection

Crystal-growth conditions for untagged GAPDHA were screened by the sitting-drop vapour-diffusion technique using 100 nl drops of protein solution mixed with 100 nl precipitant. The drops were dispensed from a Genomic Solutions Cartesian Honeybee 8+1 (Harvard Bioscience) onto 96-well MRC plates (Innovadyne) with reservoirs containing 80 μl precipitant from the crystal screen kits Wizard I and II and Cryo I and II (Emerald Biosciences) in a humidified chamber (a total of 192 conditions). Plates were sealed with transparent tape and monitored for crystal growth using CrystalProHT (TriTek) plate-storage and imaging systems at 293 and 277 K. Crystallization hits at 293 K were recorded approximately 12 h after the plates were sealed from conditions containing non-ionic precipitants. Refinement of these conditions focused on the higher molecular-weight PEGs (as these appeared to be the most promising) and was performed using a Tecan 75 liquid-handling system (Tecan) and larger crystallization drops dispensed from the Cartesian Honeybee 8+1. The condition that consistently gave reproducible diffraction-quality crystals consisted of 500 nl untagged GAPDHA (12 mg ml^{-1}) mixed with an equal volume of precipitant [100 mM HEPES pH 9.0, 17% (w/v) PEG 8000] equilibrated against a reservoir containing 80 μl precipitant. Crystals grew after one week at 293 K (Fig. 1).

The strategy for exploring crystallization conditions for His₆-GAPDHA (in the presence of 1 mM NADP) was the same as described for untagged GAPDHA. Refinement of the initial hit [34% (v/v) ethanol, 100 mM Tris pH 7.2] was performed using a Tecan 75 liquid-handling system (Tecan) to explore the effects of different alcohols and pHs. The final refined condition consisted of 500 nl His₆-GAPDHA (12 mg ml^{-1}) containing 1 mM NADP mixed with an equal volume of reservoir comprising 100 mM Tris pH 7.3, 36% (v/v) methanol. Crystals grew after 72 h at 293 K with a clear dependence on the pH of the precipitant (Fig. 2). Attempts to grow crystals of untagged GAPDHA in the presence of NADP were unsuccessful. Similarly, we found no conditions in which crystals of His₆-GAPDHA would grow in the absence of NADP.

Crystals of untagged GAPDHA (Fig. 1) were prepared for cryo-crystallography by transferring a single crystal in a litho-loop (Molecular Dimensions) into a final concentration of 15% (v/v) glycerol in reservoir solution and flash-cooling to 110 K in a stream of boiled-off liquid nitrogen. This method was not successful for His₆-GAPDHA crystals grown using methanol owing to the rapid



Figure 1
Crystals of untagged apo-GAPDHA grown in 100 mM HEPES pH 9.0, 17% PEG 8000. Crystals grew to $200 \times 50 \mu\text{m}$.

crystallization communications

Table 1
Summary of data-collection statistics for apo- and holo-GAPDHA.

Values in parentheses are for the highest resolution shell.

	GAPDHA	His ₆ -GAPDHA + NADP
Wavelength (Å)	0.933	0.934
Space group	<i>P</i> 2 ₁	<i>P</i> 3 ₂
Unit-cell parameters (Å, °)	<i>a</i> = 75.2, <i>b</i> = 100.6, <i>c</i> = 97.8, α = 90.0, β = 93.7, γ = 90.0	<i>a</i> = 116.9, <i>b</i> = 116.9, <i>c</i> = 95.4, α = 90.0, β = 90.0, γ = 120.0
Resolution limits (Å)	95.4–1.74 (1.82–1.74)	69.5–2.6 (2.74–2.60)
No. of observations	308994 (15449)	158217 (15086)
No. of unique observations	147140 (7357)	43460 (5857)
Completeness	99.0 (99.0)	96.8 (88.3)
$\langle I/\sigma(I) \rangle$	19.0 (2.2)	17.9 (5.7)
R_{merge}^\dagger	0.161 (0.534)	0.045 (0.193)

$^\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an individual measurement of the reflection with Miller indices hkl and $\langle I(hkl) \rangle$ is the mean intensity of that reflection.

evaporation of methanol upon removal of the plate seal, which caused the crystals to dissolve. This could not be alleviated by harvesting at lower temperatures or attempting to create a methanol-saturated environment. To overcome this problem, various cryoprotectant solutions were injected onto the sitting drops by piercing through the transparent tape with a Hamilton syringe, which allowed efficient harvesting and cryocooling of the crystals prior to data collection. The ability of different cryoprotectants to protect the crystal during the cryocooling process was monitored by obtaining diffraction images of two orientations of the crystal using home-source X-rays and detector (Rigaku RU2HB, Xenocs optics, R-AXIS

IV detector). The optimum cryoprotection (as judged by diffraction quality) was afforded by a mixture of 35% (v/v) 2-methyl-2,4-pentanediol and 100 mM Tris pH 7.2.

A complete data set was collected from a single cryocooled crystal of untagged GAPDHA on beamline ID14-2 at the ESRF. An ADSC Q4R CCD detector was used (151.3 mm crystal-to-detector distance) to collect 360 images of 0.5° oscillation at a fixed wavelength of 0.933 Å (a typical image is shown in Fig. 3a). Intensities were measured using *MOSFLM* (Leslie, 1992), with the autoindexing routines giving a solution consistent with a primitive monoclinic cell; examination of the principal axes showed an intensity distribution consistent with space group *P*2₁ (statistics are given in Table 1). Solvent-content calculations using four molecules of untagged GAPDHA in the asymmetric unit indicated 50.7% solvent content (V_s) with a Matthews coefficient (V_M) of 2.51 Å³ Da⁻¹ (Matthews, 1968). This is in agreement with all known GAPDH structures, which are also tetramers.

In the case of His₆-GAPDHA (grown in the presence of NADP), diffraction data were collected on beamline ID14-1 at the ESRF with an ADSC Q4R CCD detector (300.7 mm crystal-to-detector distance). 120 images were collected with 0.5° oscillation width at a fixed wavelength of 0.934 Å for a single cryocooled crystal (a typical image is shown in Fig. 3b). Analysis using the autoindexing routines of *MOSFLM* suggested a primitive trigonal cell. The data were scaled using *SCALA* (Evans, 2006) in space group *P*3 (statistics are given in Table 1) and examination of the distribution of intensities along the principal axes indicated either a 3₁ or a 3₂ screw axis. Matthews analysis (Matthews, 1968) suggested the presence of a tetramer in the asymmetric unit ($V_M = 2.5$ Å³ Da⁻¹; $V_s = 50.9\%$ solvent). Molecular replacement using PDB entry 1gd1 (GAPDH from *Bacillus stearo-*

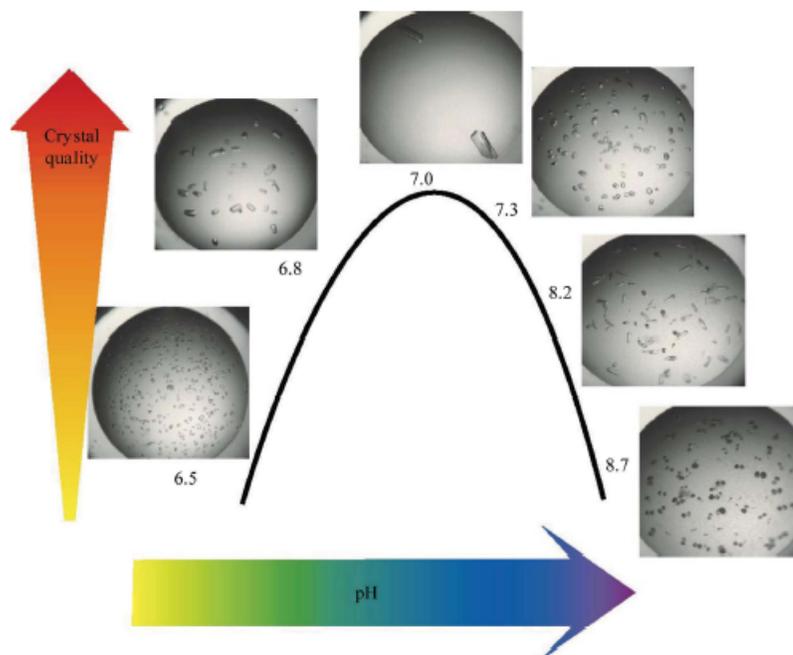


Figure 2

Crystals of His₆-GAPDHA grown in the presence of NADP. It was observed that as the pH increased the quality of the crystals increased until an optimum was reached, after which the quality decreased, with multiple nucleations occurring at the ends of the crystals.

crystallization communications

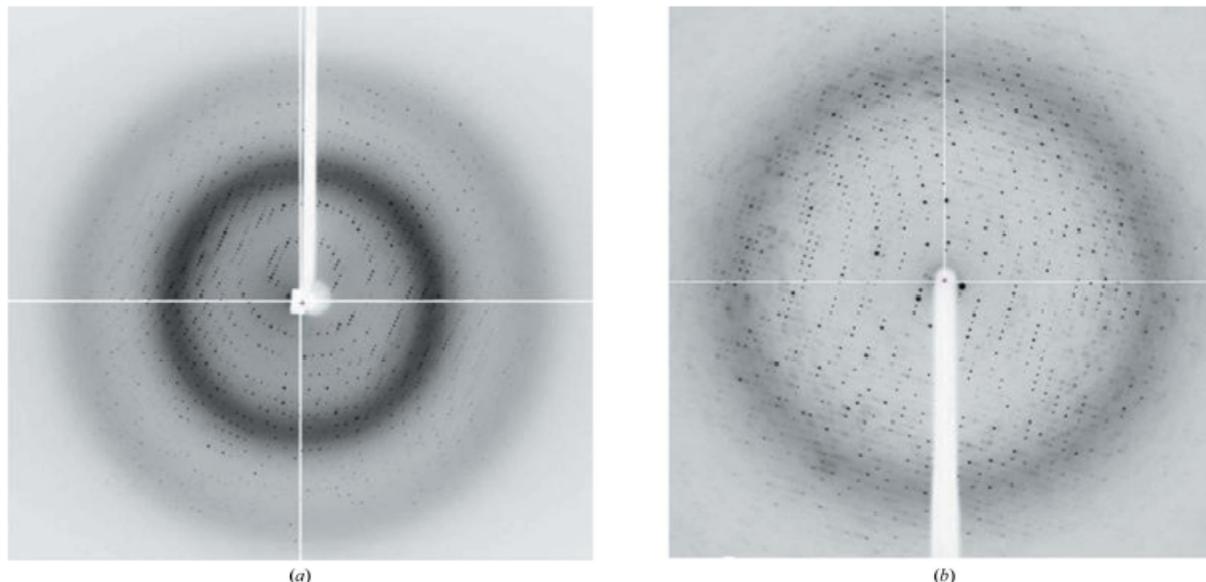


Figure 3

(a) Typical diffraction pattern for apo-GAPDHA. 360 images were collected with 0.5° oscillation at a fixed wavelength of 0.933 \AA . The crystal-to-detector distance was 151.3 mm . (b) Typical diffraction pattern of His₆-GAPDHA grown in the presence of NADP. The crystal-to-detector distance was 300.7 mm . 120 images were collected with 0.5° oscillation at a fixed wavelength of 0.934 \AA .

thermophilus; Skarzynski *et al.*, 1987) as a search model in *Phaser* (Read, 2001) gave a solution that was only consistent with space group $P3_2$. The sequence identity between the two enzymes is 46%.

3. Results and discussion

GAPDHA has been cloned, expressed and purified. Biochemical studies, which will be presented elsewhere, have shown that GAPDHA is NADP-dependent. Crystals of holo-GAPDHA were obtained when the hexahistidine tag was present and grew using 100 mM Tris pH 7.3, 36% methanol as the reservoir buffer. Crystals of apo-GAPDHA only grew when the hexahistidine affinity tag had been removed using TEV protease and belonged to a different space group ($P2_1$) to the holoenzyme ($P3_2$). Preliminary analysis of the electron density derived from the holo-His₆-GAPDHA crystals (which diffracted to 2.6 \AA) confirmed the presence of NADP bound to GAPDHA. Work is currently ongoing to obtain a suitable site-directed mutant that is able to form a stable ternary complex with NADP and glyceraldehyde-3-phosphate.

We would like to thank the beamline scientists at ID14-1 and ID14-2 of ESRF. PRE was supported by an MRC studentship.

References

- Alm, R. A. *et al.* (1999). *Nature (London)*, **397**, 176–180.
 Chalk, P. A., Roberts, A. D. & Blow, W. M. (1994). *Microbiology*, **140**, 2085–2092.
 Dunn, B. E., Cohen, H. & Blaser, M. J. (1997). *Clin. Microbiol. Rev.* **10**, 720–741.
 Eidt, S., Stolte, M. & Fischer, R. (1994). *J. Clin. Pathol.* **47**, 436–439.
 Ernst, P. B. (2000). *Annu. Rev. Microbiol.* **54**, 615–640.
 Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
 Kuipers, E. J. (1999). *Aliment. Pharmacol. Ther.* **13**, Suppl. 1, 3–11.
 Kuipers, E. J., Thijis, J. C. & Festen, H. P. (1995). *Aliment. Pharmacol. Ther.* **9**, Suppl. 2, 59–69.
 Leslie, A. G. W. (1992). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
 Marshall, B. J. & Warren, J. R. (1983). *Lancet*, **16**, 1311–1315.
 Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
 Mendz, G. L. & Hazell, S. L. (1993). *Arch. Biochem. Biophys.* **300**, 522–525.
 Mendz, G. L., Hazell, S. L. & Burns, B. P. (1993). *J. Gen. Microbiol.* **139**, 3023–3028.
 Mendz, G. L., Hazell, S. L. & Burns, B. P. (1994). *Arch. Biochem. Biophys.* **312**, 349–356.
 Mendz, G. L., Hazell, S. L. & Srinivasan, S. (1995). *Arch. Biochem. Biophys.* **321**, 153–159.
 Read, R. J. (2001). *Acta Cryst.* **D57**, 1373–1382.
 Skarzynski, T., Moody, P. C. E. & Wonacott, A. J. (1987). *J. Mol. Biol.* **193**, 171–187.
 Tomb, J.-F. *et al.* (1997). *Nature (London)*, **388**, 539–547.
 Wanken, A. E., Conway, T. & Eaton, K. A. (2003). *Infect. Immun.* **71**, 2920–2923.

Acta Crystallographica Section F
**Structural Biology
 and Crystallization
 Communications**

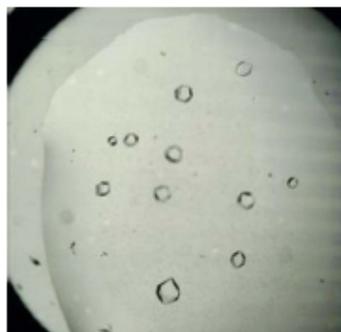
ISSN 1744-3091

**Paul R. Elliott, Shabaz
 Mohammad, Helen J. Melrose
 and Peter C. E. Moody***

Henry Wellcome Laboratories for Structural
 Biology, University of Leicester,
 Leicester LE1 9HN, England

Correspondence e-mail: pce1@leicester.ac.uk

Received 24 February 2008
 Accepted 2 July 2008



© 2008 International Union of Crystallography
 All rights reserved

Expression, purification, crystallization and preliminary X-ray analysis of an NAD-dependent glyceraldehyde-3-phosphate dehydrogenase from *Helicobacter pylori*

Helicobacter pylori is a dangerous human pathogen that resides in the upper gastrointestinal tract. Little is known about its metabolism and with the onset of antibiotic resistance new treatments are required. In this study, the expression, purification, crystallization and preliminary X-ray diffraction of an NAD-dependent glyceraldehyde-3-phosphate dehydrogenase from *H. pylori* are reported.

1. Introduction

Helicobacter pylori, a micro-aerophilic Gram-negative rod-shaped bacteria, colonizes the human upper gastrointestinal tract. *H. pylori* infection is more common in developing countries than in developed countries, with 70–90% of the population of developing countries carrying the bacteria, compared with 25–50% of the population of developed countries (Dunn *et al.*, 1997). Infection with *H. pylori* is implicated as being the cause of a variety of diseases including gastric and peptic ulcers, gastric cancer and mucosal-associated lymphoma (Kuipers *et al.*, 1995). Patients infected with *H. pylori* have a 20% lifetime risk of developing peptic ulcer disease and a 2% risk of developing gastric cancer (Ernst, 2000; Kuipers, 1999).

Treatment of *H. pylori* infection consists of a triple therapy comprising of treatment with two antibiotics, typically metronidazole and amoxicillin, coupled with proton-pump inhibitors (Cavallaro *et al.*, 2006). With high incidences of antibiotic resistance, quadruple therapies involving bismuth salts and three antibiotics are commonly being used (Gisbert & Pajares, 2002). These treatments have a variety of side effects (Perri *et al.*, 2001). A combination of these side effects and the prolonged and intense nature of treatment often results in patients not completing treatment, resulting in further incidences of antibiotic resistance and making eradication of the bacteria increasingly difficult (Egan *et al.*, 2007). Therefore, new alternative methods of treatment are required.

Analysis of the genome of *H. pylori* suggested that its only carbohydrate source was glucose (Tomb *et al.*, 1997). This was confirmed by the *in vitro* growth requirements of the bacteria (Mendz *et al.*, 1993) and the fact that the only carbohydrate transporter is glucose permease (Alm *et al.*, 1999). Genomic analysis also suggested that the main carbohydrate catabolic pathway, glycolysis is incomplete (Alm *et al.*, 1999; Tomb *et al.*, 1997). However, the early steps in glycolysis can be bypassed by the Entner–Doudoroff pathway (which is constitutively induced in *H. pylori*; Mendz *et al.*, 1994). The entry point of the Entner–Doudoroff pathway into glycolysis is through glyceraldehyde-3-phosphate, which is also a key substrate for the pentose-phosphate pathway (required for the synthesis of five-carbon sugars for nucleotide biosynthesis and for the generation of NADPH for oxidative-stress prevention). In glycolysis and gluconeogenesis, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) catalyses the reversible oxidative phosphorylation of glyceraldehyde-3-phosphate to 1,3-bisphosphoglycerate.

Two GAPDH genes have been identified in *H. pylori*: *gapA* (gi:6626253) and *gapB* (gi:2494645). The expression and crystallization of *gapA* is described in an accompanying paper (Elliott *et al.*,

crystallization communications

2008). Here, we report the cloning of *gapB*, its expression, and the purification and preliminary X-ray analysis of its gene product GAPDHB. Solution work to be published elsewhere confirms that GAPDHB is an NAD-dependent GAPDH that is able to utilize both glyceraldehyde-3-phosphate and erythrose-4-phosphate in a phosphate-dependent manner.

2. Materials and methods

2.1. Cloning and overexpression

The full nucleotide sequences encoding the *gapB* sequence (gi:2494645, NCBI, NIH) were cloned into TOPO pET151/D (Invitrogen) containing an N-terminal His₆ tag linked by a TEV protease site. The primer sequences for the forward and reverse amplification of *gapB* were CACCATGAAAATTTTATCATTGGATTG and TTAATAATGATACATAACTGG, respectively. Dideoxy sequencing confirmed the presence of the full-length *gapB* sequence. *Escherichia coli* strain Rosetta DE3 transformed with pET151/D-GAPDHB was grown at 303 K in the rich medium 2YT supple-

mented with 100 µg ml⁻¹ ampicillin and 35 µg ml⁻¹ chloramphenicol. Upon reaching an OD₆₀₀ of 0.6, the cultures were cooled to 291 K and isopropyl β-D-1-thiogalactopyranoside was added to a final concentration of 100 µM. Cells were left overnight at 291 K, harvested by centrifugation approximately 20 h after induction and immediately frozen.

2.2. Purification

Purification of His₆-tagged GAPDHB took place as follows. Frozen cell pellets were thawed in buffer A (20 mM Na₂HPO₄, 500 mM NaCl, 50 mM imidazole pH 7.4 supplemented with protease-inhibitor cocktail VII; Calbiochem). The suspension was sonicated and cell debris was removed by centrifugation. The supernatant was loaded onto a pre-equilibrated 5 ml Hi-Trap Nickel Sepharose (Amersham Biosciences) column and eluted with a linear gradient of buffer B (20 mM Na₂HPO₄, 500 mM NaCl, 500 mM imidazole pH 7.4). Fractions containing enzymatic activity were pooled and dialysed extensively against buffer C (20 mM Na₂HPO₄, 50 mM NaCl, 1 mM DTT pH 7.2). The sample was loaded onto a 5 ml Hi-Trap Sulfopropyl Sepharose cation-exchange column and eluted with a linear gradient of buffer D (20 mM Na₂HPO₄, 1 M NaCl, 1 mM DTT pH 7.2). A single peak was collected at 550 mM NaCl and was judged to be ~95% pure by SDS-PAGE analysis. GAPDHB was concentrated to 8 mg ml⁻¹ using an Amicon Ultra-15 centrifugal filter unit (10 kDa molecular-weight cutoff; Millipore) and the buffer was exchanged to 20 mM MES, 100 mM NaCl, 1 mM DTT pH 6.5 prior to crystallization.

2.3. Crystallization and data collection

Crystallization trials for the hexahistidine-tagged GAPDHB were performed in the presence of 1 mM NAD. A total of 192 crystal-growth conditions were screened by vapour diffusion using 100 nl drops of protein solution mixed with 100 nl precipitant from the

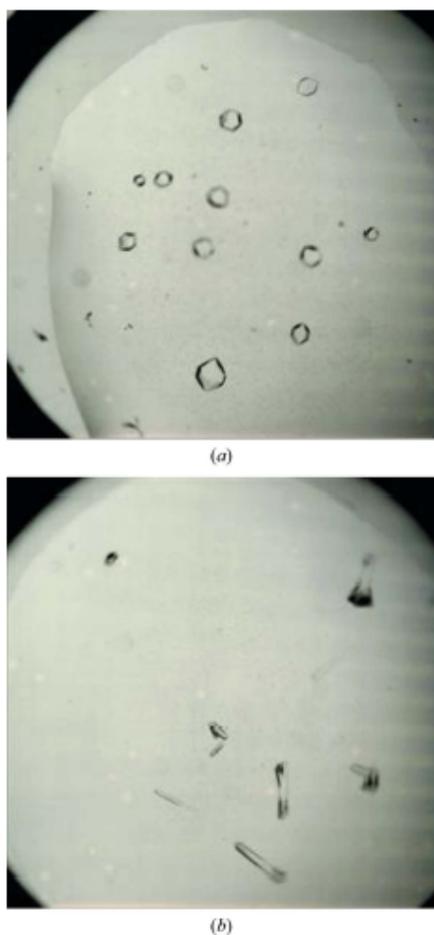


Figure 1
Two distinct crystal forms of GAPDHB were obtained after 48 h in 100 mM acetate pH 4.0, 38% 2-methyl-2,4-pentanediol. The form A crystals in (a) diffracted to 2.8 Å resolution, whilst the form B crystals in (b) did not diffract beyond 11 Å. 500 µl drops are shown.

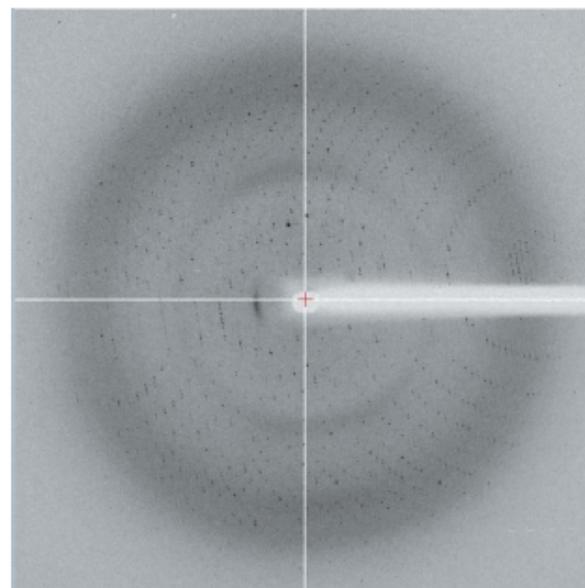


Figure 2
Typical diffraction pattern of crystal form A. Data were collected with a crystal-to-detector distance of 269.6 mm. Images of 1° oscillation were collected over 120° at a fixed wavelength of 0.931 Å.

Table 1
Summary of data-collection statistics for GAPDHB.

Values in parentheses are for the highest resolution shell.

Space group	<i>P6₃22</i>
Unit-cell parameters (Å, °)	<i>a</i> = <i>b</i> = 166.1, <i>c</i> = 253.1, <i>α</i> = <i>β</i> = 90.0, <i>γ</i> = 120.0
Resolution limits	83.1–2.8 (2.95–2.80)
No. of observations	393489 (57245)
No. of unique observations	50370 (7302)
Completeness	98.4 (99.6)
<i>I</i> (<i>hkl</i>)	17.0 (6.1)
<i>R</i> _{merge} †	0.094 (0.32)

† $R_{merge} = \frac{\sum_{obs} \sum_{i,j} |I_i(hkl) - I_j(hkl)|}{\sum_{obs} \sum_{i,j} I_i(hkl)}$, where $I_i(hkl)$ is the intensity of an individual measurement of the reflection with Miller indices hkl and $I_j(hkl)$ is the mean intensity of that reflection.

crystal screen kits Wizard I and II and Cryo I and II (Emerald Biosciences) dispensed with a Genomics Solutions Cartesian Honeybee 8+1 (Harvard Bioscience) onto 96-well MRC plates (Innovadyne) with reservoirs containing 80 µl of precipitant in a humidity chamber. Plates were sealed with transparent tape and monitored for crystal growth using CrystalProHT (TriTek) plate-storage and imaging systems at 277 and 293 K.

Several crystallization hits were recorded approximately 4 d after the plates were sealed. Crystallization conditions were predominately from the Cryo screens (Emerald Biosciences), with a preference for low pH and small-molecular-weight PEGs as a precipitant. Conditions yielding crystals were optimized using a Tecan 75 liquid-handling robot (Tecan) and 500 nl drops of protein and precipitant were dispensed from the Cartesian Honeybee 8+1 (Harvard Bioscience). Suitable crystals of diffraction quality grew from 8 mg ml⁻¹ GAPDHB containing 1 mM NAD mixed with an equal volume of reservoir containing 100 mM acetate pH 4.0 and 38% (v/v) 2-methyl-2,4-pentanediol. Crystals grew after 48 h at 277 K, with two distinct crystal forms present in the drops (Figs. 1*a* and 1*b*).

Crystals were prepared for cryocrystallography by harvesting the crystals directly into a stream of boiled-off liquid nitrogen at 110 K. The crystallization condition acted as a suitable cryoprotectant and no further rounds of optimization were required.

Diffraction data were collected on the ID14-3 beamline at the ESRF, Grenoble at 0.931 Å on an ADSC Q4R CCD detector; a typical diffraction pattern is shown in Fig. 2. Intensities were measured using *MOSFLM* (Leslie, 1992), with the autoindexing routines giving a solution consistent with a primitive hexagonal cell. The data were scaled using *SCALA* (Evans, 2006) and were consistent with the Laue group 622. Analysis of the distribution of intensities along the principal axes indicated the presence of either a 6₁ or a 6₅ screw axis. Molecular replacement using GAPDH from *Bacillus*

stearothermophilus (PDB code 1gd1; Skarzynski *et al.*, 1987) as a search model in *Phaser* (Read, 2001; the sequence identity between the two enzymes is 43%) gave a solution that was only consistent with space group *P6₃22* (data-collection statistics are given in Table 1).

3. Results and discussion

GAPDHB has been cloned, expressed and purified. Crystals of GAPDHB grew with the hexahistidine tag present and diffracted to 2.8 Å resolution. Preliminary electron-density analysis of GAPDHB confirmed that NAD was bound. Solution work, which will be presented elsewhere, demonstrated that GAPDHB catalyses the oxidative phosphorylation of glyceraldehyde-3-phosphate and erythrose-4-phosphate. Work is currently in progress to improve the diffraction quality of the crystals and to produce ternary complex analogues with glyceraldehyde-3-phosphate and with erythrose-4-phosphate in order to understand the apparent dual substrate-specificity.

We would like to thank the beamline scientists at ID14-3 of ESRF. PRE was supported by an MRC studentship.

References

- Alm, R. A. *et al.* (1999). *Nature (London)*, **397**, 176–180.
 Cavallaro, L. G., Egan, B., O'Morain, C. & Di Mario, F. (2006). *Helicobacter*, **11**, 36–39.
 Dunn, B. E., Cohen, H. & Blaser, M. J. (1997). *Clin. Microbiol. Rev.* **10**, 720–741.
 Egan, B. J., Katicic, M., O'Connor, H. J. & O'Morain, C. A. (2007). *Helicobacter*, **12**, Suppl. 1, 31–37.
 Elliott, P. R., Evans, D., Greenwood, J. A. & Moody, P. C. E. (2008). *Acta Cryst.* **F64**, 723–726.
 Ernst, P. B. (2000). *Annu. Rev. Microbiol.* **54**, 615–640.
 Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
 Gisbert, J. P. & Pajares, J. M. (2002). *Aliment. Pharmacol. Ther.* **16**, 1047–1057.
 Kuipers, E. J. (1999). *Aliment. Pharmacol. Ther.* **13**, Suppl. 1, 3–11.
 Kuipers, E. J., Thijs, J. C. & Festen, H. P. (1995). *Aliment. Pharmacol. Ther.* **9**, Suppl. 2, 59–69.
 Leslie, A. G. W. (1992). *Int. CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
 Mendz, G. L., Hazell, S. L. & Bums, B. P. (1993). *J. Gen. Microbiol.* **139**, 3023–3028.
 Mendz, G. L., Hazell, S. L. & Bums, B. P. (1994). *Arch. Biochem. Biophys.* **312**, 349–356.
 Perri, F., Festa, V., Clemente, R., Villani, M. R., Quitadamo, M., Caruso, N., Bergoli, M. & Andriulli, A. (2001). *Am. J. Gastroenterol.* **96**, 58–62.
 Read, R. J. (2001). *Acta Cryst.* **D57**, 1373–1382.
 Skarzynski, T., Moody, P. C. E. & Wonacott, A. J. (1987). *J. Mol. Biol.* **193**, 171–187.
 Tomb, J.-F. *et al.* (1997). *Nature (London)*, **388**, 539–547.

Bibliography

- ABAD-ZAPATERO, C., GRIFFITH, J. P., SUSSMAN, J. L. & ROSSMANN, M. G. (1987) Refined crystal structure of dogfish M4 apo-Lactate Dehydrogenase. *J Mol Biol.*, 198, 445-467.
- ALLEN, L.-A. H. (2001) The role of the neutrophil and phagocytosis in infection caused by *Helicobacter pylori*. *Curr Opin Infect Dis*, 14, 273-277.
- ALLOS, B. M. (2001) *Campylobacter jejuni* infections: update on emerging issues and trends. *Clin Infect Dis*, 33, 1201-1206.
- ALM, R. A., LING, L. S. L., MOIR, D. T., KING, B. L., BROWN, E. D., DOIG, P. C., SMITH, D. R., NOONAN, B., GUILD, B. C., DEJONGE, B. L., CARMEL, G., TUMMINO, P. J., CARUSO, A., NICKELSEN, M., MILLS, D. M., IVES, C., GIBSON, R., MERBERG, D., MILLS, S. D., JIANG, Q., TAYLOR, D. E., VOVIS, G. F. & TRUST, T. J. (1999) Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature*, 397, 176-180.
- ANDERSSON, J. O. & ANDERSSON, S. G. (1999) Insights into the evolutionary process of genome degradation. *Curr. Opin. Genet. Dev.*, 9, 664-671.
- ARMSTRONG, J. A., WEE, S. H., GOODWIN, C. S. & WILSON, D. H. (1987) Response of *Campylobacter pyloridis* to antibiotics, bismuth and an acid reducing agent in vitro-an ultrastructural study. *J. Med. Microbiol.*, 24, 343-350.
- ASHDOWN, L. R. (1978) Stuart's transport medium. *Med J Aust*, 25, 224.
- BAAR, C., EPPINGER, M., RADDATZ, G., SIMON, J., LANZ, C., KLIMMEK, O., NADAKUMMAR, R., GROSS, R., ROSSINUS, A., KELLER, H., JAGTAP, P., LINKE, B., MEYER, F., LEDERER, H. & SCHUSTER, S. C. (2003a) The complete genome sequence and analysis of *Wolinella succinogenes*. *Proc Nat Acad Sci*, 100, 11690-11695.
- BAAR, C., EPPINGER, M., RADDATZ, G., SIMON, J., LANZ, C., KLIMMEK, O., NADAKUMMAR, R., GROSS, R., ROSSINUS, A., KELLER, H., JAGTAP, P., LINKE, B., MEYER, F., LEDERER, H. & SCHUSTER, S. C. (2003b) Complete genome sequence and analysis of *Wolinella succinogenes*. *PNAS*, 100, 11690-11695.
- BAILEY, S., FAIRLAMB, A. H. & HUNTER, W. N. (1994) Structure of trypanothione reductase from *Crithidia fasciculata* at 2.6 Å resolution; enzyme-NADP interactions at 2.8 Å resolution. *Acta Crystallogr D Biol Crystallogr*, 1, 139-154.
- BARR, C., EPPINGER, M., RADDATZ, G., SIMON, J., LANZ, C., KLIMMEK, O., NADAKUMMAR, R., GROSS, R., ROSSINUS, A., KELLER, H., JAGTAP, P., LINKE, B., MEYER, F., LEDERER, H. & SCHUSTER, S. C. (2003) The complete genome sequence and analysis of *Wolinella succinogenes*. *Proc Nat Acad Sci*, 100, 11690-11695.
- BASSET, C., HOLTON, J., GATTA, L., RICCI, C., BERNABUCCI, V., LIUZZI, G. & VAIRA, D. (2004) *Helicobacter pylori* infection: anything new should we know? *Alimentary Pharmacology & Therapeutics*, 20, S31-41.
- BAUERFEIND, P., GARNER, R., DUNN, B. E. & MOBLEY, H. L. (1997) Synthesis and activity of *Helicobacter pylori* urease and catalase at low pH. *Gut*, 40, 25-30.
- BAYERDORFFER, E., NEUBAUER, A., RUDOLPH, B., THIEDE, C., LEHN, N., EIDT, S. & STOLTE, M. (1995) Regression of primary gastric lymphoma of mucosa-associated lymphoid tissue type after cure of *Helicobacter pylori* infection. MALT Lymphoma Study Group. *Lancet*, 345, 1591-1594.

- BELL, G. D., POWELL, K., BURRIDGE, S. M., PALLECAROS, A. & JONES, P. H. (1992) Experience with "triple" anti-Helicobacter eradication therapy: side effects and the importance of testing the pre-treatment bacterial isolate for metronidazole resistance. *Alimentary Pharmacology & Therapeutics*, 6, 427-435.
- BIESECKER, G. & WONACOTT, A. (1977) Coenzyme binding and co-operativity in D-glyceraldehyde 3-phosphate dehydrogenase. *Biochem Soc Trans*, 5, 647-652.
- BLANCO, J., MOORE, R. A., KABALEESWARAN, V. & VIOLA, R. E. (2003) A structural Basis for the Mechanism of Aspartate-beta-semialdehyde Dehydrogenase from *Vibrio Cholerae*. *Protein Sci.*, 12, 27-33.
- BLASER, M. J. & ATHERTON, J. C. (2004) Helicobacter pylori persistence: biology and disease. *Journal of Clinical Investigation*, 113, 321-333.
- BLISS, C. M., GOLENBOCK, D. T., KEATES, S., LINEVSKY, J. K. & KELLY, C. P. (1998) Helicobacter pylori lipopolysaccharide binds to CD14 and stimulates release of interleukin-8, epithelial neutrophil-activating peptide 78, and monocyte chemotactic protein 1 by human monocytes. *Infect. Immun.*, 66, 5357-5363.
- BLOM, K., LUNDIN, B. S., BOLIN, I. & SVENNERHOLM, A. (2001) Flow cytometric analysis of the localization of Helicobacter pylori antigens during different growth phases. *FEMS Immunology and Medical Microbiology*, 30, 173-179.
- BLUNDELL, T. L. & JOHNSON, L. N. (1976) *Protein Crystallography*, Academic Press.
- BORK, P. & GRUNWALD, C. (1990) Recognition of different nucleotide-binding sites in primary structures using a property-pattern approach. *Eur J Biochem*, 191, 347-358.
- BOSCHI-MULLER, S., AZZA, S., POLLASTRO, D., CORBIER, C. & BRANLANT, G. (1997) Comparative Enzymatic Properties of GapB-encoded Erythrose-4-Phosphate Dehydrogenase of *Escherichia coli* and Phosphorylating Glyceraldehyde-3-Phosphate Dehydrogenase. *The Journal of Biological Chemistry*, 272, 15106-15112.
- BRICOGNE, G. (1997) *Bayesian statistical viewpoint on structure determination: basic concepts and examples*, Academic Press.
- BROWN, L. M. (2000) Helicobacter Pylori: Epidemiology and Routes of Transmission. *Epidemiol Rev*, 22, 283-297.
- BRUNNER, N., SIEBERS, B. & HENSEL, R. (2001) Role of two different glyceraldehyde-3-phosphate dehydrogenases in controlling the reversible Embden-Meyerhof-Parnas pathway in *Thermoproteus tenax*: regulation on protein and transcript level. *Extremophiles*, 5, 101.
- BUEHNER, M., FORD, G. C., OLSEN, K. W., MORAS, D. & ROSSMANN, M. G. (1974) Three-dimensional structure of D-glyceraldehyde-3-phosphate dehydrogenase. *J Mol Biol.*, 90, 25-49.
- CAMPBELL, B. J., ENGEL, A. S., PORTER, M. L. & TAKAI, K. (2006) The versatile [epsi]-proteobacteria: key players in sulphidic habitats. *Nat Rev Micro*, 4, 458.
- CARTER, C. W. J. & CARTER, C. W. (1979) Protein crystallization using incomplete factorial experiments. *J. Biol. Chem.*, 254, 12219-12223.
- CARUGO, O. & ARGOS, P. (1997) NADP-dependent enzymes. I: Conserved stereochemistry of cofactor binding. *Proteins*, 28, 10-28.
- CASTILHO, M. S., PAVAO, F. & OLIVA, G. (2003) Evidence for the Two Phosphate Binding Sites of an Analogue of the Thioacyl Intermediate for the *Trypanosoma cruzi* Glyceraldehyde-3-phosphate Dehydrogenase-Catalyzed Reaction, from Its Crystal Structure. *Biochemistry*, 42, 7143-7151.
- CAVALLARO, L. G., EGAN, B., O'MORAIN, C. & DI MARIO, F. (2006) Treatment of Helicobacter pylori Infection. *Helicobacter*, 11, 36-39.
- CENSINI, S., LANGE, C., XIANG, Z., CRABTREE, J. E., GHIARA, P., BORODOVSKY, M., RAPPUOLI, R. & COVACCI, A. (1996) cag, a pathogenicity island of Helicobacter

- pylori, encodes type I-specific and disease-associated virulence factors. *PNAS*, 93, 14648-14653.
- CERFF, R. & CHAMBERS, S. (1979) Subunit structure of higher plant glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.*, 254, 6094-6098.
- CHALK, P. A., ROBERTS, A. D. & BLOWS, W. M. (1994) Metabolism of pyruvate and glucose by intact cells of *Helicobacter pylori* studied by ¹³C NMR spectroscopy. *Microbiology*, 140, 2085-2092.
- CHANG, H.-T., MARCELLI, S. W., DAVISON, A. A., CHALK, P. A., POOLE, R. K. & MILES, R. J. (1995) Kinetics of substrate oxidation by whole cells and cell membranes of *Helicobacter pylori*. *FEMS Microbiology*, 129, 33-38.
- CHARRON, C., TALFOURNIER, F., ISUPOV, M. N., LITTLECHILD, J. A., BRANLANT, G., VITOUX, B. & AUBRY, A. (2000) The crystal structure of -glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic archaeon *Methanothermus fervidus* in the presence of NADP⁺ at 2.1 Å resolution. *Journal of Molecular Biology*, 297, 481.
- CHEN, M., ANDERSEN, L. P., ZHAI, L. & KHARAZMI, A. (1999) Characterization of the respiratory chain of *Helicobacter pylori*. *FEMS Immunology and Medical Microbiology*, 24, 169-174.
- CHRISTIE, P. J. & VOGEL, J. P. (2000) Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. *Trends in Microbiology*, 8, 354-360.
- COBESSI, D., TETE-FAVIER, F., MARCHAL, S., AZZA, S., BRANLANT, G. & AUBRY, A. (1999) Apo and holo crystal structures of an NADP-dependent aldehyde dehydrogenase from *Streptococcus mutans*. *J Mol Biol.*, 290, 161-173.
- COBESSI, D., TETE-FAVIER, F., MARCHAL, S., BRANLANT, G. & AUBRY, A. (2000) Structural and biochemical investigations of the catalytic mechanism of an NADP-dependent aldehyde dehydrogenase from *Streptococcus mutans*. *J Mol Biol.*, 300, 141-152.
- COPELAND, A., LUCAS, S., LAPIDUS, A., BARRY, K., DETTER, J. C., GLAVINA, T., HAMMON, N., ISRANI, S., PITLUCK, S., CHAIN, P., MALFATTI, S., SHIN, M., VERGEZ, L., SCHMUTZ, J., LARIMER, F., LAND, M., KYRPIDES, N., LYKIDIS, A. & RICHARDSON, P. (2006) Complete sequence of *Thiomicrospira denitrificans* ATCC 33889. *NCBI database (unpublished)*.
- CORRE, E., REYSENBACH, A.-L. & PRIEUR, D. (2001) [epsilon]-Proteobacterial diversity from a deep-sea hydrothermal vent on the Mid-Atlantic Ridge. *FEMS Microbiology Letters*, 205, 329.
- CORTHESEY-THEULAZ, I. E., BERGONZELLI, G. E., HENRY, H., BACHMANN, D., SCHORDERET, D. F., BLUM, A. L. & ORNSTON, L. N. (1997) Cloning and characterisation of *Helicobacter pylori* succinyl CoA:acetoacetate CoA-transferase, a novel prokaryotic member of the CoA-transferase family. *J. Biol. Chem.*, 272, 25659-25667.
- COVACCI, A., TELFORD, J. L., DEL GIUDICE, G., PARSONNET, J. & RAPPUOLI, R. (1999) *Helicobacter pylori* virulence and genetic geography. *Science*, 284, 1328-1333.
- COVER, T. L. & BLASER, M. J. (1996) *Helicobacter pylori* infection, a paradigm for chronic mucosal inflammation: pathogenesis and implications for eradication and prevention. *Adv Intern Med*, 41, 85-117.
- CUTLER, A. F., HAVSTAD, S., MA, C. K., BLASER, M. J., PEREZ-PEREZ, G. I. & SCHUBERT, T. T. (1995) Accuracy of invasive and noninvasive tests to diagnose *Helicobacter pylori* infection. *Gastroenterology*, 109, 136-141.
- D'AMBROSIO, K., PAILOT, A., TALFOURNIER, F., DIDIERJEAN, C., BENEDETTI, E., AUBRY, A., BRANLANT, G. & CORBIER, C. (2006) The first crystal structure of a

- thioacylenzyme intermediate in the ALDH family: new coenzyme conformation and relevance to catalysis. *Biochemistry*, 45, 2978-2986.
- DAVIES, I. W., LEAVER-FAY, A., CHEN, V. B., BLOCK, J. N., KARPAL, G. J., WANG, X., MURRAY, L. W., ARENDALL, W. B., SNOEYINK, J., RICHARDSON, J. S. & RICHARDSON, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research*, 35, w375-w383.
- DIDIERJEAN, C., CORBIER, C., FAITH, M., FAVIER, F., BOSCHI-MULLER, S., BRANLANT, G. & AUBRY, A. (2003) Crystal Structure of Two Ternary Complexes of Phosphorylating Glyceraldehyde-3-phosphate Dehydrogenase from *Bacillus stearothermophilus* with NAD and D-Glyceraldehyde 3-Phosphate. *The Journal of Biological Chemistry*, 278, 12968-12976.
- DIDIERJEAN, C., RAHUEL-CLERMONT, S., VITOUX, B., DIDEBERG, O., BRANLANT, G. & AUBRY, A. (1997) A Crystallographic Comparison Between Mutated Glyceraldehyde-3-phosphate Dehydrogenases from *Bacillus stearothermophilus* Complexed with Either NAD or NADP. *Journal of Molecular Biology*, 268, 739-759.
- DIXON, M. F., GENTA, R. M., YARDLEY, J. H. & CORREA, P. (1994) Classification and grading of gastritis. The updated Sydney System. International Workshop on the Histopathology of Gastritis, Houston 1994. *American Journal of Surgical Pathology*, 20, 1161-1181.
- DRUMM, B., PEREZ-PEREZ, G. I., BLASER, M. J. & SHERMAN, P. M. (1990) Intrafamilial clustering of *Helicobacter pylori* infection. *N. Engl. J. Med.*, 322, 359-363.
- DUEE, E., OLIVIER-DEYRIS, L., FANCHON, E., CORBIER, C., BRANLANT, G. & DIDEBERG, O. (1996) Comparison of the Structures of Wild-type and a N313T Mutant of *Escherichia coli* Glyceraldehyde 3-Phosphate Dehydrogenases: Implication for NAD Binding and Cooperativity. *Journal of Molecular Biology*, 257, 814-838.
- DUNDON, W. G., NISHIOKA, H., POLENGHI, A., PAPINUTTO, E., ZANOTTI, G., MONTEMURRO, P., DEL GIUDICE, G., RAPPUOLI, R. & MONTECUCCO, C. (2001) The neutrophil-activating protein of *Helicobacter pylori*. *International Journal of Medical Microbiology*, 291, 545.
- DUNN, B. E., COHEN, H. & BLASER, M. J. (1997) *Helicobacter pylori*. *Clin Microbiol Rev*, 10, 720-741.
- EATON, K. A., BROOKS, C. L., MORGAN, D. R. & KRAKOWKA, S. (1991) Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. *Infect. Immun.*, 59, 2470-2475.
- EATON, K. A., MORGAN, D. R. & KRAKOWKA, S. (1992) Motility as a factor in the colonisation of gnotobiotic piglets by *Helicobacter pylori*. *J Med Microbiol*, 37, 123-127.
- EHRENFELD, M., JECK, R., KLATTE, W., KUHN, N. & WOENCKHAUS, C. (1981) Half-of-the-sites reactivity of glyceraldehyde-3 phosphate dehydrogenase from rabbit muscle with structural analogs of NAD. *Z. Naturforsch*, 36, 545-551.
- EMSLEY, P. & COWTAN, K. (2004) Coot: model-building tools for molecular graphics. *Acta Cryst.*, D60, 2126-2132.
- ENGEL, A. S., LEE, N., PORTER, M. L., STERN, L. A., BENNETT, P. C. & WAGNER, M. (2003) Filamentous "Epsilonproteobacteria" Dominate Microbial Mats from Sulfidic Cave Springs. *Appl. Environ. Microbiol.*, 69, 5503-5511.
- EPPINGER, M., BAAR, C., RADDATZ, G., HUSON, D. H. & SCHUSTER, S. C. (2004) Comparative analysis of four Campylobacteriales. *Nat Rev Micro*, 11, 872-885.
- ERNST, P. B. (2000) THE DISEASE SPECTRUM OF HELICOBACTER PYLORI: The Immunopathogenesis of Gastrointestinal Ulcer and Gastric Cancer. *Annual Review Microbiology*, 54, 615-640.

- ETTEMA, T., AHMED, H., GEERLING, A., VAN DER OOST, J. & SIEBERS, B. (2008) The non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN) of *Sulfolobus solfataricus*: a key-enzyme of the semi-phosphorylative branch of the Entner–Doudoroff pathway. *Extremophiles*, 12, 75.
- EVANS, D. J. J., EVANS, D. J., TAKEMURA, T., NAKANO, H., LAMPERT, H. C., GRAHAM, D. Y., GRANGER, D. N. & KVIETYS, P. R. (1995) Characterization of a *Helicobacter pylori* neutrophil-activating protein. *Infect. Immun.*, 63, 2213-2220.
- EVANS, P. & MCCOY, A. (2008) An introduction to molecular replacement. *Acta Crystallogr D Biol Crystallogr*, 64, 1-10.
- EVANS, P. R. (2006) Scaling and assessment of data quality. *Acta Cryst.*, D62, 72-82.
- FAEHNLE, C. R., LE COQ, J., LIU, X. & VIOLA, R. E. (2006) Examination of key intermediates in the catalytic cycle of aspartate-beta-semialdehyde dehydrogenase from a gram-positive infectious bacteria. *J. Biol. Chem.*, 281, 31031-31040.
- FAKHERI, H., MERAT, S., HOSSEINI, V. & MALEKZADEH, R. (2004) Low-dose furazolidone in triple and quadruple regimens for *Helicobacter pylori* eradication. *Alimentary Pharmacology & Therapeutics*, 19, 89-93.
- FALINI, G., FERMANI, S., RIPAMONTI, A., SABATINO, P., SPARLA, F., PUPILLO, P. & TROST, P. (2003) Dual Coenzyme Specificity of Photosynthetic Glyceraldehyde-3-phosphate Dehydrogenase Interpreted by the Crystal Structure of A4 Isoform Complexed with NAD. *Biochemistry*, 42, 4631-4639.
- FERMANI, S., RIPAMONTI, A., SABATINO, P., ZANOTTI, G., SCAGLIRINI, S., SPARLA, F., TROST, P. & PUPILLO, P. (2001) Crystal Structure of the Non-regulatory A4 Isoform of Spinach Chloroplast Glyceraldehyde-3-phosphate Dehydrogenase Complexed with NADP. *Journal of Molecular Biology*, 314, 527-542.
- FERMANI, S., SPARLA, F., FALINI, G., MARTELLI, P. L., CASADIO, R., PUPILLO, P., RIPAMONTI, A. & TROST, P. (2007) Molecular mechanism of thioredoxin regulation in photosynthetic A2B2-glyceraldehyde-3-phosphate dehydrogenase. *PNAS*, 104, 11109-11114.
- FERREIRA-DA-SILVA, F., PERELRA, P. J. B., GALES, L., ROESSLE, M., SVERGUN, D. I., MORADAS-FERREIRA, P. & DAMAS, A. M. (2006) The Crystal and Solution Structures of Glyceraldehyde-3-phosphate Dehydrogenase Reveal Different Quaternary Structures. *The Journal of Biological Chemistry*, 281, 33433-33440.
- FERRI, G., STOPPINI, M., MELONI, M., ZAPPONI, M. C. & IADAROLA, P. (1990) Chloroplast glyceraldehyde-3-phosphate dehydrogenase (NADP): amino acid sequence of the subunits from isoenzyme I and structural relationship with isoenzyme II. *Biochim. Biophys. Acta*, 1041, 36-42.
- FIEDOREK, S. C., MALATY, H. M., EVANS, D. L., PLUMPHERY, C. L., CASTEEL, H. B., EVANS, D. J. & GRAHAM, D. Y. (1991) Factors influencing the epidemiology of *Helicobacter pylori* infection in children. *Pediatrics*, 88, 578-582.
- FILLINGER, S., BOSHI-MULLER, S., AZZA, S., DERVYN, E., BRANLANT, G. & AYMERICH, S. (2000) Two Glyceraldehyde-3-Phosphate Dehydrogenases with opposite physiological roles in a nonphotosynthetic bacterium. *The Journal of Biological Chemistry*, 275, 14031-14037.
- FINEL, M. (1998) Does NADH play a central role in energy metabolism in *Helicobacter pylori*? *TIBS*, 23, 412-414.
- FISCHER, W., PULS, J., BUHRDORF, R., GEBERT, B., ODENBREIT, S. & HAAS, R. (2001) Systematic mutagenesis of the *Helicobacter pylori* cag pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8. *Mol. Microbiol.*, 42, 1337-1348.

- FONTANA, C., FAVARO, M., MINELLI, S., CRISCUOLO, A. A., PIETROIUSTU, A., GALANTE, A. & FAVALLI, C. (2002) New site of modification of 23S rRNA associated with clarithromycin resistance of *Helicobacter pylori* clinical isolates. *Antimicrob. Agents. Chemother.*, 46, 3765-3769.
- FORMAN, D., SITAS, F., NEWELL, D. G., STACEY, A. R., BOREHAM, J., PETO, R., CAMPBELL, T. C., LI, J. & CHEN, J. (1990) Geographical association of *Helicobacter pylori* antibody prevalence and gastric cancer mortality in rural China. *Int. J. Cancer*, 46, 608-611.
- FOUTS, D. E., MONGODIN, E. F., MANDRELL, R. E., MILLER, W. G., RASKO, D. A., RAVEL, J., BRINKAC, L. M., DEBOY, R. T., PARKER, C. T., DAUGHERTY, S. C., DODSON, R. J., DURKIN, A. S., MADUPU, R., SULLIVAN, S. A., SHETTY, J. U., AYODEJI, M. A., SHVARTSBEYN, A., SCHATZ, M. C., BADGER, J. H., FRASER, C. M. & NELSON, K. E. (2005) Major structural differences and novel potential virulence mechanisms from the genomes of multiple campylobacter species. *PLoS Biol*, 3, e15.
- FRENCH, G. S. & WILSON, K. S. (1978) *Acta Cryst A*, A34, 517.
- GARMAN, E. F. & SCHNEIDER, T. R. (1997) Macromolecular Cryocrystallography. *J. Appl. Cryst.*, 30, 211-237.
- GEIS, G., LEYING, H., SUERBAUM, S., MAI, U. & OPFERKUCH, W. (1989) Ultrastructure and chemical analysis of *Campylobacter pylori* flagella. *J. Clin. Microbiol.*, 27, 436-441.
- GENNIS, L. S. (1976) Negative homotropic cooperativity and affinity heterogeneity: preparation of yeast glyceraldehyde-3-phosphate dehydrogenase with maximal affinity homogeneity. *PNAS*, 73, 3928-3932.
- GIUDICE, G. D., COVACCI, A., TELFORD, J. L., MONTECUCCO, C. & RAPPUOLI, R. (2001) THE DESIGN OF VACCINES AGAINST HELICOBACTER PYLORI AND THEIR DEVELOPMENT. *Annual Review Immunology*, 19, 523-563.
- GUPTA, R. S. (2006) Molecular signatures (unique proteins and conserved indels) that are specific for the epsilon proteobacteria (campylobacteriales). *BMC Genomics*, 167, 1-17.
- HA, N. C., OH, S. T., SUNG, J. Y., CHA, K. A., LEE, M. H. & OH, B. H. (2001) Supramolecular assembly and acid resistance of *Helicobacter pylori* urease. *Nat. Struct. Biol.*, 8, 505-509.
- HABENICHT, A., HELLMAN, U. & CERFF, R. (1994) Non-phosphorylating GAPDH of Higher Plants is a Member of the Aldehyde Dehydrogenase Superfamily with no Sequence Homology to Phosphorylating GAPDH. *Journal of Molecular Biology*, 237, 165.
- HADFIELD, A. T., KRYGER, G., OUYANG, J., RINGE, D., PETSKO, G. A. & VIOLA, R. E. (2001) Active site analysis of the potential antimicrobial target aspartate semialdehyde dehydrogenase. *Biochemistry*, 40, 14475-14483.
- HAZELL, S. L., EVANS, D. J. J. & GRAHAM, D. Y. (1991) *Helicobacter pylori* catalase. *J. Gen. Microbiol*, 137, 57-61.
- HILVERS, A. G., DAM, V. & SLATER, E. C. (1964) THE REACTION MECHANISM OF D-GLYCERALDEHYDE-3-PHOSPHATE:NAD⁺ OXIDOREDUCTASE (PHOSPHORYLATING) OF RABBIT MUSCLE. *Biochim Biophys Acta*, 85, 206-227.
- HUGHES, N. J., CLAYTON, C. L., CHALK, P. A. & KELLY, D. J. (1998) The porCDAB and oorDACB genes encode distinct pyruvate:flavodoxin and 2-oxoglutarate:acceptor oxidoreductases which mediate electron transport to NADP. *J. Bacteriol.*, 180, 2694-2697.
- INTERNATIONAL AGENCY FOR RESEARCH ON CANCER (1994) IARC monographs on the evaluation of carcinogenic risks to humans, vol 61; schistosomes, liver flukes and *Helicobacter pylori*. *International Agency for Research on Cancer, Lyon, France*.

- ISHII, Y., HASHIMOTO, T., MINAKAMI, S. & YOSHIKAWA, H. (1964) The formation of erythronic acid 4-phosphate from erythrose 4-phosphate by glyceraldehyde 3-phosphate. *The journal of Biochemistry*, 56, 111-112.
- ISRAEL, D. A., SALAMA, N., KRISHNA, U., RIEGER, U. M., ATHERTON, J. C., FALKOW, S. & PEEK, R. M. (2001) *Helicobacter pylori* genetic diversity within the gastric niche of a single human host. *PNAS*, 98, 14625-14630.
- ISUPOV, M. N., FLEMING, T. M., DALBY, A. R., CROWHURST, G. S., BOURNE, P. C. & LITTLECHILD, J. A. (1999) Crystal Structure of the Glyceraldehyde-3-phosphate Dehydrogenase from the Hyperthermophilic Archaeon *Sulfolobus solfataricus*. *Journal of Molecular Biology*, 291, 651-660.
- JANCARIK, J. & KIM, S.-H. (1991) Sparse-matrix sampling—a screening method for crystallisation of proteins. *Journal of Applied Crystallography*, 24, 409-411.
- KABSCH, A. (1976) A solution for the best rotation to relate two sets of vectors. *Acta Cryst.*, A32, 922-923.
- KABSCH, W. (1988) Automatic indexing of rotation diffraction patterns. *J. Appl. Cryst.*, 21, 67-72.
- KATHER, B., STINGL, K., VAN DER REST, M. E., ALTENDORF, K. & MOLENAAR, D. (2000) Another unusual type of citric acid cycle enzyme in *Helicobacter pylori*: the malate:quinone oxidoreductase. *Journal of Bacteriology*, 182, 3204-3209.
- KIESSLICH, R., GOETZ, M., BURG, J., STOLTE, M., SIEGEL, E., MAEURER, M. J., THOMAS, S., STRAND, D., GALLE, P. R. & NEURATH, M. F. (2005a) Diagnosing *Helicobacter pylori* In Vivo by Confocal Laser Endoscopy. *Gastroenterology*, 128, 2119-2123.
- KIESSLICH, R., GOETZ, M., VIETH, M., GALLE, P. R. & NEURATH, M. F. (2005b) Confocal Laser Endomicroscopy. *Gastrointestinal Endoscopy Clinics of North America*, 15, 715-731.
- KJØLLER, M., FISCHER, A. & JUSTESEN, T. (1991) Transport conditions and number of biopsies necessary for culture of *Helicobacter pylori*. *Eur J Clin Microbiol Infect Dis*, 10, 166-167.
- KODAMA, Y. & WATANABE, K. (2004) *Sulfuricurvum kujiense* gen. nov., sp. nov., a facultatively anaerobic, chemolithoautotrophic, sulfur-oxidizing bacterium isolated from an underground crude-oil storage cavity. *Int J Syst Evol Microbiol*, 54, 2297-2300.
- KONNO, M., FUJII, N., YOKOTA, S., SATO, K., TAKAHASHI, M., MILNO, E. & SUGIYAMA, T. (2005) Five-year follow-up study of mother-to-child transmission of *Helicobacter pylori* infection detected by a random amplified polymorphic DNA fingerprinting method. *J. Clin. Microbiol.*, 43, 2246-2250.
- KORNDORFER, I., STEIPE, B. S., HUBER, R., TOMSCHY, A. & JAENICKE, R. (1995) The Crystal Structure of Holo-glyceraldehyde-3-phosphate Dehydrogenase from the Hyperthermophilic Bacterium *Thermotoga maritima* at 2.5 Å Resolution. *Journal of Molecular Biology*, 246, 511.
- KRIMSKY, I. & RACKER, E. (1955) Acyl derivatives of glyceraldehyde-3-phosphate dehydrogenase. *Science*, 122, 319-321.
- KUIPERS, E. J. (1999) Review article: exploring the link between *Helicobacter pylori* and gastric cancer. *Alimentary Pharmacology and Therapeutics*, 13, 3-11.
- KUSTERS, J. G., GERRITS, M. M., VAN STRIJP, J. A. & VANDENBROUCKE-GRAULS, C. M. (1997) Coccoid forms of *Helicobacter pylori* are the morphologic manifestation of cell death. *Infect. Immun.*, 65, 3672-3679.
- KUSTERS, J. G., VILET, A. H. M. & KUIPERS, E. J. (2006) Pathogenesis of *Helicobacter pylori* Infection. *Clinical Microbiology Reviews*, 19, 449-490.

- LABIGNE, A. & DE REUSE, H. (1996) Determinants of *Helicobacter pylori* pathogenicity. *Infect Agents Dis*, 4, 191-202.
- LASKOWSKI, A., MACARTHUR, M. W., MOSS, D. S. & THORNTON, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, 26, 283-291.
- LESLIE, A. G. & WONACOTT, A. J. (1984) Structural evidence for ligand-induced sequential conformational changes in glyceraldehyde 3-phosphate dehydrogenase. *J. Mol. Biol.*, 178, 743-772.
- LESLIE, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Jnt CCP4 + ESF-EAMCB Newsl. on Protein Crystallography*, 26.
- LEUNK, R. D., JOHNSON, P. T., DAVID, B. C., KRAFT, W. G. & MORGAN, D. R. (1988) Cytotoxic activity in broth-culture filtrates of *Campylobacter pylori*. *J. Med. Microbiol.*, 26, 93-99.
- LEVITZKI, A. (1974) Half-of-the-sites and all-of-the-sites reactivity in rabbit muscle glyceraldehyde 3-phosphate dehydrogenase. *J Mol Biol.*, 90, 451-468.
- LISBY, G. (1999) Application of nucleic acid amplification to clinical microbiology. *Mol. Biotechnol.*, 12, 75-99.
- LUPETTI, P., HEUSER, J. E., MANETTI, R., MASSARI, P., LANZAVECCHIA, S., BELLON, P. L., DALLAI, R., RAPPUOLI, R. & TELFORD, J. L. (1996) Oligomeric and subunit structure of the *Helicobacter pylori* vacuolating cytotoxin. *J. Cell Biol.*, 133, 801-807.
- MAGALHAES QUEIROZ, D. M. & LUZZA, F. (2006) Epidemiology of *Helicobacter pylori* Infection. *Helicobacter*, 11, 1-5.
- MAIER, R. J., FU, C., GUILBERT, J., MOSHIRI, F., OLSON, J. & PLAUT, A. G. (1996) Hydrogen uptake hydrogenase in *Helicobacter pylori*. *FEMS Microbiol. Lett.*, 141, 71-76.
- MARAI, A., MENDZ, G. L., HAZELL, S. L. & MEGRAUD, F. (1999) Metabolism and Genetics of *Helicobacter pylori*: the Genome Era. *Microbiology and Molecular Biology Reviews*, 63, 642-674.
- MARCUS, E. A. & SCOTT, D. R. (2001) Cell lysis is responsible for the appearance of extracellular urease in *Helicobacter pylori*. *Helicobacter*, 6, 93-99.
- MARSHALL, B. J., ARMSTRONG, J. A., MCGECHIE, D. B. & GLANCY, R. J. (1985) Attempt to fulfil Koch's postulates for pyloric *Campylobacter*. *Med J Aust*, 142, 436-439.
- MARSHALL, B. J. & WARREN, J. R. (1983) Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet*, 16, 1311-1315.
- MAURICE, M., CREMADES, N., CROXEN, M. A., SISSON, G., SANCHO, J. & HOFFMAN, P. S. (2007) Flavodoxin:Quinone Reductase (FqrB): a Redox Partner of Pyruvate:Ferredoxin Oxidoreductase That Reversibly Couples Pyruvate Oxidation to NADPH Production in *Helicobacter pylori* and *Campylobacter jejuni*. *Journal of Bacteriology*, 189, 4764-4773.
- MCPHERSON, A. (1998) *Crystallization of Biological Micromolecules*, Cold Spring Harbor Laboratory Press.
- MENDZ, G. L., BURNS, B. P. & HAZELL, S. L. (1995a) Characterisation of glucose transport in *Helicobacter pylori*. *Biochem Biophys Acta*, 1244, 269-276.
- MENDZ, G. L. & HAZELL, S. L. (1993) Glucose phosphorylation in *Helicobacter pylori*. *Arch Biochem Biophys*, 300, 522-525.
- MENDZ, G. L. & HAZELL, S. L. (1995) Aminoacid Utilization by *Helicobacter pylori*. *International journal of Biochemistry & Cell Biology*, 27, 1085-1093.
- MENDZ, G. L., HAZELL, S. L. & BURNS, B. P. (1993) Glucose utilization and lactate production by *Helicobacter pylori*. *J. Gen. Microbiol*, 139, 3023-3028.
- MENDZ, G. L., HAZELL, S. L. & BURNS, B. P. (1994a) The Entner-Doudoroff pathway in *Helicobacter pylori*. *Arch Biochem Biophys*, 312, 349-356.

- MENDZ, G. L., HAZELL, S. L. & SRINIVASAN, S. (1995b) Fumarate Reductase: A target for therapeutic intervention against *Helicobacter pylori*. *Arch Biochem Biophys*, 321, 153-159.
- MENDZ, G. L., HAZELL, S. L. & VANGORKOM, L. (1994b) Pyruvate metabolism in *Helicobacter pylori*. *Arch. Microbiol.*, 162, 187-192.
- MENDZ, G. L., HOLMES, E. M. & FERRERO, R. L. (1998) In situ characterisation of *Helicobacter pylori* arginase. *Biochim. Biophys. Acta*, 1388, 465-477.
- MENDZ, G. L., SHEPLEY, A. J., HAZELL, S. L. & SMITH, M. A. (1997) Purine metabolism and the microaerophily of *Helicobacter pylori*. *Arch. Microbiol.*, 168, 448-456.
- MERCER, W. D., WINN, S. I. & WATSON, H. C. (1976) Twinning in crystals of human skeletal muscle -glyceraldehyde-3-phosphate dehydrogenase. *Journal of Molecular Biology*, 104, 277.
- MISHRA, K. K., SRIVASTAVA, S., GARG, A. & AYYAGARI, A. (2006) Antibiotic susceptibility of *Helicobacter pylori* clinical isolates: comparative evaluation of disk-diffusion and E-test methods. *Curr Microbiol*, 53, 329-334.
- MOESE, S., SELBACH, M., KWOK, T., BRINKMANN, V., KONIG, W., MEYER, T. F. & BACKERT, S. (2004) *Helicobacter pylori* induces AGS cell motility and elongation via independent signaling pathways. *Infect. Immun.*, 72, 3646-3649.
- MOLINARI, M., GALLI, C., DE BERNARD, M., NORAI, N., RUYSSCHAERT, J. M., RAPPUOLI, R. & MONTECUCCO, C. (1998a) The acid activation of *Helicobacter pylori* toxin VacA: structural and membrane binding studies. *Biochem Biophys Res Commun.*, 248, 334-340.
- MOLINARI, M., SALIO, M., GALLI, C., NORAI, N., RAPPUOLI, R., LANZAVECCHIA, S. & MONTECUCCO, C. (1998b) Selective inhibition of Ii-dependent antigen presentation by *Helicobacter pylori* toxin VacA. *J. Exp. Med.*, 187, 135-140.
- MONIOT, S., BRUNO, S., VONRHEIN, C., DIDIERJEAN, C., BOSCHI-MULLER, S., VAS, M., BRICOGNE, G., BRANLANT, G., MOZZARELLI, A. & CORBIER, C. (2008) Trapping of the thioacylglyceraldehyde-3-phosphate dehydrogenase intermediate from *Bacillus stearothermophilus*. Direct evidence for a flip-flop mechanism. *J. Biol. Chem.*, 283, 21693-21702.
- MORAN, A. P. (1996) Pathogenic properties of *Helicobacter pylori*. *Scand. J. Gastroenterol. Suppl.*, 215, S22-S31.
- MORAS, D., OLSEN, K. W., SABESAN, M. N., BUEHNER, M., FORD, G. C. & ROSSMANN, M. G. (1975) Studies of Asymmetry in the Three-dimensional Structure of Lobster D-Glyceraldehyde-3-phosphate Dehydrogenase". *The Journal of Biological Chemistry*, 250, 9137-9162.
- MORRIS, A. & NICHOLSON, G. (1987) Ingestion of *Campylobacter pyloridis* causes gastritis and raised fasting gastric pH. *The American journal of gastroenterology*, 82, 192-199.
- MOUGIN, A., CORBIER, C., SOUKRI, A., WONACOTT, A., BRANLANT, C. & BRANLANT, G. (1988) Use of site-directed mutagenesis to probe the role of Cys149 in the formation of charge-transfer transition in glyceraldehyde-3-phosphate dehydrogenase. *Protein Eng.*, 2, 45-48.
- MUOTIALA, A., HELANDER, I. M., PYHALA, L., KOSUNEN, T. U. & MORAN, A. P. (1992) Low biological activity of *Helicobacter pylori* lipopolysaccharide. *Infect. Immun.*, 60, 1714-1716.
- MURSHUDOV, G. N., VAGIN, A. A., WILSON, K. S. & DODSON, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst.*, D53, 240-255.

- MURTHY, M. R. N., GARAVITO, R. M., JOHNSON, J. E. & ROSSMANN, M. G. (1980) Structure of lobster apo--glyceraldehyde-3-phosphate dehydrogenase at 3.0 Å resolution. *Journal of Molecular Biology*, 138, 859.
- NACHAMKIN, I., ALLOS, B. M. & HO, T. (1998) Campylobacter Species and Guillain-Barre Syndrome. *Clin. Microbiol. Rev.*, 11, 555-567.
- NAGATA, K., NAGATA, Y., SATO, T., FUJINO, M. A., NAKAJIMA, K. & TAMURA, T. (2003) L-Serine, D- and L-proline and alanine as respiratory substrates of *Helicobacter pylori*: correlation between in vitro and in vivo amino acid levels. *Microbiology*, 149, 2023-2030.
- NAUMANN, M., WESSLER, S., BARTSCH, C., WIELAND, B., COVACCI, A., HAAS, R. & MEYER, T. F. (1999) Activation of Activator Protein 1 and Stress Response Kinases in Epithelial Cells Colonized by *Helicobacter pylori* Encoding the *cag* Pathogenicity Island. *J. Biol. Chem.*, 274, 31655-31662.
- NELSON, K., WHITTMAM, T. S. & SELANDER, R. K. (1991) Nucleotide polymorphism and evolution in the glyceraldehyde-3-phosphate dehydrogenase gene (*gapA*) in natural populations of *Salmonella* and *Escherichia coli*. *PNAS*, 88, 6667-6671.
- OSHIMA, K.
- OSHIMA, K., KAKIZAWA, S., NISHIGAWA, H., JUNG, H.-Y., WEI, W., SUZUKI, S., ARASHIDA, R., NAKATA, D., MIYATA, S.-I., UGAKI, M. & NAMBA, S. (2004) Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. *Nature Genetics*, 36, 27-29.
- OTWINOWSKI, Z. & MINOR, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology, Macromolecular Crystallography, part A*, 276, 307-326.
- PAILLOT, A., D'AMBROSIO, K., CORBIER, C., TALFOURNIER, F. O. & BRANLANT, G. (2006) Invariant Thr244 is essential for the efficient acylation step of the non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase from *Streptococcus mutans*. *Biochem J*, 400, 521-530.
- PARKHILL, J., WREN, B. W., MUNGALL, K., KETLEY, J. M., CHURCHER, C., BASHAM, D., CHILLINGWORTH, T., DAVIES, R. M., FELTWELL, T., HOLROYD, S., JAGELS, K., KARLYSHEV, A. V., MOULE, S., PALLEN, M. J., PENN, C. W., QUALL, M. A., RAJANDREAM, M.-A., RUTHERFORD, K. M., VAN VILET, A. H. M., WHITEHEAD, S. & BARRELL, B. G. (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature*, 403, 665-668.
- PARSONNET, J. (1995) The incidence of *Helicobacter pylori* infection. *Alimentary Pharmacology & Therapeutics*, S9, 45-51.
- PARSONNET, J., FRIEDMAN, G. D., ORENTREICH, N. & VOGELMAN, H. (1997) Risk for gastric cancer in people with CagA positive or CagA negative *Helicobacter pylori* infection. *Gut*, 40, 297-301.
- PATTERSON, A. L. (1934) A Fourier series method for the determination of the components of interatomic distances in crystals. *Physics Review*, 46, 372-376.
- PAVAO, F., CASTILHO, M. S., PUPO, M. T., DIAS, R. L. A., CORREA, A. G., FERNANDES, J. B., SILVA, M. F. G. F., MAFEZOLI, J., VIEIRA, P. C. & OLIVA, G. (2002) Structure of *Trypanosoma cruzi* glycosomal glyceraldehyde-3-phosphate dehydrogenase complexed with chalepin, a natural product inhibitor, at 1.95 Å resolution. *FEBS Letters*, 520, 13-17.
- PECZON, B. D. & SPIVEY, H. O. (1972) Catalytic sites in rabbit muscle glyceraldehyde-3-phosphate dehydrogenase. Their number and their kinetic and spectral properties. *Biochemistry*, 11, 2209-2217.

- PEREZ-PEREZ, G. I., ROTHENBACHER, D. & BRENNER, H. (2004) Epidemiology of *Helicobacter pylori* infection. *Helicobacter*, S9, 1-6.
- PERRI, F., FESTA, V. & ANDRIULLI, A. (1998) Treatment of antibiotic resistant *Helicobacter pylori*. *N. Engl. J. Med.*, 339, 53.
- PESCI, E. C. & PICKETT, C. L. (1994) Genetic organization and enzymatic activity of a superoxide dismutase from the microaerophilic human pathogen, *Helicobacter pylori*. *Gene*, 143, 111-116.
- PETERSON, W. L. (1991) *Helicobacter pylori* and peptic ulcer disease. *N. Engl. J. Med.*, 324, 1043-1048.
- PETERSON, W. L., GRAHAM, D. Y., MARSHALL, B. J., BLASER, M. J., GENTA, R. M., KLEIN, P. D., STRATTON, C. W., DRNEE, J., PROKOCIMER, P. & SIEPMAN, N. (1993) Clarithromycin as monotherapy for eradication of *Helicobacter pylori*: a randomized, double-blind trial. *Am. J. Gastroenterol.*, 88, 1860-1864.
- PISTON, S. M., MENDZ, G. L., SRINIVASAN, S. & HAZELL, S. L. (1999) The tricarboxylic acid cycle of *Helicobacter pylori*. *Eur. J. Biochem*, 260, 258-267.
- POUNDER, R. E. & NG, D. (1995) The prevalence of *Helicobacter pylori* infection in different countries. *Alimentary Pharmacology & Therapeutics*, 9, 33-39.
- POWELL, H. R. (1999) The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta Crystallogr D Biol Crystallogr*, 55, 1690-1695.
- RACHER, E. (1955) Mechanism of action and properties of pyridine nucleotide-linked enzymes. *Phys. Rev*, 35, 1-56.
- RACKER, E. & KRIMSKY, I. (1952) THE MECHANISM OF OXIDATION OF ALDEHYDES BY GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE. *J. Biol. Chem.*, 198, 731-743.
- RAMACHANDRAN, G. N. & SASISEKHARAN, V. (1968) Conformation of polypeptides and proteins. *Adv Protein Chem*, 23, 283-248.
- RAUTELIN, H., VON BONSDORFF, C. H., BLOMBERG, B. & DANIELSSON, D. (1994) Ultrastructural study of two patterns in the interaction of *Helicobacter pylori* with neutrophils. *J. Clin. Pathol.*, 47, 667-669.
- READ, R. J. (1986) Improved fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst A*, A42, 140-149.
- READ, R. J. (2001) Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Cryst.*, D57, 1373-1382.
- REYNOLDS, D. J. & PENN, W. C. (1994) Characteristics of *Helicobacter pylori* growth in a defined medium and determination of its amino acid requirement. *Microbiology*, 140, 2649-2656.
- RICCI, C., HOLTON, J. & VAIRA, D. (2007) Diagnosis of *Helicobacter pylori*: Invasive and non-invasive tests. *Best Practice & Research Clinical Gastroenterology*, 21, 299.
- ROGERS, A. B. & FOX, J. G. (2004) Inflammation and Cancer. I. Rodent models of infectious gastrointestinal and liver cancer. *American Journal of Physiology and Gastrointestinal Liver Physiology*, 286, G361-G366.
- ROITEL, O., VACHETTE, P., AZZA, S. & BRANLANT, G. (2003) P but not R-axis interface is involved in cooperative binding of NAD on tetrameric phosphorylating glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus*. *J. Mol. Biol.*, 326, 1513-1522.
- ROTIMI, O., CAIRNS, A., GRAY, S., MOAYYEDI, P. & DIXON, M. F. (2000) Histological identification of *Helicobacter pylori*: comparison of staining methods. *J Clin Pathol*, 53, 756-759.
- SACHS, G., WEEKS, D. L., WEN, Y., MARCUS, E. A., SCOTT, D. R. & MELCHERS, K. (2005) Acid Acclimation by *Helicobacter pylori*. *Physiology*, 20, 429-438.

- SAITO, M., NISHIMURA, K., HASEGAWA, Y., SHINOHARA, T., WAKABYASHI, S., KURIHARA, T., ISHIZUKA, M. & NAGATA, Y. (2007) Alanine racemase from *Helicobacter pylori* NCTC 11637: Purification, characterization and gene cloning. *Life Sciences*, 80, 788-794.
- SAMUEL, M. C., VUGIA, D. L. & SHALLOW, S. (2004) Epidermology of sporadic *Campylobacter* infection in the United States and declining trend in incidence, FoodNet. *Clin Infect Dis*, 38, S165-S174.
- SATIN, B., DEL GIUDICE, G., DELLA BIANCA, V., DUSI, S., LAUDANNA, C., TONELLO, F., KELLECHER, D., RAPPUOLI, R., MONTECUCCO, C. & ROSSI, F. (2000) The neutrophil-activating protein (HP-NAP) of *Helicobacter pylori* is a protective antigen and a major virulence factor. *J. Exp. Med.*, 191, 1467-1476.
- SCOTT, D. R., MARCUS, E. A., WEEKS, D. L. & SACHS, G. (2002) Mechanisms of acid resistance due to the urease system of *Helicobacter pylori*. *Gastroenterology*, 123, 187-195.
- SELBACH, M., MOESE, S., HURWITZ, R., HAUCK, C. R., MEYER, T. F. & BACKERT, S. (2003) The *Helicobacter pylori* CagA protein induces cortactin dephosphorylation and actin rearrangement by c-Src inactivation. *EMBO J.*, 22, 515-528.
- SHEN, Y. Q., LI, J., SONG, S. Y. & LIN, Z. J. (2000) Structure of Apo-glyceraldehyde-3-phosphate Dehydrogenase from *Palinurus versicolor*. *Journal of Structural Biology*, 130, 1-9.
- SHERWOOD, D. (1976) *Crystals, X-rays and Proteins*, Longman.
- SHUMAN, S. (1994) Novel Approach to Molecular Cloning and Polynucleotide Synthesis Using Vaccinia DNA Topoisomerase. *J. Biol. Chem.*, 269, 32678-32684.
- SIEBERS, B. & SCHONHEIT, P. (2005) Unusual pathways and enzymes of central carbohydrate metabolism in Archaea. *Curr Opin Microbiol*, 8, 695-705.
- SIMPSON, F. J., PERLIN, A. S. & SIEBEN, A. S. (1966) Erythrose-4-phosphate. *Methods Enzymol*, 9, 35-38.
- SKARZYNSKI, T., MOODY, P. C. E. & WONACOTT, A. J. (1987) Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1.8 Å resolution. *J. Mol. Biol.*, 193, 171-187.
- SOPHOS, N. A. & VASILIOU, V. (2003) Aldehyde dehydrogenase gene superfamily: the 2002 update. *Chemico-Biological Interactions*, 143-144, 5.
- SPIEGELHALDER, C., GERSTENECKER, B., KERSTEN, A., SCHILTZ, E. & KIST, M. (1993) Purification of *Helicobacter pylori* superoxide dismutase and cloning and sequencing of the gene. *Infect. Immun.*, 61, 5315-5325.
- STINGL, K., ALTENDORF, K. & BAKKER, E. P. (2002) Acid survival of *Helicobacter pylori*: how does urease activity trigger cytoplasmic pH homeostasis? *Trends in Microbiology*, 10, 70-74.
- SUERBAUM, S. & JOSEPHANS, C. (2007) *Helicobacter pylori* evolution and phenotypic diversification in a changing host. *Nature Reviews Microbiology*, 5, 441-451.
- SUERBAUM, S., JOSEPHANS, C., STERZENBACH, T., DRESCHER, B., BRANDT, P., BELL, M., DROGE, M., FARTMANN, B., FISCHER, H. P., GE, Z. M., HORSTER, A., HOLLAND, R., KLEIN, K., KONIG, J., MACKO, L., MENDZ, G. L., NYAKATURA, G., SCHAUER, D. B., SHEN, Z. L., WEBER, J., FROSCH, M. & FOX, J. G. (2003) The complete genome sequence of the carcinogenic bacterium *Helicobacter hepaticus*. *Proc Nat Acad Sci*, 100, 7901-7906.
- SUERBAUM, S., SMITH, J. M., BAPUMIA, K., MORELLI, G., SMITH, N. H., KUNSTMANN, E., DYREK, E. & ACHTMAN, M. (1998) Free recombination within *Helicobacter pylori*. *PNAS*, 95, 12619-12624.

- SUZUKI, T., MATSUO, K., ITO, H., SAWAKI, A., HIROSE, K., WAKAI, K., SATO, S., NAKAMURA, T., YAMAO, K., UEDA, R. & TAJIMA, K. (2006) Smoking increases the treatment failure for *Helicobacter pylori* eradication. *Am. J. Med.*, 119, 217-224.
- TANIGUCHI, Y., KIMURA, K., SOHARA, H., SHIRASAKI, A., KAWADA, H., SATOH, K., KIHARA, K., WANG, X. M., TAKIMOTO, T., GOTO, Y., TAKATORI, K., IDA, K. & KAJIWARA, M. (1996) Simple ¹³C-urea breath test with infra-red spectrophotometer. *J. Gastroenterol.*, 31, S37-S40.
- TANNER, J. J., HECHT, R. M. & KRAUSE, K. L. (1996) Determinants of Enzyme Thermostability Observed in the Molecular Structure of *Thermus aquaticus* D-Glyceraldehyde-3-phosphate Dehydrogenase at 2.5 Å Resolution. *Biochemistry*, 35, 2597-2609.
- TEN EYCK, L. F. (1985) Fast Fourier transform calculation of electron density maps. *Methods in Enzymology*, 115, 324-337.
- TERAI, S., IJIMA, K., KATIO, K., DAIRAKU, N., SUZUKI, T., YOSHIDA, M., KOIKE, T., KITAGAWA, Y., IMATANI, A., SEKINE, H., OHARA, S. & SHIMOSEGAWA, T. (2008) Long-term outcomes of gastric mucosa-associated lymphoid tissue lymphomas after *Helicobacter pylori* eradication therapy. *Tohoku. J. Exp. Med.*, 214, 79-87.
- TOMB, J.-F., WHITE, O., KERLAVAGE, A. R., CLAYTON, R. A., SUTTON, G. G., FLEISCHMAN, R. D., KETCHUM, K. A., KLENK, H. P., GILL, S., DOUGHTERY, B. A., NELSON, K., QUACKENBUSH, J., ZHOU, L., KIRKNESS, E. F., PETERSON, S., LOFTUS, B., RICHARDSON, D., DODSON, R., KHALAK, H. G., GLODEK, A., MCKENNY, K., FITZGERALD, L. M., N., L., ADAMS, M. D., HICKEY, E. K., BERG, D. E., GOCAYNE, J. D., UTTERBACK, T. R., PETERSON, J. D., KELLY, J. M., COTTON, M. D., WEIDMAN, J. M., FUJII, C., BOWMAN, C., WATTHEY, L., WALLIN, E., HAYES, W. S., BORODOVSKY, M., KARP, P. D., SMITH, H. O., FRASER, C. M. & VENTER, J. C. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature*, 388, 539-547.
- TONELLO, F., DUNDON, W. G., SATIN, B., MOLINARI, M., TOGNON, G., GRANDI, G., DEL GIUDICE, G., RAPPUOLI, R. & MONTECUCCO, C. (1999) The *Helicobacter pylori* neutrophil-activating protein is an iron-binding protein with dodecameric structure. *Mol. Microbiol.*, 34, 238-246.
- TRENTHAM, D. R. (1971) Rate-determining processes and the number of simultaneously active sites of D-glyceraldehyde 3-phosphate dehydrogenase. *Biochem J*, 122, 71-77.
- TSUTSUMI, R., HIGASHI, H., HIGUCHI, M., OKADA, M. & HATAKEYAMA, M. (2003) Attenuation of *Helicobacter pylori* CagA x SHP-2 signaling by interaction between CagA and C-terminal Src kinase. *J. Biol. Chem.*, 278, 3664-3670.
- UEDO, N., ISHIHARA, R., IISHI, H., YAMAMOTO, S., YAMAMOTO, S., YAMADA, T., IMANAKA, K., TAKEUCHI, Y., HIGASHINO, K., ISHIGURO, S. & TATSUTA, M. (2006) A new method of diagnosing gastric intestinal metaplasia: narrow-band imaging with magnifying endoscopy. *Endoscopy*, 819-824.
- VALVERDE, F., LOSUDA, M. & SERRANO, A. (1997) Functional complementation of an *E. coli* gap mutant supports an amphibolic role for NAD(P) dependent GAPDH of *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.*, 179, 4513-4522.
- VAN DER OOST, J., SCHUT, G., KENGEN, S. W., HAGEN, W. R., THOMM, M. & DE VOS, W. M. (1998) The ferredoxin-dependent conversion of glyceraldehyde-3-phosphate in the hyperthermophilic archaeon *Pyrococcus furiosus* represents a novel site of glycolytic regulation. *J. Biol. Chem.*, 273, 28149-28154.
- VELAYUDHAN, J., JONES, M. A., BARROW, P. A. & KELLY, D. J. (2004) L-Serine Catabolism via an Oxygen-Labile L-Serine Dehydratase Is Essential for Colonization of the Avian Gut by *Campylobacter jejuni*. *Infection and Immunity*, 72, 260-268.

- VELDHUYZEN VAN ZANTEN, S. O. J., DIXON, M. F. & LEE, A. (1999) The gastric transition zones: neglected links between gastroduodenal pathology and Helicobacter ecology. *Gastroenterology*, 116, 1217-1229.
- VELICK, S. F. & HAYES, J. E., JR. (1953) PHOSPHATE BINDING AND THE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE REACTION. *J. Biol. Chem.*, 203, 545-562.
- WALSH, F. S., CRONIN, M. & KOBLAR, S. (1991) Association between glycoconjugate antibodies and Campylobacter infection in patients with Guillain-Barre syndrome. *J Neuroimmunol*, 34, 43-51.
- WIERENGA, R. K. & HOL, W. G. (1983) Predicted nucleotide-binding properties of p21 protein and its cancer-associated variant. *Nature*, 302, 842-844.
- WOLIN, M. J., WOLIN, E. A. & JACOBS, N. J. (1961) CYTOCHROME-PRODUCING ANAEROBIC VIBRIO, VIBRIO SUCCINOGENES, SP. N. *J. Bacteriol.*, 81, 911-917.
- XIANG, Z., CENSINI, S., BAYELI, P. F., TELFORD, P. F., FIGURA, N., RAPPUOLI, R. & COVACCI, A. (1995) Analysis of expression of CagA and VacA virulence factors in 43 strains of Helicobacter pylori reveals that clinical isolates can be divided into two major types and that CagA is not necessary for expression of the vacuolating cytotoxin. *Infect. Immun.*, 63, 94-98.
- YAMAZAKI, S., YAMAKAWA, A., ITO, Y., OHTANI, M., HIGASHI, H., HATAKEYAMA, M. & AZUMA, T. (2003) The CagA protein of Helicobacter pylori is translocated into epithelial cells and binds to SHP-2 in human gastric mucosa. *J. Infect. Dis.*, 187, 334-337.
- YEUNG, C. K., FU, K. H., YUEN, K. Y., NG, W. F., TSANG, T. M., BRANICKI, F. J. & SAING, H. (1990) Helicobacter pylori and associated duodenal ulcer. *Arch. Dis. Child.*, 65, 1212-1216.
- YUKI, N. (2001) Infectious origins of, and molecular mimicry in, Guillain-Barre and Fisher syndromes. *The Lancet Infectious Diseases*, 1, 29-37.
- YUKI, N., YOSHINO, H., SATO, S. & MIYATAKE, T. (1990) Acute axonal polyneuropathy associated with anti-GM1 antibodies following Campylobacter enteritis. *Neurology*, 40, 1900-1902.
- YUN, M., PARK, C. G., KIM, J. Y. & PARK, H. W. (2000) Structural Analysis of Glyceraldehyde 3-Phosphate Dehydrogenase from Escherichia coli: Direct Evidence of Substrate Binding and Cofactor-Induced Conformational Changes. *Biochemistry*, 39, 10702-10710.
- ZHAO, G., PEASE, A. J., BHARANI, N. & WINKLER, M. E. (1995) Biochemical characterisation of gapB-encoded Erythrose 4-Phosphate Dehydrogenase of Escherichia coli K-12 and its possible role in pyridinol 5'-phosphate biosynthesis. *Journal of Bacteriology*, 177, 2804-2812.
- ZOOROB, R. J. (1994) NIH Consensus Conference. Helicobacter pylori in peptic ulcer disease. NIH Consensus Development Panel on Helicobacter pylori in Peptic Ulcer Disease. *NIH Consensus Conference*. JAMA.