L-RIBULOSE-5-PHOSPHATE 4-EPIMERASE:
EPIMERISATION THROUGH CARBON-CARBON BOND CLEAVAGE

by

ANNE ELIZABETH JOHNSON
BSc., University of Toronto, 1992

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
in
THE FACULTY OF GRADUATE STUDIES
Department of Chemistry

We accept this thesis as conforming
to the required standard:

THE UNIVERSITY OF BRITISH COLUMBIA
April, 1998
© Anne E. Johnson, 1998
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Chemistry

The University of British Columbia
Vancouver, Canada

Date 9 April 1998

DE-6 (2/88)
Abstract

L-Ribulose-5-phosphate 4-epimerase, a bacterial enzyme which catalyses the final step in L-arabinose metabolism, interconverts L-ribulose-5-phosphate and D-xylulose-5-phosphate independently of NAD⁺. The labile stereocentre does not bear an acidic proton; therefore, a simple deprotonation-reprotonation mechanism cannot be followed.

The epimerase was cloned from *Escherichia coli* into a highly efficient overexpression vector. The recombinant enzyme was found to contain a mixture of divalent zinc, manganese and copper ions. A preparation of a homogeneous sample of the Zn²⁺ form of the recombinant enzyme displayed kinetic constants similar to those of the naturally abundant epimerase from *E. coli*.

Amino acid sequence similarity between the epimerase and the Class II L-fuculose-1-phosphate aldolase suggests that these two enzymes may be evolutionarily related and that the epimerisation, which is metal-dependent, may occur through carbon-carbon bond cleavage and reformation. Three conserved residues (H95, H97 and D76) which are thought to be the metal ion ligands in the epimerase have been independently altered by site directed mutagenesis to asparagine. The resulting mutant epimerases exhibit low $k_{cat}$ values. The H95N and H97N epimerases have a reduced affinity for Zn²⁺ and lose metal readily, while
the D76N epimerase which has an affinity for Zn\textsuperscript{2+} comparable to that of the wild type epimerase loses metal upon extended dialysis. These observations serve to establish a structural link between the active sites of the epimerase and the aldolase.

The H97N epimerase was found capable of catalysing the aldol addition between dihydroxyacetone and glycolaldehyde phosphate (the unbound forms of the proposed reaction intermediates) to form an equilibrium mixture of L-ribulose-5-phosphate and D-xylulose-5-phosphate. In addition, the epimerase was able to release dihydroxyacetone from an equilibrating pool of L-ribulose-5-phosphate and D-xylulose-5-phosphate. These observations of aldolase activity establish that the active site of the epimerase is capable of catalysing carbon-carbon bond cleavage, and support the notion that the epimerase and the aldolase are evolutionarily related. Glycolaldehyde phosphate was shown to be a competitive inhibitor of the H97N enzyme with a $K_I$ of 0.37 mM. The wild type enzyme was not significantly inhibited at 5 mM. The H97N mutation appears to have created a “leaky” epimerase which can bind to the normal reaction intermediates and generate them from the unbound aldol cleavage products.
Table of Contents

Abstract ........................................................................................................... ii

Table of Contents .......................................................................................... iv

List of Tables ................................................................................................... viii

List of Figures .................................................................................................. ix

Abbreviations and Symbols ........................................................................... xii

Acknowledgements ........................................................................................ xvii

Chapter I: Enzyme Catalysed Stereochemical Inversion
and Carbon-Carbon Bond Cleavage ............................................................... 1
Stereochemical Inversion ............................................................................... 1
Carbohydrate Epimerases ............................................................................ 3
L-Ribulose-5-Phosphate 4-Epimerase ............................................................. 8
Role and Distribution .................................................................................. 8
Previous Mechanistic Studies ....................................................................... 10
Proposed Mechanisms ............................................................................... 11
Relationship with Other Enzymes ............................................................... 13
Aldolases ....................................................................................................... 15
Class I Aldolases ......................................................................................... 16
Class II Aldolases ....................................................................................... 17
Summary ....................................................................................................... 22

Chapter II: Cloning, Overexpression and Characterization of
L-Ribulose-5-Phosphate 4-Epimerase from Escherichia coli ......................... 24
Introduction .................................................................................................. 24
Overexpression Systems ............................................................................... 24
The araBAD Operon .................................................................................... 28
Physical Characterization of the Mutant L-Ribulose-5-Phosphate 4-Epimerases ................................................................. 87
Activity of the Wild Type and Mutant Enzymes ......................................... 90
Zinc Ion Content of the Zn$^{2+}$ Reconstituted Epimerases ...................... 94
Activity of the Cobalt-Substituted Epimerases ..................................... 95
UV-Visible Spectra of the Co(II) Substituted Enzymes .............................. 97
Co$^{3+}$-Modification of the Co$^{2+}$-Substituted Wild Type 4-Epimerase .... 97
Conclusion ......................................................................................... 99

Experimental Methods ......................................................................... 101
Data Base Searches and Protein Sequence Alignments ......................... 101
Bacterial and Phage Strains ................................................................. 101
ssDNA Rescue of pRE1 ......................................................................... 102
Site-Directed Mutagenesis ................................................................... 102
H95N ................................................................................................... 103
H97N ................................................................................................... 106
D76N ................................................................................................... 107
H171N ................................................................................................... 107
DNA sequencing .................................................................................... 108
Enzyme Purification ............................................................................... 108
Calculation and Determination of Mutant Enzymes’ Subunit Masses .... 109
Determination of Tetrameric Structure of the Native Enzymes ............. 109
Preparation of Metal-Free Buffers and Glassware ................................ 110
Preparation of Apoenzyme and Reactivation with Zn(II) or Co(II) ...... 110
Circular Dichroism and Thermal Stability ........................................... 111
Determination of Zinc(II) Content ......................................................... 112
Assay for L-ribulose-5-phosphate 4-epimerase activity ......................... 112
Assay for Apoenzyme Activity ............................................................. 113
UV-Visible Spectra of the Co(II)-Substituted 4-Epimerases ................ 113
Co(III) Modification of the Co(II)-Substituted Wild Type L-Ribulose-
Phosphate 4-Epimerase ..................................................................... 114

Chapter IV: L-Ribulose-5-Phosphate 4-Epimerase Follows a
Retroaldol/Aldol Mechanism ............................................................... 115
Introduction ......................................................................................... 115
Results and Discussion ......................................................................... 119
Intermediate Identification with the Wild Type L-Ribulose-5-
Phosphate 4-Epimerase ..................................................................... 119
Intermediate Identification Using Mutant 4-Epimerases ..................... 122
Detection of Intermediate Release ...................................................... 131
An Aldol-Like Mechanism for L-Ribulose-5-Phosphate
4-Epimerase ......................................................................................... 133
Conclusion ......................................................................................... 135
List of Tables

Table 2.1. Predicted Subunit Molecular Masses of L-Ribulose-5-Phosphate 4-Epimerase ................................................................. 41
Table 2.2. Comparison of $K_m$ and $K_{eq}$ Constants Determined for L-Ribulose-5-Phosphate 4-Epimerase .................................................. 47
Table 2.3. Primers Used for PCR Amplification of the araB Gene .................. 53
Table 2.4. Primers Used for PCR Amplification of the araD Gene ................ 55
Table 2.5. Primers Used in Sequencing the araD Gene ................................. 58
Table 3.1. Molecular Masses of Mutant 4-Epimerase Subunits ....................... 86
Table 3.2. Melting Temperature of Wild Type and Mutant 4-Epimerases .......... 89
Table 3.3. Kinetic Parameters of the Zinc-Reconstituted Wild Type and Mutant L-Ribulose-5-Phosphate 4-Epimerases ................................. 92
Table 3.4. Kinetic Parameters of the Zinc-Reconstituted Wild Type and Mutant L-Ribulose-5-Phosphate 4-Epimerases In the Presence of Additional Zn$^{2+}$ 93
Table 3.5. Zinc Content of Wild Type and Mutant 4-Epimerases .................... 94
Table 3.6. Kinetic Parameters of the Cobalt-Substituted Wild Type and Mutant L-Ribulose-5-Phosphate 4-Epimerases ................................. 96
Table 3.6. Primers Used for Site-Directed Mutagenesis of the araD Gene ............ 104
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Stereochemical Inversion</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>The Mechanism of Epimerisation of D-Ribulose-5-Phosphate 3-Epimerase</td>
<td>4</td>
</tr>
<tr>
<td>1.3</td>
<td>The Mechanism of Epimerisation of UDP-Glucose 4-Epimerase</td>
<td>4</td>
</tr>
<tr>
<td>1.4</td>
<td>Proposed Mechanism of GDP-D-Mannose 3,5-Epimerisation</td>
<td>5</td>
</tr>
<tr>
<td>1.5</td>
<td>Proposed Mechanism for Epimerisation by UDP-N-Acetylglucosamine 2-Epimerase</td>
<td>7</td>
</tr>
<tr>
<td>1.6</td>
<td>Interconversion of L-Ribulose-5-Phosphate and D-Xylulose-5-Phosphate</td>
<td>8</td>
</tr>
<tr>
<td>1.7</td>
<td>Bacterial L-Arabinose Utilization Pathway</td>
<td>9</td>
</tr>
<tr>
<td>1.8</td>
<td>Proposed Pathways for L-Ribulose-5-Phosphate 4-Epimerase Consistent with Observations</td>
<td>12</td>
</tr>
<tr>
<td>1.9</td>
<td>Reaction Catalysed by D-erythro-Dihydroneopterin Triphosphate 2'-Epimerase</td>
<td>14</td>
</tr>
<tr>
<td>1.10</td>
<td>Mechanism of Class I Aldolases</td>
<td>16</td>
</tr>
<tr>
<td>1.11</td>
<td>Mechanism of Class II Aldolases</td>
<td>17</td>
</tr>
<tr>
<td>1.12</td>
<td>Coordination of the Active Site Zn$^{2+}$ Ion by Phosphoglycolohydroxamate and the Enediolate of Dihydroxyaceton Phosphate</td>
<td>19</td>
</tr>
<tr>
<td>1.13</td>
<td>A Proposed Mechanism for Catalysis by the Class II Aldolase L-Fuculose-1-Phosphate Aldolase</td>
<td>19</td>
</tr>
<tr>
<td>1.14</td>
<td>Epimerisation of D-Tagatose-1,6-diphosphate Catalysed by D-Tagatose-1,6-Diphosphate Aldolase</td>
<td>20</td>
</tr>
<tr>
<td>1.15</td>
<td>Parallel Pathways of L-Fucose and L-Rhamnose Dissimilation</td>
<td>21</td>
</tr>
<tr>
<td>2.1</td>
<td>A genetic map of the <em>E. coli</em> araC and <em>araBAD</em> operons indicating the proteins encoded by the <em>araBAD</em> operon and the reactions which these proteins catalyse</td>
<td>29</td>
</tr>
<tr>
<td>2.2</td>
<td>L-Ribulose-5-Phosphate 4-Epimerase Interconverts L-Ribulose-5-Phosphate and D-Xylulose-5-Phosphate</td>
<td>30</td>
</tr>
<tr>
<td>2.3</td>
<td>Phosphorylation of L-Ribulose by L-Ribulokinase Provides L-Ribulose-5-Phosphate, the Substrate for L-Ribulose-5-Phosphate 4-Epimerase</td>
<td>32</td>
</tr>
<tr>
<td>2.4</td>
<td>Proton-NMR Spectrum of Purified L-Ribulose-5-Phosphate (free acid) at 500 MHz (in D$_2$O)</td>
<td>36</td>
</tr>
</tbody>
</table>
Figure 2.5. SDS-PAGE of Crude and Purified Protein ...................................... 38
Figure 2.6. A Coupled Assay for \( \text{L-ribulose-5-phosphate 4-epimerase} \) Activity ......................... 43
Figure 2.7. An Alternate Assay for \( \text{L-ribulose-5-phosphate 4-epimerase} \) Activity ........................................... 44
Figure 2.8. Equilibrium between \( \text{L-ribulose-5-phosphate} \) and \( \text{D-xylulose-5-phosphate} \) as Determined by \( ^1 \text{H-NMR} \) (500 MHz) .................................................. 49
Figure 3.1. \( ^1 \text{H-NMR} \) of the Co\(^{2+}\)-substituted \( \text{B. fragilis} \) metallo-\( \beta \)-lactamase (1 mM) at pH 7.4 ............................................. 73
Figure 3.2. Coordination of \( \text{Zn}^{2+} \) in the Active Site of \( \text{L-fuculose-1-phosphate aldolase} \) .................................................. 76
Figure 3.3. Rotation of E73 of \( \text{L-fuculose-1-phosphate aldolase} \) Away from \( \text{Zn}^{2+} \) on Binding of Phosphoglycolohydroxamate ........................................... 77
Figure 3.4. Amino Acid Residues ................................................................. 78
Figure 3.5. The RCPCR Method of Site-Directed Mutagenesis ......................... 81
Figure 3.6. Creation of a Megaprimer Prior to RCPCR ....................................... 82
Figure 3.7. Amino Acid Sequence Alignment .................................................. 84
Figure 3.8. CD Spectra of the Wild Type and Mutant \( \text{L-ribulose-5-phosphate 4-epimerases} \) ........................................... 88
Figure 3.9. Native Structure of \( \text{L-ribulose-5-phosphate 4-epimerases} \) ......................... 91
Figure 3.10. Deconvoluted Electrospray Ionization Mass Analysis of the Co(III) Wild Type 4-epimerase ................................ .................. 98
Figure 3.11. A Schematic Diagram of pRE1 Indicating Relative Primer Binding Sites and Locations of Restriction Sites on pRE1 and Restriction Sites Introduced by Site Directed Mutagenesis .......................................... 105
Figure 4.1. Class II \( \text{L-fuculose-1-phosphate aldolase (A) and L-ribulose-5-phosphate 4-epimerase (B)} \) Can Utilise Similar Mechanistic Strategies ........................................... 116
Figure 4.2. Reduction of Dihydroxyacetone to Glycerol by Glycerol Dehydrogenase ................................................................. 117
Figure 4.3. Formation of An Epimeric Mixture of \( \text{L-ribulose-5-phosphate} \) and \( \text{D-xylulose-5-phosphate} \) From Dihydroxyacetone and Glycolaldehyde Phosphate (5 mM) ........................................... 124
Figure 4.4. \( ^1 \text{H-NMR} \) (500 MHz, \( D_2O \), solvent suppressed) of the H97N Catalysed Coupling Reaction Products at 72 h ...................... 127
Figure 4.5. Glycolaldehyde Phosphate is a Competitive Inhibitor of Epimerisation by H97N L-ribulose-5-phosphate 4-epimerase ...................... 129
Figure 4.6. \( K_i \) Determination for H97N L-ribulose-5-phosphate 4-epimerase Inhibition by Glycolaldehyde Phosphate .............................................. 130
Figure 4.7. Dihydroxyacetone Release by H97N and Wild Type L-ribulose-5-phosphate 4-epimerases Equilibrating 10 mM L-ribulose-5-phosphate and D-xylulose-5-phosphate .......................... 132
Figure 4.8. Synthetic Route to Glycolaldehyde Phosphate from Allyl Alcohol ................. 136
Figure 4.9. Determination of Sugar Content .................................................... 141
Figure 5.1. Hydroxamate Derivatives Which Could Inhibit L-Ribulose-5-Phosphate 4-Epimerase. .................................................. 147
Figure 5.2. Preparation of Pentitol Phosphates as Potential Inhibitors of L-Ribulose-5-Phosphate 4-Epimerase. .............................. 148
Figure 5.3. Addition of Glycolate to Glycidol Phosphate. ................................. 149
Figure 5.4. Potential Active Site Labelling Agents. ........................................... 150
Figure B.1. 500 MHz Proton-NMR Spectrum of Commercial D-Xylulose-5-Phosphate (Sodium Salt) in D$_2$O (Solvent Suppressed). .............. 162
Figure B.2. 500 MHz DQFCOSY of L-Ribulose-5-Phosphate (Free Acid) in D$_2$O (Solvent Suppressed). .............................................. 163
Figure B.3. 500 MHz DQFCOSY of Commercial D-Xylulose-5-Phosphate (Sodium Salt) in D$_2$O (Solvent Suppressed). .............................. 164
Figure B.4. 200 MHz Proton-NMR of Allyl Phosphate (Cyclohexylammonium Salt) in D$_2$O. ................................................................. 165
Figure B.5. 200 MHz Proton-NMR Spectrum of Glycolaldehyde Phosphate (Ca$^{2+}$ Salt) in Acidic D$_2$O. .................................................. 166
Figure C.1. Proteins Sharing a Homologous Domain with E. coli L-Ribulose-5-Phosphate 4-Epimerase. ................................................. 168
Figure D.1. Direct Plot of Kinetic Data from Zinc(II)-Substituted L-Ribulose-5-Phosphate 4-Epimerases. ............................................... 169
Figure D.2. Direct Plots of Kinetic Data for Zinc(II)-Substituted L-Ribulose-5-Phosphate 4-Epimerases Assayed in the Presence of 0.1 mM Zn$^{2+}$ 170
Figure D.3. Direct Plots of Kinetic Data for Cobalt(II)-Substituted L-Ribulose-5-Phosphate 4-Epimerases. ............................................. 171
Abbreviations and Symbols

αGDH  \(\alpha\)-glycerophosphate dehydrogenase
δ chemical shift (in NMR)
ε extinction coefficient
\(\varepsilon_d\) extinction coefficient of denatured protein
\(\varepsilon_n\) extinction coefficient of native protein
\(\theta_{\text{obs}}\) observed ellipticity
Å Angstrom(s)
A (deoxy)adenosine (in nucleic acids); alanine (Ala) (in proteins); absorbance
\(A_d\) absorbance of a solution of denatured protein
\(A_n\) absorbance of a solution of native protein
ADP adenosine diphosphate
araA gene encoding L-arabinose isomerase
AraA L-arabinose isomerase
araB gene encoding L-ribulokinase
AraB L-ribulokinase
araBAD the group of araA, araB and araD genes.
araC gene encoding the regulatory protein for the araBAD operon
araD gene encoding L-ribulose-5-phosphate 4-epimerase
AraD L-ribulose-5-phosphate 4-epimerase
Asn asparagine; N
Asp aspartic acid; D
ATCC American Type Culture Collection
ATP adenosine triphosphate
B. fragilis Bacteroides fragilis
B. subtilis Bacillus subtilis
BSA bovine serum albumin
C (deoxy)cytosine (in nucleic acids); cysteine (Cys) (in proteins)
ca. circa
CD circular dichroism
CoA coenzyme A
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>d</td>
<td>doublet</td>
</tr>
<tr>
<td>D</td>
<td>aspartic acid; Asp</td>
</tr>
<tr>
<td>D₂O</td>
<td>deuterium oxide</td>
</tr>
<tr>
<td>D76</td>
<td>aspartate at position 76 of the AraD sequence</td>
</tr>
<tr>
<td>D76N</td>
<td>AraD in which D76 has been replaced with an asparagine</td>
</tr>
<tr>
<td>Da</td>
<td>dalton(s)</td>
</tr>
<tr>
<td>dd</td>
<td>doublet of doublet</td>
</tr>
<tr>
<td>DHAP</td>
<td>dihydroxyacetone phosphate</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dR5P</td>
<td>D-ribose-5-phosphate</td>
</tr>
<tr>
<td>ds</td>
<td>double stranded</td>
</tr>
<tr>
<td>dt</td>
<td>doublet of triplet</td>
</tr>
<tr>
<td>dTDP</td>
<td>deoxy thymidine diphosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>dut</td>
<td>gene encoding dUTPase (dut indicates an inability to make active dUTPase)</td>
</tr>
<tr>
<td>DXu5P</td>
<td>D-xyulose-5-phosphate</td>
</tr>
<tr>
<td>E</td>
<td>glutamic acid; Glu</td>
</tr>
<tr>
<td>E73</td>
<td>glutamate at position 73 of the FucA sequence</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EC</td>
<td>Enzyme Commission (classification number) of the International Union of Biochemistry</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamine tetraacetic acid, disodium salt</td>
</tr>
<tr>
<td>ESIMS</td>
<td>electrospray ionization mass spectrometry</td>
</tr>
<tr>
<td>F1</td>
<td>a bacteriophage</td>
</tr>
<tr>
<td>FucA</td>
<td>L-fuculose-1-phosphate aldolase</td>
</tr>
<tr>
<td>FucI</td>
<td>L-fucose isomerase</td>
</tr>
<tr>
<td>FucK</td>
<td>L-fuculose kinase</td>
</tr>
<tr>
<td>G</td>
<td>(deoxy)guanosine (in nucleic acids); or glycine (Gly) (in proteins)</td>
</tr>
<tr>
<td>GAP</td>
<td>glyceraldehyde-3-phosphate</td>
</tr>
<tr>
<td>GAPDH</td>
<td>glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>αGDH</td>
<td>α-glycerophosphate dehydrogenase</td>
</tr>
<tr>
<td>GDP</td>
<td>guanosine diphosphate</td>
</tr>
<tr>
<td>Glu</td>
<td>glutamic acid; E</td>
</tr>
<tr>
<td>h</td>
<td>hour(s)</td>
</tr>
<tr>
<td>H</td>
<td>histidine; His</td>
</tr>
<tr>
<td>H92</td>
<td>histidine at position 92 of FucA sequence</td>
</tr>
<tr>
<td>H94</td>
<td>histidine at position 94 of FucA sequence</td>
</tr>
<tr>
<td>H95</td>
<td>histidine at position 95 of AraD sequence</td>
</tr>
<tr>
<td>H95N</td>
<td>AraD in which H95 has been replaced with an asparagine</td>
</tr>
<tr>
<td>H97</td>
<td>histidine at position 97 of AraD sequence</td>
</tr>
<tr>
<td>H97N</td>
<td>AraD in which H97 has been replaced with an asparagine</td>
</tr>
</tbody>
</table>
H155  histidine at position 155 of FucA sequence
H171  histidine at position 171 or the AraD sequence
H171N AraD in which H171 has been replaced with an asparagine
HEPES N-[2-hydroxyethyl]piperazine-N'-[2-ethane sulphonic acid]
His  histidine; H
HPLC  high pressure/performance liquid chromatography
Hz  Hertz
I  isoleucine; Ile
ICPMS  inductively coupled plasma mass spectrometry
ID  inner diameter
J  coupling constant (in NMR)
K. pneumoniae  Klebsiella pneumoniae
$k_{cat}$  catalytic rate constant (turnover number)
kDa  kilodalton(s)
$K_{eq}$  equilibrium constant
$K_I$  dissociation constant for an enzyme-inhibitor complex
$K_M$  Michaelis constant
lacI  gene encoding the lac repressor
lacZ  gene encoding β-galactosidase
L. pentoaceticus Lactobacillus pentoaceticus
L. pentosus Lactobacillus pentosus
L. plantarum Lactobacillus plantarum
LB  Luria Bertani medium
LRu5P  L-ribulose-5-phosphate
LSIMS  liquid soft ionization mass spectrometry
m  multiplet
M13K07 a “helper” phage
mdeg  millidegree(s)
mRNA  messenger RNA
N  asparagine; Asn
NAD$^+$  nicotinamide adenine dinucleotide
NADH  nicotinamide adenine dinucleotide, reduced form
NMR  nuclear magnetic resonance
NMWL  nominal molecular weight limit
nt  nucleotide(s)
dNTP  deoxynucleotide triphosphate
PAGE  polyacrylamide gel electrophoresis
pAJ1 plasmid which overexpresses the H95N epimerase
pAJ2 plasmid which overexpresses the H97N epimerase
pAJ3 plasmid which overexpresses the D76N epimerase
pBS a protein overexpression vector
pBSTIM a modified pBS plasmid which overexpresses triosephosphate
XV

isomerase

P/C/I
phenol: chloroform: isoamyl alcohol

PCR
polymerase chain reaction

PDB
Protein Data Bank

PEG
polyethylene glycol

pET-11a
a protein overexpression vector

pHC5
a plasmid containing the polB and araD genes

ppb
parts per billion

ppm
parts per million

pRE1
a overexpression plasmid for L-ribulose-5-phosphate 4-epimerase

psi
pounds per square inch

RBS
ribosomal binding site

RCPCR
recombinant circle polymerase chain reaction

RhaA
L-rhamnose isomerase

RhaB
L-rhamnulose kinase

RhaD
L-rhamnulose-1-phosphate aldolase

RNA
ribonucleic acid

rpm
rotations per minute

s
singlet

S7P
sedoheptulose-7-phosphate

S. typhimurium
Salmonella typhimurium

SD
Shine-Dalgarno

SDS
sodium dodecyl sulphate

ss
single stranded

t
triplet

T
(deoxy)thymidine (in nucleic acids); threonine (Thr) (in proteins)

T4
a bacteriophage

T7
a bacteriophage

TIM
triosephosphate isomerase

Tn
transposon

tpi
gene encoding triosephosphate isomerase

TPP
thiamine pyrophosphate; cocarboxylase

Tris
tris(hydroxy)amino methane

Trp.
tryptophan

tt
triplet of triplets

Tyr
tyrosine

V
valine; Val

V₀
initial reaction rate

vol
(by) volume

U
unit (of enzyme activity), which is the amount of enzyme required to catalyse formation of 1 µmol of product per minute

UDP
uridine diphosphate
gene encoding uracil-N-glycosylase (*ung* indicates the inability to make the active enzyme)

**UV**
ultraviolet

**Vis**
visible
Acknowledgements

I would like to thank Mum, Dad, and Jean for their support and encouragement throughout the duration of my doctoral studies. I cherish the wonderful friendship Claire Johnson and I have had during the last 5½ years. Finally, I thank my husband Philip Johnson for his love and encouragement, and for running NMR experiments for me.
Chapter I

Enzyme Catalysed Stereochemical Inversion
and Carbon-Carbon Bond Cleavage

Stereochemical Inversion

Bacteria benefit from their ability to utilize a wide variety of sugars and amino acids of unusual stereochemistry that are not metabolized by animals. These compounds are often enantiomers or diastereomers of common biomolecules and can serve as unique biosynthetic building blocks or energy sources. Thus, stereochemical control is an important feature in many biosynthetic and metabolic pathways, and often involves the inversion of configuration about an asymmetric carbon centre catalysed by enzymes known as racemases and epimerases (Adams, 1976; Glaser, 1972). Racemases are enzymes that act on substrates with only one asymmetric centre (interconverting enantiomers), while epimerases catalyse
stereochemical inversion on substrates with more than one asymmetric centre (interconverting diastereomers).

These enzymes can be divided into three groups based on their substrates (see Figure 1.1). The first group, where $X=\text{OH}$ and $Y=\text{carbonyl or carboxylate}$, includes the carbohydrate epimerases and α-hydroxyacid racemases. A second group in which $X=\text{NH}_2$ or $\text{NHR}'$ and $Y=\text{COOH}$, comprises the amino acid racemases and epimerases, and a third group, where $X=\text{COOH}$, aryl or alkyl and $Y=\text{acyl-CoA}$ includes the acyl-CoA racemases.

No enzymatic inversion of stereochemistry is known around stereocentres that are not hydrogen-substituted (Adams, 1976). For comprehensive reviews on epimerases and their mechanisms, see the works of Glaser (1972), Adams (1976), and Tanner & Kenyon (1998).

Inversion could, in principle, occur by breaking and reforming any of the bonds at the carbon centre; however, almost all racemases and epimerases ultimately operate by a proton-transfer pathway (Adams, 1976; Barber, 1979; Faraci & Walsh, 1988; Frey, 1987; Glaser, 1972; McDonough & Wood, 1961; Melo & Glaser, 1968). The enzyme first deprotonates the substrate and then reprotonates the resulting intermediate on the opposite face, bringing about inversion. Often, the centre is "activated" toward proton-transfer by
an adjacent carbonyl or imminium functionality, which increases the acidity of the labile proton (Faraci & Walsh, 1988; McDonough & Wood, 1961). If the centre is not already activated, an activating group can be introduced by the enzyme with the use of a NAD$^+$ or pyridoxal phosphate cofactor.

Carbohydrate Epimerases

Carbohydrates are densely functionalised, and therefore have an abundance of stereochemical configurations. In order to make more use of these molecules, a number of carbohydrate epimerases have evolved that interconvert sugars by stereochemical inversion. These epimerases can be categorized according to whether or not they use a NAD$^+$ cofactor.

The vast majority of cofactor-independent carbohydrate epimerases invert stereochemistry at a position adjacent to a carbonyl or carboxylate functionality, using a deprotonation-reprotonation mechanism.

An example of epimerisation at a centre activated by an adjacent carbonyl group is the interconversion of D-ribulose-5-phosphate and D-xylulose-5-phosphate by D-ribulose-5-phosphate 3-epimerase. This epimerisation (Figure 1.2) occurs by deprotonation at C-3 to form the enediol(ate) intermediate, followed by reprotonation on the opposite face to form the epimeric product (Adams, 1976; Glaser, 1972).
Chapter I: Stereochemical Inversion

Figure 1.2. The Mechanism of Epimerisation of D-Ribulose-5-Phosphate 3-Epimerase.

Alternatively, epimerisation may occur at an “unactivated” centre by oxidation of that centre with a cofactor such as NAD\(^+\), followed by re-reduction of the intermediate species from the opposite face (Figure 1.3). Oxidation and re-reduction occur by hydride-transfer,

Figure 1.3. The Mechanism of Epimerisation of UDP-Glucose 4-Epimerase
such as in UDP-glucose 4-epimerase. In this enzyme, the C-4 hydride is transferred to enzyme-bound NAD$^+$ to produce the intermediate UDP-4-ketopyranose. Rotation about the bond connecting the glycosyl anomeric oxygen and the UDP moiety allows the intermediate to become reoriented in the active site so that the hydride can add back to C-4 on the opposite face of the carbohydrate ring (Frey, 1987).

NAD$^+$ can also be used to activate an adjacent centre: GDP-\(\text{d-mannose}\) epimerase of *Chlorella pyrenoidosa* appears to operate in this manner (Barber, 1979). In this mechanism (Figure 1.4), the epimerase transiently oxidizes the C-4 hydroxyl group to a keto functionality. It is thought that the epimerisation then occurs through two sequential deprotonation/reprotonation events, with two enediols as intermediates. The catalytic cycle

![Figure 1.4. Proposed Mechanism of GDP-\(\text{d-mannose}\) 3,5-Epimerisation.](attachment:image.png)
Chapter I: Stereochemical Inversion

is completed by stereospecific reduction of the C-4 keto group. An epimerase in dTDP-L-
rhamnose synthetase of Pseudomonas aeruginosa also appears to share this mechanism, although
the reductase activity is found in an associated protein (Glaser, 1972; Glaser et al., 1972; Melo & Glaser, 1968; Tanner & Kenyon, 1998). This enzyme epimerises the C-3 and C-5
positions of dTDP-4-keto-6-deoxy-D-glucose.

Epimerases and racemases which appear to invert unactivated stereocentres without
the use of any cofactors and which, therefore, must employ unique reaction mechanisms are
of great interest. Bacterial UDP-N-acetylglucosamine 2-epimerase catalyses the
interconversion of UDP-N-acetylglucosamine and UDP-N-acetylmannosamine by
epimerisation at C-2, an unactivated stereocentre. Solvent isotope was incorporated into C-2
of the equilibrating epimers (Sala et al., 1996; Salo, 1976). Furthermore, a primary kinetic
isotope effect is observed upon epimerisation of [2-2H]UDP-N-acetylglucosamine (Morgan
et al., 1997). These results suggest that epimerisation occurs through removal of the proton
at C-2 followed by its replacement in an opposite stereochemical sense. This epimerisation
must occur without NAD⁺-induced transient oxidation at C-3 since addition of NAD⁺ to
the epimerase did not have an effect on its activity and a tightly-bound NAD⁺ molecule was
not found in association with the enzyme (Morgan et al., 1997). Studies on the epimerase
have provided evidence that the anomeric carbon-oxygen bond is broken (and reformed)
during the course of the reaction (Sala et al., 1996). These observations point toward a
mechanism involving trans-elimination of UDP to generate an enzyme-bound 2-
acetamidoglucal intermediate, followed by the syn-addition of UDP to the C-1 - C-2 double
bond with reprotonation of the opposite face to form UDP-N-acetylmannosamine (Figure 1.5). Further evidence for this mechanism is that on extended incubation in the presence of the enzyme, an equilibrating pool of the epimers was converted into free UDP and 2-acetamidoglucal (Morgan et al., 1997).

Figure 1.5. Proposed Mechanism for Epimerisation by UDP-N-Acetylglucosamine 2-Epimerase.

Another enzyme which operates at an "unactivated" stereocentre without the aid of a cofactor is L-ribulose-5-phosphate 4-epimerase, which plays a key role in bacterial L-arabinose utilization, interconverting L-ribulose-5-phosphate and D-xylulose-5-phosphate by epimerisation at C-4 (Figure 1.6) (Burma & Horecker, 1958b; Lee et al., 1968; Wolin et al., 1957; Wolin et al., 1958). Determining the mechanism of this enzyme will be the focus of this thesis.
**L-Ribulose-5-Phosphate 4-Epimerase**

**Role and Distribution**

L-Arabinose is an unusual sugar, having an L- rather than a D- stereoconfiguration, and is found in plant pectins, gums and complex polysaccharides. Various bacteria are capable of utilizing this sugar as an energy source through a pathway elucidated in 1958 (Burma & Horecker, 1958b; Simpson *et al.*, 1958). In this pathway (Figure 1.7), L-arabinose is first isomerized to L-ribulose by a specific isomerase. L-Ribulose is then phosphorylated with ATP by a typical Mg²⁺-requiring kinase. A key step in L-arabinose metabolism is the interconversion of L-ribulose-5-phosphate (LRu5P) and D-xylulose-5-phosphate (DXu5P) by epimerisation at C-4. DXu5P can then be shuttled into the pentose-phosphate pathway.

---

**Figure 1.6.** Interconversion of L-Ribulose-5-Phosphate and D-Xylulose-5-Phosphate.
Chapter I: Stereochemical Inversion

![Diagram showing the bacterial L-Arabinose Utilization Pathway]

**Figure 1.7.** Bacterial L-Arabinose Utilization Pathway

L-Ribulose-5-phosphate 4-epimerase has been isolated from *Klebsiella pneumoniae* (Wolin *et al.*, 1958) (formerly *Aerobacter aerogenes* (Orskov, 1984)), *Escherichia coli* (Lee *et al.*, 1968), and *Lactobacillus plantarum* (Burma & Horecker, 1958b), while its activity has been detected in a number of other bacteria including *L. pentosus* (Burma & Horecker, 1957), and *L. pentoaceticus* (Rappoport *et al.*, 1951). Furthermore, the DNA sequence of the gene encoding this epimerase, *araD*, has been determined for *E. coli* (Bonner *et al.*, 1990; Chen *et al.*, 1990; Iwasaki *et al.*, 1991; Lee *et al.*, 1986; Mineno *et al.*, 1990; Yura *et al.*, 1992), *Salmonella typhimurium* (Lin *et al.*, 1985), *Haemophilus influenzae* (Fleischmann *et al.*, 1995), *Mycoplasma pneumoniae* (Himmelreich *et al.*, 1996) and most recently, *Bacillus subtilis* (Sá-Nogueira *et al.*, 1997). Undoubtedly, with more bacterial genome sequencing, further examples will be discovered.
Previous Mechanistic Studies

An examination of the substrate’s structure clearly indicates that this enzyme does not employ a simple proton-transfer mechanism since the proton at the unactivated C-4 stereocentre is quite non-acidic. The mechanistic studies to date show that there is no appreciable kinetic isotope effect upon epimerisation of D-[4-3H]-Xu5P or of L-[4-2H]-Ru5P (this effect is expected in both proton and hydride transfer pathways), and that epimerisation does not cause exchange of either hydrogen or oxygen isotopes with solvent (Davis et al., 1972; McDonough & Wood, 1961; Salo et al., 1972). All known epimerases and racemases that operate by proton transfer pathways show either complete or some degree of proton exchange with solvent. Exchange of oxygen with solvent could occur if epimerisation was achieved through direct displacement of the C-4 hydroxyl group by solvent. Furthermore, careful studies have indicated that this enzyme does not use the cofactor NAD$^+$ to oxidize the substrate transiently (several known sugar epimerases employ this strategy). A particularly elegant experiment showed the lack of NAD$^+$ by isolating the epimerase from a culture of a K. pneumoniae nicotinic acid auxotroph which was supplemented with $^{14}$C-labelled nicotinic acid. No radioactivity was associated with the purified enzyme, which was fully active (Deupree & Wood, 1970).

The epimerase was, however, found to require divalent metal ions for activity (Deupree & Wood, 1972). The epimerase was completely inhibited by treatment with EDTA. Activity of the resultant apoenzyme could be restored to various extents by addition
of divalent metal ions to the assay cuvettes, in the order Mn$^{2+}$ > Co$^{2+}$ > Ni$^{2+}$ > Ca$^{2+}$ > Zn$^{2+}$ > Mg$^{2+}$ (greatest to least degree of reactivation).

**Proposed Mechanisms**

This enzyme is unique among epimerases in that it presumably operates without breaking the carbon-hydrogen bond at the stereolabile position of the sugar. Two alternative pathways (see Figure 1.8) involving carbon-oxygen or carbon-carbon bond breakage and reformation are consistent with the present mechanistic information. It was postulated (Deupree & Wood, 1972) that the reaction catalysed by L-ribulose-5-phosphate 4-epimerase may proceed via a dehydration/rehydration mechanism with a sequestered water molecule (Path A) or via a retroaldol/aldol mechanism involving carbon-carbon bond breakage (Path B).

Path A is an elimination/readdition mechanism which proceeds with a sequestered water molecule. The first step in this reaction would be a base-catalysed removal of the C-3 proton, which is activated by the adjacent carbonyl group at C-2. This would result in a carbanion at C-3 that could eliminate the C-4 hydroxyl group to form a double bond between C-3 and C-4. It is not clear how the enzyme would achieve the *syn* elimination which must occur in one direction, and the *anti* elimination in the other direction, without a major conformational change. The presence of a divalent metal ion in the active site could facilitate removal of hydroxide and prevent its exchange with bulk solvent. The enone
Figure 1.8. Proposed Pathways for L-Ribulose-5-Phosphate 4-Epimerase Consistent with Observations.
Chapter I: Stereochemical Inversion

intermediate species would be transformed back to a ketopentose by addition of the hydroxyl
group to C-4 on the opposite face of the double bond from which it was removed.

Path B is an unprecedented means of achieving stereochemical inversion. In this
retroaldol/aldol mechanism, deprotonation of the C-4 hydroxyl group leads to breakage of
the C-3 – C-4 bond and to formation of the intermediates glycolaldehyde-2-phosphate and
an enediolate anion. The carbonyl group of the substrate may be chelated to the metal ion
and facilitate enolate formation. The carbonyl group of the aldehyde intermediate must
rotate in the active site before the same face of the enediolate adds back to the opposite face
of the aldehydic carbonyl group in an aldol addition. Thereby the bond between C-3 and
C-4 is reformed with the geometry differing only at C-4 in the product.

A third possibility, not shown in Figure 1.8, involves a series of proton transfer steps
that result in the isomerization of the substrate to a C-3 ketose. This carbonyl group at C-3
would acidify the proton at C-4 and allow a deprotonation/reprotonation event to occur,
inverting the stereochemistry at this position. Finally, the product would be isomerized back
to a C-2 ketose. This seems unlikely because all of the steps of this mechanism would have
to occur without incorporation of solvent hydrogen atoms.

Relationship with Other Enzymes

An enzyme which may have a similar mechanism to L-ribulose-5-phosphate 4-
epimerase is dihydroneopterin triphosphate 2'-epimerase, which interconverts D-erythro-
dihydroneopterin triphosphate to L-threo-dihydromonapterin triphosphate (Figure 1.9). This
epimerase also acts at an "unactivated" centre and has a requirement for divalent metal ions, 
$\text{Mn}^{2+}$ and $\text{Mg}^{2+}$ being the most effective (Heine & Brown, 1975). This epimerase shares 
amino acid similarity to a putative dihydroneopterin aldolase domain but lacks aldolase 
activity (Haussmann et al., 1997).

![Figure 1.9. Reaction Catalysed by $\text{D-erythro}$-Dihydroneopterin Triphosphate 2'-Epimerase.](image)

Recently, there was an observation of significant sequence homology between $\text{L}$-ribulose-5-
phosphate 4-epimerase and bacterial $\text{L}$-fuculose-1-phosphate aldolase (a member of a group 
of enzymes which catalyse metal-dependent carbon-carbon bond breaking reactions) (Dreyer 
& Schulz, 1993; Dreyer & Schulz, 1996b). It seems likely that the epimerase has 
evolutionarily diverged from the aldolases. For this reason we favour mechanistic Path B,
in which the enzyme generates a metal-bound enolate and instead of protonating and releasing it, adds it back to either face of the bound glycolaldehyde phosphate.

Aldolases

Aldolases are enzymes which catalyse aldol cleavages (or aldol condensations in the reverse direction) (Horecker et al., 1972; Morse & Horecker, 1968) yielding an aldehyde and a carbonyl-containing molecule for which the enzyme has very high specificity (Feingold & Hoffee, 1972; Gijsen et al., 1996; Wood, 1972). Aldol cleavage reactions involve the formation of a resonance-stabilized enolate intermediate. Aldolases catalyse this reaction by enhancing the stability of the enolate structure either through Schiff base formation or by use of a divalent metal ion which coordinates to the carbonyl oxygen atom. These two classes of aldolase may represent analogous proteins which have presumably arisen by convergent evolution (Rutter, 1964; Rutter, 1965).

Aldolases are currently of interest to synthetic chemists as catalysts, especially because of their ability to form enantiomerically pure vicinal diols (Gijsen et al., 1996).
Class I Aldolases

Class I aldolases are found in plants and animals and use a Schiff base mechanism for catalysis. This group of enzymes, in particular fructose-bisphosphate aldolase (often referred to as "aldolase") are well-studied and have been extensively reviewed (Horecker et al., 1972; Morse & Horecker, 1968). This class of aldolase does not require divalent metal ions for catalysis and is not inhibited by treatment with EDTA or other metal chelator. They are, however, inhibited by reduction with NaBH₄ in the presence of substrate. In these enzymes, the ε-amino group of an active site lysine residue reacts with the carbonyl group of the substrate to form an imminium cation or protonated Schiff base. This is illustrated in Figure 1.10 using fructose-1,6-diphosphate aldolase as an example. In this aldolase, base-catalysed deprotonation of the C-4 hydroxyl group leads to formation of a carbonyl group at C-4 and cleavage of the carbon-carbon bond between C-3 and C-4. The aldehydic moiety,
glyceraldehyde-3 phosphate (GAP) is then released from the enzyme’s active site. The reaction is completed by protonation of the dihydroxyacetone phosphate (DHAP) anion by an active site proton donor, followed by hydrolysis of the Schiff base and release of DHAP.

Class II Aldolases

Class II aldolases, found in microorganisms, use a divalent metal ion to polarize the carbonyl oxygen of the substrate and stabilize the enolate intermediate in the reaction. These enzymes are not inactivated by reduction with NaBH$_4$ in the presence or absence of substrate. They are, however, inhibited by metal chelating agents such as EDTA, indicating their requirement for a divalent metal ion for catalysis. This metal ion is most often a tightly-bound Zn$^{2+}$ ion.

In the catalytic reaction, the divalent metal ion coordinates to and polarizes the carbonyl oxygen of DHAP or the substrate, and may also orient the C-1-phosphate group. In this mechanism (Figure 1.11), base-catalysed deprotonation of the C-4 hydroxyl group leads to formation of a carbonyl group at C-4 and cleavage of the bond between C-3 and C-4.

Figure 1.11. Mechanism of Class II Aldolases.
in a retro-aldol reaction. Polarization of the carbonyl bond at C-2 by Zn$^{2+}$ facilitates formation of a metal-bound enolate anion. The aldehydic product is then released from the enzyme’s active site. The reaction is completed by protonation of the DHAP anion by an active site base, thereby regenerating active enzyme in a protonation state suitable for the retro-aldol reaction direction.

Crystallographic work on two DHAP-dependent Class II aldolases, fructose-bisphosphate aldolase (Cooper et al., 1996) and fuculose-1-phosphate aldolase (Dreyer & Schulz, 1993; Dreyer & Schulz, 1996b), indicates that the DHAP-dependent Class II aldolases should be further divided into two groups based on their structure – those which are homodimers and have two Zn$^{2+}$ ions per subunit (eg. fructose-bisphosphate aldolase), and those which are homotetramers with one Zn$^{2+}$ ion per active site (eg. fuculose-1-phosphate aldolase) (Cooper et al., 1996). There is little amino acid sequence similarity between these two groups of Class II aldolases.

Recent work on L-fuculose-1-phosphate aldolase has led to the proposal of a new mechanism. This was the result of the observation of the structure of this aldolase with the bound inhibitor phosphoglycolohydroxamate (Figure 1.12). As a potent inhibitor ($K_i = 5 \mu$M), this molecule mimics the enediolate intermediate or perhaps a transition state of the Class II aldolases. Phosphoglycolohydroxamate forms a bidentate chelate in which both the enol and hydroxamate ligands are coordinated to the Zn$^{2+}$ ion (Dreyer & Schulz, 1996a).
In the new mechanism for Class II aldolases (Figure 1.13), polarization of the carbonyl bond increases the acidity of the protons on C-3 of DHAP, one of which is abstracted by a general base, most likely the carboxylate group of E73, resulting in a metal-bound enediolate intermediate. Nucleophilic attack by this species on the carbonyl group of the incoming L-lactaldehyde forms a new carbon-carbon bond. The aldehydic oxygen atom subsequently accepts a proton from a tyrosine hydroxyl group. This residue is thought to help stabilize the developing negative charge and to control the stereochemistry of the nucleophilic attack by hydrogen bonding to the aldehyde (Fessner et al., 1996).
Although aldolases are noted for their exceptional stereospecificity, Class II tagatose-1,6-diphosphate aldolase occasionally binds the aldehydic moiety inversely, and condenses it to DHAP (see Figure 1.14). Conversely, this enzyme is able to cleave the epimeric sugar, fuculose-1,6-diphosphate, although at a much slower rate (1%) than it cleaves its normal substrate, tagatose-1,6-diphosphate. The net effect is an overall epimerization (Fessner & Eyrisch, 1992).

![Figure 1.14. Epimerisation of D-Tagatose-1,6-diphosphate Catalysed by D-Tagatose-1,6-Diphosphate Aldolase.](image)

L-Fuculose-1-phosphate aldolase, which has amino acid homology (34% identity) (Dreyer & Schulz, 1996b) to L-ribulose-5-phosphate 4-epimerase, is a metalloenzyme (Ghalambor & Heath, 1962; Ghalambor & Heath, 1966) and is the third enzyme in the bacterial L-fucose metabolic pathway (Figure 1.15) (Chen et al., 1987). Rhamnulose-1-phosphate aldolase and fuculose-1-phosphate aldolase, despite the similarity between the reactions they catalyse and their roles in the parallel metabolic pathways (see Figure 1.15) for bacterial rhamnose and fucose utilization, have only a small amount of homology in their
amino acid sequences. Together with some additional genetic evidence, these two aldolases are thought to have evolved convergently (Moralejo et al., 1993).

Figure 1.15. Parallel Pathways of L-Fucose and L-Rhamnose Dissimilation. These sugars are degraded to dihydroxyacetone phosphate and L-lactaldehyde, which are further metabolized. The enzymes in each pathway are L-fucose isomerase (FucI), L-fuculose kinase (FucK), L-fuculose-1-phosphate aldolase (FucA), L-rhamnose isomerase (RhaA), L-rhamnulose kinase (RhaB), and L-rhamnulose-1-phosphate aldolase (RhaD).
Summary

Enzymes are among the most efficient catalysts known. Therefore, an understanding of their mechanisms can provide valuable insights into the nature of chemical catalysis and shed light on the course of natural evolution. This thesis strives to discover the nature of the mechanism used by L-ribulose-5-phosphate 4-epimerase, whether it follows a dehydration/rehydration pathway or a retroaldol/aldol mechanism like that used by L-fuculose-1-phosphate aldolase. This work necessitates a plentiful source of the epimerase. Therefore, L-ribulose-5-phosphate 4-epimerase was cloned from *Escherichia coli* (Chapter II) and the recombinant enzyme was shown to have comparable properties to the naturally abundant *E. coli* enzyme. The epimerase has been shown to be dependent on metal ions for activity (Deupree & Wood, 1972) and have amino acid sequence similarity to L-fuculose-1-phosphate aldolase (Dreyer & Schulz, 1993; Dreyer & Schulz, 1996b). The metal ion ligands of the epimerase are determined (Chapter III) and are shown to be the same as those in the aldolase. After establishing a structural link between the aldolase and L-ribulose-5-phosphate 4-epimerase, it was thought that these two enzymes likely share a mechanistic strategy. Therefore, the epimerase and mutant epimerases were assayed for aldolase activity (Chapter IV). Low levels of activity were observed, supporting the notion that this enzyme utilizes a retroaldol/aldol mechanism. A number of possible future experiments are described in Chapter V which will provide further evidence for a retroaldol/aldol mechanism, identify
Chapter I: Stereochemical Inversion

the catalytic residues in the epimerase and show the differences in the active sites of the epimerase and the aldolase.
Chapter II

Cloning, Overexpression and Characterization of

L-Ribulose-5-Phosphate 4-Epimerase from *Escherichia coli*

Introduction

Overexpression Systems

A vital component in any study of a protein or enzyme is a source of homogeneous, soluble, and active material, preferably in milligram quantities. Many small proteins or peptides may be purified from their normal biological source, or, if small enough, peptide synthesis may be feasible. Frequently, however, overexpression in a suitable biological host system is required to obtain sufficient quantities of material to carry out the planned studies.

The term overexpression applies to a technique in which the DNA encoding a protein is introduced into bacteria and the amount of this protein that is produced is significantly higher than that of a typical endogenous protein. The theoretical maximum
yield of a protein from one litre of an *Escherichia coli* culture at $10^9$ cells per mL is 75 mg, or 50% of the total protein (New England Biolabs, 1997 catalogue).

Besides providing larger quantities of the desired protein, use of an overexpression system often provides a more economical means of obtaining the protein. Many natural sources contain only small quantities of the desired protein, and much time is spent purifying it from the necessarily large quantities of source material.

A variety of overexpression systems are available to suit various needs: bacterial (usually *E. coli*), yeast, insect cell and mammalian (tissue culture) are currently in use. The molecular biology of *E. coli* is well understood, and consequently, many overexpression systems are available for this host. In addition, *E. coli* is inexpensive to grow, and it grows quickly. Methods leading to the optimization of these systems are numerous and generally exploit every aspect of protein biosynthesis. A major disadvantage with the use of bacterial overexpression systems is their inability to perform many of the post-translational modifications found in eukaryotic proteins.

In order to achieve overexpression, the gene encoding the protein of interest must first be cloned into an overexpression vector. A vector is a small piece of closed circular extrachromosomal, double stranded DNA and normally confers on the host cell antibiotic resistance which helps in vector selection and propagation. Vectors, have an origin of replication which determines the number of copies of the plasmid in the cell and which is where replication begins. Once the gene of interest is inserted into the vector, the resulting construct is called a plasmid. Not all vectors can be used for protein expression. Besides
the elements found on cloning vectors, overexpression vectors also have a promoter, a ribosomal binding site (RBS), a Shine-Dalgarno sequence (SD) and a translation terminator. Many of these features have been optimized in commercially available overexpression vectors.

The first step in protein expression is transcription of DNA into RNA (by RNA polymerase); the resulting transcript is called mRNA. A promoter is a sequence of DNA which enhances transcription of the mRNA for a gene. The promoter is positioned approximately 10 to 100 base pairs upstream of the ribosomal binding site and is normally under the control of a regulatory gene. This gene may be either on the plasmid itself or integrated into the host chromosome. *E. coli* promoters consist of a hexanucleotide sequence located about 35 base pairs upstream of the transcription initiation codon (the "-35 region") which is separated from another hexanucleotide sequence (the "-10 region") by a short nucleotide spacer. Promoters useful for overexpression should be sufficiently strong that at least 10% of the total soluble cellular protein is from overexpression. If the protein to be overexpressed is toxic to the host cell, it is essential that the control over gene transcription and translation be tight.

The second step in protein biosynthesis is translation of the mRNA message into a polypeptide by ribosomes. The ribosomal binding site is a region of DNA sequence downstream from the promoter (at the 5' end of the mRNA) and spans up to approximately 54 nucleotides between positions -35 and +22 of the mRNA coding sequence. The Shine-Dalgarno (SD) is a sequence which interacts with the complementary 3' end of the 16S
rRNA (part of the ribosome) during translation initiation. Spacing between the SD and the start codon varies from 5 to 13 nucleotides, and influences the efficiency of translation initiation. The sequence of the spacer should eliminate the potential of any secondary-structure formation in the mRNA transcript, as these structures can reduce the efficiency of translation.

A stop codon at the 3'-end of the coding sequence is the signal for termination of translation. A number of stem-loop structures in the mRNA transcript (of the DNA) located downstream of the stop codon protect the mRNA from exonucleolytic degradation and extend the mRNA half-life. An excellent review by Makrides (1996) on overexpression in *E. coli* provides more detail on considerations in choosing and designing overexpression vectors for this host.

The overexpression vector into which the *araD* gene was cloned is a modified version of the pBS vector available from Stratagene. Modification of this vector in the laboratory of Jeremy Knowles (Hermes *et al.*, 1990) removed a portion of the multiple cloning site and rendered the regulatory genes *lacI* and *lacZ* non-functional. In addition, the T7 promoter sequence was replaced with a trc promoter. This promoter is a hybrid between the lac and tac promoters and is much stronger than either of them (Amann & Brosius, 1985; Amann *et al.*, 1983). This plasmid requires no induction because it is not repressed, and was found to express triosephosphate isomerase (TIM) at roughly 100 mg of protein per litre of cell culture (Hermes *et al.*, 1990). This level of overexpression appears to be higher than the theoretical yield, but could arise from cultures which were grown to higher cell density.
Because of the high levels of overexpression achieved with pBSTIM, *araD* was cloned into this vector in place of *tpi*, the gene encoding TIM.

**The *araBAD* Operon**

Mammals neither metabolize nor intestinally absorb the plant sugar L-arabinose. Therefore, the bacteria which inhabit the intestine are periodically presented with a feast of this pentose. The *araBAD* operon is a transcriptional unit which encodes the three structural genes for bacterial L-arabinose dissimilation (Figure 2.1). The operon can be transcribed (i.e. mRNA is formed) under conditions of low glucose concentration and in the presence of L-arabinose (Lee & Bendet, 1967; Lee et al., 1986; Sá-Nogueira *et al.*, 1997). Even when grown in minimal medium and fully induced with L-arabinose, these enzymes, which are present in approximately equimolar amounts (Lin *et al.*, 1985), each comprise only three to four percent of the total cellular protein (Lee & Bendet, 1967). Therefore, in order to obtain large quantities of any of these enzymes, and to allow for future site-directed mutagenesis studies, their genes must be cloned into an overexpression vector.
After entering the cell, L-arabinose is sequentially converted to L-ribulose, L-ribulose-5-phosphate, and D-xylulose-5-phosphate by the action of the three structural gene products (see Figure 2.2): L-arabinose isomerase (AraA, EC 5.3.1.4), L-ribulokinase (AraB, EC 2.7.1.16), and L-ribulose-5-phosphate 4-epimerase (AraD, EC 5.1.3.4), respectively (Burma & Horecker, 1957; Lee et al., 1986; Sá-Nogueira et al., 1997; Simpson et al., 1958). D-Xylulose-5-phosphate is further metabolized through the pentose phosphate pathway (Heath et al., 1958; Rappoport et al., 1951; Sá-Nogueira et al., 1997; Simpson et al., 1958). The two enzymes of particular interest to this study are L-ribulose-5-phosphate 4-epimerase, and L-ribulokinase.
L-Ribulose-5-Phosphate 4-Epimerase: An Enigmatic Reaction Mechanism

This epimerase is of particular interest because, unlike the vast majority of epimerases and racemases, this enzyme appears to invert the stereochemistry (Figure 2.2) at an unactivated centre without the aid of a cofactor. Therefore, this enzyme must use a unique mechanism which does not involve direct deprotonation of the non-acidic proton at the C-4 stereocentre.

Several studies were performed on this enzyme, but a clear description of the mechanism remains elusive. The enzyme was definitively demonstrated not to use a NAD$^+$ cofactor (Deupree & Wood, 1970). Furthermore, experiments using isotopically-labelled LRu5P, DXu5P or solvent (Davis et al., 1972; McDonough & Wood, 1961; Salo et al., 1972) showed that epimerisation appears to occur without breaking the carbon-hydrogen bond between C-4 and C-3 (see Chapter I for details).
A divalent metal ion was required for activity. The naturally occurring metal ion was not determined; however, in the presence of at least 1 mM EDTA, the enzyme was rendered inactive. Various degrees of activity could be restored by addition of divalent metal ions (1 μM to 10 mM) to the assay cuvettes in the order Mn$^{2+}$ > Co$^{2+}$ > Ni$^{2+}$ > Ca$^{2+}$ > Zn$^{2+}$ > Mg$^{2+}$ (greatest to least degree of reactivation).

These studies led to the proposal of two possible reaction mechanisms in which either the carbon-oxygen bond or a carbon-carbon bond at C-4 was broken and then reformed since the epimerase appears not to break the carbon-hydrogen bond at C-4 (Deupree & Wood, 1972). These mechanisms are discussed and illustrated in Chapter I.

The native enzyme was found to have a molecular mass of about 110 kDa (Deupree & Wood, 1975; Gielow & Lee, 1975) as determined by high speed sedimentation equilibrium analysis and, until the gene was sequenced, was thought to have three subunits of about 35 kDa each (Gielow & Lee, 1975). After the araD gene was sequenced and analysed, the enzyme was determined to have four subunits of about 25.5 kDa, as calculated from the predicted amino acid sequence of the epimerase from *E. coli* (Lee et al., 1986; Mineno et al., 1990), *Bacillus subtilis* (Sá-Nogueira et al., 1997), and *Salmonella typhimurium* (Lin et al., 1985, corrected by A. Bairoch, unpublished data, Swiss Protein Data Bank, 1995).

Kinetic parameters were determined for the enzyme, and pH rate profiles established that it is most active between pH 7 to 9 (Burma & Horecker, 1958b; Lee et al., 1968; Wolin et al., 1958).
L-Ribulokinase: A Means of Producing L-Ribulose-5-Phosphate

L-Ribulokinase is a typical magnesium-requiring kinase (Burma & Horecker, 1958a; Lee & Bendet, 1967; Lee & Englesberg, 1966; Simpson & Wood, 1956; Simpson & Wood, 1958; Wolin et al., 1958) which phosphorylates L-ribulose with adenosine triphosphate (ATP) to form L-ribulose-5-phosphate and adenosine diphosphate (ADP) (Figure 2.3). Like the epimerase, this enzyme has been purified from a number of bacterial sources, including E. coli (Lee & Bendet, 1967; Lee & Englesberg, 1966), Klebsiella pneumoniae (formerly Aerobacter aerogenes) (Simpson & Wood, 1958), and Lactobacillus plantarum (Burma & Horecker, 1958a).

![Figure 2.3. Phosphorylation of L-Ribulose by L-Ribulokinase Provides L-Ribulose-5-Phosphate, the Substrate for L-Ribulose-5-Phosphate 4-Epimerase.](image)

Since the substrate for L-ribulose-5-phosphate 4-epimerase is not commercially available, it was necessary to prepare the substrate ourselves. Using L-ribulokinase to phosphorylate L-ribulose is a convenient method of preparing this phosphopentoketose. Although the kinase has been purified to homogeneity, this is not necessary for production of L-ribulose-5-phosphate. According to the method of R.L. Anderson (1966), L-
ribulokinase may be partially purified from an L-arabinose induced strain of bacteria which cannot produce active L-ribulose-5-phosphate 4-epimerase. When the bacteria are grown on minimal media in the presence of arabinose the levels of expression of the araBAD gene products increase so that the bacteria may utilize this sugar as an energy source. This crude enzyme preparation can then be used as a catalyst for phosphorylation of L-ribulose with ATP. In the absence of the epimerase, arabinose can proceed only to L\(\text{Ru5P}\), which accumulates. It is important to maintain the pH of this reaction mixture between 6 and 7.5; the enzyme is unstable below pH 6 and above pH 7.5 (Burma & Horecker, 1958a) and the maximal activity lies around pH 7.6 (Lee & Bendet, 1967).

Even when fully induced, this enzyme is estimated to comprise only 3 or 4% of the total cellular protein (Lee & Bendet, 1967). For our purposes this was sufficient, although it would be more convenient to have this enzyme overexpressed. Besides phosphorylating L-ribulose, L-ribulokinase can phosphorylate its enantiomer, D-ribulose (Burma & Horecker, 1958a; Lee & Bendet, 1967; Lee & Englesberg, 1966; Simpson & Wood, 1958), as well as adonitol (also known as ribitol) and L-arabitol (Simpson & Wood, 1958).
Results and Discussion

Preparation of L-Ribulose-5-Phosphate

L-Ribulokinase was initially obtained from L-arabinose induced *E. coli* Y1090 (modified) (see the experimental section for details), a strain which lacks a functional L-ribulose-5-phosphate 4-epimerase. Initial difficulties in preparing L-ribulokinase for phosphorylation were overcome by maintaining the temperature of the cell paste and the crude cell lysate between 10 and 4°C. Although the purified kinase is reported to be heat-stable (Lee & Bendet, 1967; Lee & Englesberg, 1966), in the cell lysate it was much less stable, perhaps due to proteolytic degradation. After a number of successful preparations, we were unable to obtain L-ribulokinase activity from this strain, possibly as a result of a spontaneous mutation or from movement of the transposon (Tn10) in the genome. At this point we tried cloning the L-ribulokinase, but our attempts have thus far been unsuccessful. Ultimately we found that another 4-epimerase-less strain, *E. coli* MC4100, worked very well as a stable source of L-ribulokinase when used in Anderson’s procedure.

L-Ribulose-5-phosphate was purified by anion exchange chromatography and the fractions greater than 90% pure (by 1H-NMR spectroscopy; see Figure 2.4) from a number of preparations were pooled and lyophilized.

The concentration of a stock solution of the substrate was accurately determined by using a sample of known dilution to initiate a cuvette (incubated at 37°C) containing all the
normal assay components except LRu5P. A large excess of enzyme was used to epimerise the sugar rapidly.

**Attempted Cloning of L-Ribulokinase**

The difficulties in preparing active L-ribulokinase from *E. coli* Y1090 (modified) led us to try cloning this enzyme for overexpression. The *araB* gene of *E. coli* DH5αF', encoding L-ribulokinase (EC 2.7.1.16) was amplified directly by PCR (40 cycles) on whole cells. The PCR product, when thoroughly digested by *Nde* I proved to have a restriction site for this enzyme at nt 827 giving fragments ca. 827 nt and ca. 874 nt long. It was thought that partial digestion by *Nde* I might yield a product which was digested at the 5' end of the gene (where this site had been introduced on the primers) but not in the middle. The products of partial *Nde* I digestion were separated on an agarose gel and the large (ca. 1701 nt) fragment isolated. Ligation of this fragment into a pET-11a vector prepared for cloning by *Nde* I and *BamH* I digestion was not successful. By this time, preparation of LRu5P using L-ribulokinase isolated from *E. coli* MC4100 had proven successful, so further attempts at cloning L-ribulokinase were abandoned.
Figure 2.4. Proton-NMR Spectrum of Purified L-Ribulose-5-Phosphate (free acid) at 500 MHz (in D$_2$O).
A superior cloning strategy for this gene would be to introduce a \textit{Nco} I site at the 5'-end of the gene during PCR, and then to clone this fragment into a compatible vector. \textit{Nde} I is the preferred restriction site at the 5'-end of genes cloned into overexpression vectors because it contains ATG, the start codon, as the second half of its six-base recognition sequence (CATATG). Many overexpression vectors which use this site have optimized distances and nucleotide sequences between the "-10" site and the start codon for maximal protein expression. \textit{Nco} I can also be used in this manner, as its recognition sequence (CCATGG) also contains ATG. However, use of this restriction enzyme for cloning the 5'-end of the gene into an overexpression vector restricts the second amino acid of the protein sequence to one of five residues: V, A, D, E, or G. The second residue in the L-ribulokinase sequence happens to be alanine.

\textit{Cloning, Expression, Purification and Characterization of the L-Ribulose-5-Phosphate 4-Epimerase of Escherichia coli}

The \textit{araD} gene encoding L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4) was amplified directly from whole cells of \textit{E. coli} DH1 using 40 cycles of PCR. The PCR product was cloned into the modified pBS vector described earlier, creating the new construct, pRE1 (this work was performed by Dr. M. Tanner).

Prior to sequencing the \textit{araD} gene contained on pRE1, we were concerned that there may be some PCR-derived errors in this gene because of the large number of PCR cycles used. The number of cycles necessary can be reduced by using a plasmid as a template as
opposed to genomic DNA when using whole cells. We obtained a plasmid, pHC5, from the laboratory of Dr. R.E. Moses which contained a section of the *E. coli* genome including the complete *araD* gene (Chen *et al.*, 1990). We planned to use this plasmid as the template for PCR amplification of the *araD* gene and to reclone it into the modified pBS vector. While the *araD* gene contained on the plasmid pHC5 was successfully amplified by PCR (30 cycles), this PCR product was not successfully cloned into the same modified pBS vector used in formation of pRE1. These difficulties most likely arose during isolation of DNA from agarose or during the ligation phase of cloning. Ultimately, the portion of plasmid pRE1 containing the *araD* gene was sequenced and was shown by sequencing not to contain any PCR-derived errors.

In order to avoid contamination by endogenous enzyme, pRE1 was transformed into a strain of *E. coli* which lacks a functional LRu5P 4-epimerase. For this purpose, *E. coli* Y1090 (modified) was used, although *E. coli* MC4100 should work equally well. High levels of overexpression were observed (as shown in Figure 2.5), and were estimated to account for at least 60% of the total soluble protein. A cell paste containing the highly overexpressed epimerase was

![Figure 2.5](image.png)

**Figure 2.5.** SDS-PAGE of Crude and Purified Protein. The protein components of crude and purified 4-epimerase samples were separated by 12% SDS-PAGE. The molecular weight standards were: 66 kDa, BSA; 29 kDa, carbonic anhydrase; and 18.4 kDa, β-lactoglobulin.
lysed by passage through a chilled French Press cell and clarified by ultracentrifugation. Highly negatively-charged impurities were removed from the clarified lysate by passage through a DE-52 column, from which the epimerase was eluted using 0.4 M NaCl in 10 mM potassium phosphate buffer (pH 7.6, containing 10% glycerol). The epimerase was then purified from other proteins by ion-exchange HPLC on a Q-Pak column with a linear gradient of 0 to 0.4 M NaCl in 10 mM potassium phosphate buffer (pH 7.6, containing 10% glycerol). The epimerase, which eluted at about 0.3 M NaCl, was collected as a highly concentrated fraction. The purity was estimated to be at least 95% by visual examination of a Coomassie Blue stained SDS-polyacrylamide gel (12%) (Figure 2.5). At least 100 mg of purified enzyme were obtained from one litre of bacterial culture. This level of overexpression corresponds to that observed for TIM in this vector (Hermes et al., 1990).

The modified pBS vector used in this work previously had its site for transcriptional repression deleted, which means that transcription of the gene of interest on the vector is effectively always “on”. We found that overnight broth cultures when used as the inoculum for large cultures were incapable of overexpressing the epimerase. In order to achieve the high levels of overexpression of which this vector is capable, the inoculum for cultures used to prepare the epimerase was a colony. The reason for this is not understood.

The overexpressed enzyme has electrophoretic mobility corresponding to about 31 kDa, similar to that observed for the recombinant enzyme from S. typhimurium (Lin et al., 1985). The expected mass is 25 kDa, and one would expect an electrophoretic mobility which agrees with this. The lower mobility might be caused by the epimerase having more
basic residues than an average protein. Based on translations of the genetic code, amino acid sequences can accurately be predicted. These sequences are used to calculated subunit molecular weights based on amino acid composition. The predicted subunit masses can then be verified by electrospray ionization mass spectrometry (ESIMS) which has been shown to be (Ashton et al., 1994) an accurate method for determining protein subunit mass. The subunit mass of the recombinant epimerase was determined in this fashion and was consistent with the mass calculated from the expected amino acid sequence (calculated: 25520 Da; found: 25522 ± 4 Da). The entire enzyme, being a homotetramer, has a mass four times this value. Table 2.1 lists the subunit masses of LRu5P 4-epimerases from various sources.

Protein concentration was required to be accurate for determination of the zinc-content of the zinc-reconstituted enzyme. Therefore, the extinction coefficient of the purified enzyme at 280 nm was determined first by the method of Gill & von Hippel (1989), using the relationship,

\[ \epsilon_n = \frac{A_n}{A_d} \epsilon_d \]  

(2.1)

where \( \epsilon_n \) = the extinction coefficient at 280 nm of the native protein,

\( A_n \) = the absorbance at 280 nm of the native protein,

\( A_d \) = the absorbance at 280 nm of the denatured protein, and

\( \epsilon_d \) = the extinction coefficient at 280 nm of the denatured protein.
Chapter II: Cloning, Overexpression and Characterization

Table 2.1. Predicted Subunit Molecular Masses of L-Ribulose-5-Phosphate 4-Epimerase.

<table>
<thead>
<tr>
<th>Source Organism</th>
<th>Subunit Length</th>
<th>Predicted Subunit Mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> K12</td>
<td>231</td>
<td>25520</td>
</tr>
<tr>
<td><em>E. coli</em> B/r</td>
<td>231</td>
<td>25504 f</td>
</tr>
<tr>
<td><em>S. typhimurium</em></td>
<td>248 f</td>
<td>27059</td>
</tr>
<tr>
<td><em>B. subtilis</em></td>
<td>229</td>
<td>ca. 25700</td>
</tr>
<tr>
<td><em>E. coli</em> DH1</td>
<td>231</td>
<td>25520</td>
</tr>
</tbody>
</table>

* Mineno et al., 1990.
* Lee et al., 1986.
* Lin et al., 1985.
* Sá-Nogueira et al., 1997.
* This study.
f There is a probable frame-shift mutation here, likely from an error in DNA sequencing (Lee et al., 1986; and A. Bairoch, unpublished observation, Swiss Protein Data Bank). Bairoch predicts 231 amino acids with a mass of 25540 Da.
f There are three amino acid substitutions in this sequence compared to the *E. coli* K12 sequence: I instead of V at residue 50, A rather than T at residue 70, and E rather than D at position 216. Even with these substitutions, the calculated mass should be 25518.

$\varepsilon_0$ was calculated from the number of Trp, Tyr residues and disulphide bonds in the protein as described by Edelhoch (1967):

$$4 \text{Trp} + 9 \text{Tyr} = 4(5690 \text{ cm}^{-1} \text{ M}^{-1}) + 9(1280 \text{ cm}^{-1} \text{ M}^{-1}) = 34280 \text{ cm}^{-1} \text{ M}^{-1}.$$  

Protein at several concentrations was mixed with an equal volume of either water or guanidinium-HCl, and the absorbance of each measured at 280 nm. Using this method, the native epimerase was calculated to have an extinction coefficient of $\varepsilon = 34314 \text{ cm}^{-1} \text{ M}^{-1}$ (where M$^{-1}$ refers to the subunit concentration), or 1.76 mL mg$^{-1}$cm$^{-1}$ as the average of three determinations, which differed from that given by Lee, Patrick & Masson (1968) who
reported 1.57 mL mg\(^{-1}\) based on the dry weight of the enzyme. This method can have errors of up to 30% (Gill & von Hippel, 1989), which we deemed too high. As a result the extinction coefficient was also calculated by total amino acid analysis on a pure protein solution of known absorbance to which 0.1 mM norleucine (an unnatural amino acid) had been added as an internal standard. The relative amounts of alanine and norleucine were used to determine the concentration of the protein sample (the enzyme contains 21 alanine residues per subunit). This resulted in the extinction coefficient \(\varepsilon = 1.73 \text{ mL mg}^{-1} \text{ cm}^{-1}\) (33717 cm\(^{-1}\) M\(^{-1}\) per subunit) for the epimerase. This value was used for all determinations of protein concentration measured at 280 nm. It is of interest to note that a commonly used colorimetric technique (Bradford, 1976) calibrated to BSA gave values two to three times greater than those calculated from direct measurement of the \(A_{280}\). In this method, the protein solution is mixed with an acidic solution of methanol and Commassie Brilliant blue. After a five minute incubation, the absorbance at 595 nm is measured, and a calibration curve using BSA as a standard is used to interpolate the protein concentration of the standard. BSA is often quoted in the literature as a standard, but this protein has been observed to develop a deeper colour than other proteins (Bio-Rad 1996 catalogue), and is therefore not the best standard to use. The epimerase must then develop an even deeper colour than BSA; perhaps BSA has a lower proportion of basic and aromatic residues than the epimerase.

The purified epimerase was assayed in the L-ribulose-5-phosphate (LRu5P) to D-xylulose-5-phosphate (DXu5P) direction using the assay described by Davis \textit{et al.} (1972). In
this assay (see Figure 2.6), the epimerase converts LRu5P to DXu5P. DXu5P and ribose-5-phosphate (contained in the assay mixture) are transformed into sedoheptulose-7-phosphate (S7P) and glyceraldehyde-3-phosphate by the action of transketolase. Triosephosphate isomerase (TIM) converts glyceraldehyde-3-phosphate into dihydroxyacetone phosphate (DHAP), which is then reduced with NADH by the action of α-glycerophosphate dehydrogenase (αGDH). The oxidation of NADH to NAD$^{+}$ is followed as a decrease in absorbance at 340 nm. A small epimerase-independent background rate was observed and was determined for each cuvette individually. This rate was also observed by Davis et al.
(1972) and by Deupree and Wood (1975). The background rate was seen in the absence of both LRu5P and 4-epimerase, resulting perhaps from impurities in the coupling enzymes and reagents. This rate was subtracted from the rate of NAD$^+$ formation after addition of the epimerase. An alternate assay (GAPDH assay; Figure 2.7) in which glyceraldehyde-3-phosphate is oxidized to 1-arseno-3-phosphoglycerate (an unstable molecule which decomposes rapidly) by the action of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) could also be used. This assay, which exhibited a very high background rate and was therefore thought to be unreliable, was not used routinely. Kinetic constants for the

![Chemical diagram](image_url)

**Figure 2.7.** An Alternate Assay for L-Ribulose-5-Phosphate 4-Epimerase Activity.
epimerase (determined using the assay described by Davis et al. (1972) were $K_M = 0.026 \pm 0.008$ mM and $k_{\text{cat}} = 13.5 \pm 0.3$ s$^{-1}$.

**Determination of Metal Ion Content**

L-Ribulose-5-phosphate 4-epimerase can accept a variety of divalent metal ions for activity (Deupree & Wood, 1972); however, it is not known which of these is present in the enzyme isolated from bacteria at natural abundance. A purified sample of the recombinant enzyme and a sample of its buffer were digested in 10% ultrapure nitric acid. These samples were analysed for metal ion content by inductively-coupled plasma mass spectrometry (ICPMS). The results indicated that the enzyme preparation contained a mixture of metal ions in order of decreasing abundance: Zn$^{2+} >$ Mn$^{2+} >$ Cu$^{2+}$. It is possible that the very high levels of overexpression resulted in the enzyme sequestering a variety of metal ions as it is formed. The metal ion distribution we observed may therefore vary depending on the metal ions present in the growth medium and probably does not reflect the metal ion preferences obtained under normal growth conditions. For further studies we wished to prepare an enzyme with homogenous metal ion content. Zinc is a reasonable choice because it is the divalent metal ion commonly used for electrophilic catalysis in many enzymes including Class II aldolases (see Chapter III for discussion).

In order to prepare epimerase with a homogeneous metal ion content, apoenzyme was prepared by treatment of the epimerase with EDTA as described by Deupree and Wood
(1972), followed by dialysis against two changes of metal-free HEPES/Tris buffer (50 mM, pH 7.6, containing 10% glycerol). The resultant epimerase was inactive as determined using an “end-point” assay. For this assay, metal-free tRu5P (0.94 mM) was incubated in a buffered solution of the epimerase (at 0.01 mg/mL) for five minutes, at which time it was added to a pre-incubated cuvette containing the coupling system. Apoenzyme was reconstituted by incubation with 10 equivalents of ZnCl₂. Excess metal ions were removed by dialysis against two changes of HEPES/Tris buffer (50 mM, pH 7.6, containing 10% glycerol). Inconsistent ICPMS results were obtained after extended dialysis. Therefore, an alternate method for removing metal ions, passage through a size exclusion column, was required to obtain a reasonable result from ICPMS. Incidentally, there was no change in the activity of the enzyme after passage through the size exclusion column. After Zn(II) reconstitution, the enzyme was found to have 1.05 equivalents of Zn²⁺ per subunit (by ICPMS) and kinetic parameters of $k_{cat} = 20.4 \pm 0.9 \text{ s}^{-1}$ and $K_M = 0.087 \pm 0.007 \text{ mM}$. The $K_M$ of the enzyme was similar to those reported in the literature (Table 2.2) except for that of the L. plantarum enzyme.
Table 2.2. Comparison of $K_M$ and $K_{eq}$ Constants Determined for L-Ribulose-5-Phosphate 4-Epimerase.

<table>
<thead>
<tr>
<th>Source Organism</th>
<th>$K_M$ (mM) $^a$</th>
<th>$K_{eq}$ $^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> B/r $^c$</td>
<td>0.095</td>
<td>ca. 1</td>
</tr>
<tr>
<td><em>K. pneumoniae</em> $^d$</td>
<td>0.1</td>
<td>1.86</td>
</tr>
<tr>
<td><em>L. plantarum</em> $^e$</td>
<td>1.1</td>
<td>1.2</td>
</tr>
<tr>
<td><em>E. coli</em> DH1 $^f$</td>
<td>0.087</td>
<td>1.2</td>
</tr>
</tbody>
</table>

$^a$ for L-ribulose-5-phosphate  
$^b$ for [DXu5P]/[LRu5P]  
$^c$ Lee et al., 1968; Gielow & Lee, 1975.  
$^d$ Wolin et al., 1958; Deupree & Wood, 1975.  
$^e$ Burma & Horecker, 1958.  
$^f$ This study.

**Equilibrium**

Treatment of a commercial sample of DXu5P with the enzyme and subsequent analysis of the equilibrium mixture by $^1$H-NMR spectroscopy indicated that the equilibrium was slightly in favour of the DXu5P epimer (see Figure 2.8), in agreement with a previous report by Burma & Horecker (1958b) (see Table 2.2). The equilibrium constant was calculated to be 1.2 from intergration of the peaks arising from the C-4 protons of DXu5P and LRu5P.
Conclusion

The high level of overexpression of L-ribose-5-phosphate 4-epimerase in *E. coli* transformed with pRE1 provided a rich source of the epimerase, which after purification yielded about 100 mg of enzyme per litre of culture. The recombinant enzyme, after zinc ion replacement is indistinguishable from the wild-type *E. coli* enzyme.

Large quantities of pure epimerase makes possible detailed studies on the enzyme, including studies on the nature of its reaction mechanism and possibly x-ray crystal structure determination. As the *araD* gene is cloned, site-directed mutagenesis studies are also feasible.
Figure 2.8. Equilibrium between L-Ribulose-5-Phosphate and D-Xylulose-5-Phosphate as Determined by $^1$H-NMR (500 MHz). Integration of the area below the peaks at 3.94 ppm (H-C4 of LRu5P) and 4.12 ppm (H-C4 of DXu5P) provided relative concentrations of these two sugars.
Chapter II: Cloning, Overexpression and Characterization

Experimental Methods

**Bacterial Strains and Plasmids**

*E. coli* strain DH5αF' (Gibco/BRL) was used for all plasmid manipulations. *E. coli* strain Y1090 (modified to cure it of pMC9) (a gift from Dr. Neil Gilkes, Department of Microbiology, University of British Columbia) was used as the host for overexpression of recombinant L-ribulose-5-phosphate 4-epimerase. *E. coli* strain MC4100 (ATCC #35695) was used in the preparation of substrate. Both *E. coli* Y1090 (cured) and *E. coli* MC4100 were araD139 mutants. *E. coli* DH1 was used as the source of genomic DNA for the PCR amplification of *araD* in construction of pRE1, while *E. coli* DH5αF' was the source of genomic DNA for *araB* PCR amplification. *E. coli* JM109 (DE3) was used as the host for pET-11a overexpression vectors. Cells for general cloning and overexpression were grown in LB media supplemented with ampicillin (50 μg/mL).

The vector used for cloning *araD* was a modified version of pBSTIM (Hermes et al., 1990), in which the *Nde* I restriction site on the original pBS+ vector (Stratagene) was removed by changing nucleotide residue 186 from C to A, and by introduction of an *Nde* I site at the start codon in the cloning region by changing the bases at -1 and -2 (relative to the start codon of TIM) from CC to AT. These modifications were performed by Dr. S. Pollack in the laboratory of Dr. Jeremy Knowles, Harvard. pET-11a (Novagen) was the vector used for attempts at L-ribulokinase cloning. Both of these plasmids encode a gene for ampicillin resistance.
Unless otherwise noted, general methods for handling and manipulating DNA are those described by Sambrook, Maniatis & Fritsch (1989) or Ausubel et al. (1992).

**Preparation of L-Ribulose-5-Phosphate**

L-Ribulose-5-phosphate was prepared according to the procedure of R.L. Anderson (1966). *E. coli* MC4100, an *araD139*, or 4-epimeraseless, mutant of *E. coli* was grown in 2 L of a casein hydrolysate-mineral medium (as described by Englesberg (1961): 1% potassium phosphate (pH 7.0), 0.01% MgSO₄·7H₂O, 0.1% (NH₄)₂SO₄, and 1% Casamino acids) containing yeast extract (0.25 g/L) until the optical density at 600 nm was about 0.4. L-Arabinose (2 g) was added to a concentration of 0.4%, and the incubation continued for another 4 hours. L-Ribulokinase was isolated in an ammonium sulfate fraction of the crude cell lysate, and was used without further purification.

The reaction mixture contained 4 mmole L-ribulose (601 mg, 429 μL), 5 mmole ATP (2.75 g), 5 mmole MgCl₂ (476 mg), 5 mmole NaF (210 mg), 1 mmole reduced sodium glutathione (310 mg), 2.5 mmole EDTA (12.5 μL of a 0.2 M solution), and L-ribulokinase in a total volume of 200 mL. NaOH was added periodically to maintain pH 7.5. After the reaction was complete, 2.29 mL of glacial acetic acid was added to a final concentration of 0.2 M, the solution suction filtered through Whatman #1 filter paper, and the filtrate neutralized with about 19 mL of 2 M NaOH.

L-Ribulose-5-phosphate was purified from other reaction components by ion-exchange chromatography on a 100 mL AG 1-X8 (formate) column (ID 19 mm), eluting
with a stepwise gradient of formic acid: 6 x 100 mL deionized water; 6 x 80 mL 1N formic acid; 6 x 80 mL 2 N formic acid, 12 x 40 mL 2.5N formic acid; and 6 x 80 mL 3N formic acid. LRu5P eluted between 2 and 3 N formic acid. Fractions that were greater than 90% pure by NMR were pooled and lyophilized, and brought to pH 7.0 with NaOH.

$^1$H- NMR (500 MHz, D$_2$O): $\delta$ 4.02 (apparent t, J= 5.41 Hz, 2 H, H-C5), $\delta$ 4.11 (dt J= 5.41 Hz, 1 H, H-C4), $\delta$ 4.43 (d, J= 5.41 Hz, 1 H, H-C3), $\delta$ 4.56 (d, J = 19.5 Hz, 1 H, H-Cl), $\delta$ 4.63 (d, J = 19.5 Hz, 1 H, H'-Cl); see Figure 2.4. A COSY plot is corresponding to this sample and which aided in peak identification is shown in Appendix B.

The concentration of LRu5P in the stock solution was determined by initiating an assay cuvette preincubated at 37°C and containing all the normal assay components except LRu5P, but which contained 50 µg of 4-epimerase, with 10 µL of a 1/10 dilution of the stock solution. The concentration of LRu5P in the cuvette was determined from the absolute change in absorbance at 340 nm due to oxidation of NADH to NAD$^+$. 

**Attempted Cloning of L-Ribulokinase**

The *araB* gene of *E. coli* DH5α, encoding L-ribulokinase (Lee *et al.*, 1986; see Appendix A), was amplified by 40 cycles of PCR between primers AJ8 (containing a *Nde* I restriction site) and AJ9 (containing *Pst* I and *BamH* I restriction sites) (see Table 2.3). Template DNA was the genomic DNA provided in whole DH5αF' cells which had been harvested, washed and resuspended in an equal volume of sterile distilled water. Each reaction contained (in 100 µL): 5 µL DH5αF' cell suspension, 0.1 µM each dNTP, 0.5 µM
each primer, 1.5 mM MgCl$_2$, 50 mM KCl, 20 mM Tris·HCl (pH 8.4), and 2.5 U Taq DNA polymerase (Gibco/BRL). 100 µL of mineral oil was added to prevent evaporation during thermocycling. An initial melt at 95°C (10 min), annealing at 55°C (5 min) and chain extension at 72°C (2 min) was followed by 39 cycles of 95°C (1.5 min), 55°C (2 min) and 72°C (2 min). The reaction was ended with a final extension of 4 minutes at 72°C.

| Table 2.3. Primers Used for PCR Amplification of the $\text{araB}$ Gene. |
|-----------------|-----------------|
| Primer | Sequence |
| AJ8 $^a$ | 5'-GGGAATTCCATATGGCGATTGCAATTGGCCTCGAT-3' |
| AJ9 $^b$ | 5'-CAGGCGGATCCTGCAAGTTATAGAGTGCAACGGGCGCTGTTG-3' |

$^a$ Primer AJ8, whose 24 3'-end bases are complementary to the 5'-end of the $\text{araB}$ gene has a $Nde$ I restriction site (in italics).

$^b$ Primer AJ9 has overlapping $BamH$ I and $Pst$ I restriction sites (5' to 3', in italics), and its 24 3'-end bases are complementary to the 3'-end of the $\text{araB}$ gene.

The resultant DNA fragment was isolated from a 1.2% agarose gel and purified using a Wizard PCR Preps DNA Purification System (Promega). Both the vector (pET-11a; Novagen) and the purified PCR fragment were treated with $Nde$ I and $BamH$ I DNA restriction enzymes. The vector was isolated from a 1.2% agarose gel and purified using a Wizard PCR Preps Purification System (Promega). The restricted PCR fragment was not purified since there is an additional $Nde$ I restriction site in the middle of the gene.

Partial digestion of the PCR-amplified $\text{araB}$ gene, choosing conditions under which approximately half of the DNA was restricted was also tried. These conditions were: 1 µg PCR fragment ($\text{araB}$), 6 U $BamH$ I (10 U/µL), 1 X Buffer D (Promega), water to 6.0 µL,
incubated at 37°C for 3 hours. 2 U Nde I, more Buffer D (to keep 1 X concentration), and water (to a total volume of 10 µL) were then added, and the reaction mixture incubated 15 minutes longer. The mixture was then heated at 80°C for 5 minutes, lyophilized and redissolved in a small volume of water. The restricted araB gene was isolated and purified from a 1.2% agarose gel using Wizard PCR Preps DNA Purification System (Promega), ligated into prepared pET-11a, and transformed into E. coli DH5αF'. Resultant colonies were checked for plasmid containing araB by 1.2% agarose gel electrophoresis of the uncut and restricted (Pst I and BamH I) plasmid DNA. In addition, promising colonies were checked for overexpression in JM109 (induced with 0.4 mM IPTG).

**Cloning of the araD gene of Escherichia coli**

The araD gene was cloned from E. coli strain DH1 by generating a PCR product from 50 µL whole cell suspension (harvested and resuspended in an equal volume of water) using Taq DNA polymerase and 40 cycles of thermal cycling (95°C for 1.5 min, 55°C for 2 min, and 72°C for 2 min) according to the method of Joshi, Baichwal & Ames (1991). The sequence-specific primers (MT04 and MT05) encompassed the complete gene and contained restriction sites for Nde I and Pst I for cloning into pBSTIM (see Table 2.4). The resulting PCR product was purified by passage through an Ultrafree-MC 10000 NMWL filter unit (Millipore), followed by extraction with phenol and “SEVAG” and then ethanol precipitated.
Table 2.4. Primers Used for PCR Amplification of the \textit{araD} Gene.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MT04 $^a$</td>
<td>5'-CCGCGCGGGGGATCCTGCAAGTTACTGCCCCGTAATATGCG-3'</td>
</tr>
<tr>
<td>MT05 $^b$</td>
<td>5'-GGCCCCATGGCCCATATGTTAGAAGATCTCAAAACGCGAGG-3'</td>
</tr>
</tbody>
</table>

$^a$ MT04 has overlapping \textit{BamH I} and \textit{Pst I} restriction sites (5' to 3', respectively, in italics), and the 19 bases at its 3'-end are complementary to the 3'-end of the \textit{araD} gene.

$^b$ MT05 has \textit{Nco I} and \textit{Nde I} restriction sites, (5' to 3', respectively, in italics), and the 25 bases at its 3'-end are complementary to the 5'-end of \textit{araD}.

Both the vector (pBSTIM) and the PCR product (from six pooled reactions) were digested with \textit{Nde I} and \textit{Pst I}. Additionally, the vector was treated with alkaline phosphatase. These samples were isolated by electrophoresis on a 1% low melt SeaPlaque agarose gel (FMC BioProducts) and the appropriate bands (as visualized with ethidium bromide and ultraviolet light) were excised. An in-gel ligation was then performed using T4 DNA ligase for 30 h at 15°C. The ligation mixture was melted and immediately transformed into \textit{E. coli} XL1-Blue by electroporation. Colonies were picked from LB agar plates containing 100 μg/mL ampicillin and used to inoculate 10 mL overnight cultures of LB containing 200 μg/mL ampicillin. Plasmid pRE1 was obtained from an overnight culture using the Magic Minipreps DNA Purification System (Promega). This work was performed by Dr. Martin Tanner in the laboratory of Dr. Jeremy Knowles (Harvard, MA).
Attempts to Subclone the araD gene of Escherichia coli

The araD gene of E. coli was amplified by PCR from plasmid pHC5 (Chen et al., 1990) kindly donated by Dr. R.E. Moses, and with the sequence-specific primers MT04 and MT05 used previously. The araD gene was amplified in a MJ Research MiniCycler. The reaction mixture (100 µL) contained: 0.1 µM each dNTP, 0.5 µM each primer, 0.1 µg pHC5, 1.5 mM MgCl₂, 50 mM KCl, 20 mM Tris-HCl (pH 8.4), and 2.5 U Taq DNA polymerase (Gibco/BRL). 100 µL of mineral oil was added to prevent evaporation during thermocycling. An initial melt at 94°C (5 minutes), annealing at 50°C and chain extension at 72°C (5 minutes) was followed by 28 cycles of 94°C (1 min), 50°C (2 min) and 72°C (2 min). The final cycle was 95°C (1 min), 50°C (2 min) and 72°C (10 min). The resultant PCR product was prepared for insertion into an overexpression vector by digestion with NdeI and PstI and purified from a 1.2% agarose gel using a Nucleotrap Extraction Kit for Nucleic Acids (Macherey-Nagel).

The vector, pRE1, was digested with PstI and NdeI and either treated with calf alkaline phosphatase (Promega) or purified from a 1.2% agarose gel with the Nucleotrap Extraction Kit for Nucleic Acids. The amplified, restricted DNA fragment was ligated (16 h at 15°C) into the prepared vector and transformed into E. coli DH5α. Resultant colonies were checked for plasmid DNA of appropriate size and for protein overexpression. At the time we had no other overexpression vectors into which to insert the araD gene.
Chapter II: Cloning, Overexpression and Characterization

Calculation and Determination of Subunit Molecular Mass

The subunit molecular mass was calculated to be 25520 Da based on the amino acid composition. Verification was provided by electrospray ionization mass spectrometry performed by Dr. Shichang Miao and David Chow, which determined the subunit molecular mass to be 25522 ± 4 Da. Electrospray ionization experiments used an HPLC-ESMS setup consisting of a microbore HPLC (Michrom UMA) connected in-line to either a PE-SCIEX APIII or API 300 MS as described by Hess et al. (1993). Intact epimerase samples (10 µL, ≥1µg/µL) were injected into a microbore PLRP column (1 x 50 mm) and eluted with a linear gradient of 16 to 80% acetonitrile in water (containing 0.05% trifluoroacetic acid) over three minutes and maintained at 80% acetonitrile for an additional seven minutes. The eluant was introduced into the spectrometer, which was operated in the single quadrupole mode. Protein molecular weights were determined from this data using deconvolution software supplied by Sciex.

DNA sequencing

The araD gene of pRE1 was sequenced in entirety using the Sanger dideoxynucleotide chain termination method and Sequenase version 2.0 DNA Sequencing Kit (United States Biochemical) to ensure that no errors had occurred during PCR amplification of the gene. Double-stranded plasmid DNA was isolated for this purpose using a Wizard Minipreps DNA Purification System (Promega). Primers used for sequencing included MT04 and MT05 as well as MT12, MT13, MT14 and MT15 (see Table 2.5).
### Table 2.5. Primers Used in Sequencing the \textit{araD} Gene.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MT12 $^a$</td>
<td>5'-GGCGTCGATTACAGCGTCATGA-3'</td>
</tr>
<tr>
<td>MT13 $^b$</td>
<td>5'-GGGCGCAGGCCGGTCAGTCGAT-3'</td>
</tr>
<tr>
<td>MT14 $^c$</td>
<td>5'-GGCGTTCGTTTGTTCTCCATTCCTCCACGGCCCCT-3'</td>
</tr>
<tr>
<td>MT15 $^d$</td>
<td>5'-CGTACCTCAACCACCTTCACCGGT-3'</td>
</tr>
</tbody>
</table>

$^a$ Identical to the sense strand of \textit{araD} between nucleotides 133 and 154.

$^b$ Identical to the sense strand of \textit{araD} between nucleotides 311 and 332.

$^c$ Identical to the sense strand of \textit{araD} between nucleotides 493 and 520.

$^d$ Complementary to the sense strand of \textit{araD} between nucleotides 210 and 187.

---

**Enzyme Purification**

The procedure for purification of recombinant \textit{L}-ribulose-5-phosphate 4-epimerase was achieved using a modification of the method described by Lee \textit{et al.} (Gielow & Lee, 1975; Lee \textit{et al.}, 1968).

For purification of \textit{L}-ribulose-5-phosphate 4-epimerase, pRE1 was transformed into \textit{E. coli} Y1090 (modified), an \textit{araD} mutant. LB broth (500 mL) supplemented with 50 \(\mu\)g/mL ampicillin in a 1 L flask was inoculated with a colony from the transformation and incubated overnight at 37°C with agitation. Cells were harvested by centrifugation and resuspended in 10 mM potassium phosphate buffer, pH 7.6 containing 10% glycerol, 1 \(\mu\)g/mL aprotinin and 1 \(\mu\)g/mL pepstatin. Cells were lysed using a French Press at 20000 psi. Cell debris was removed by a 20 minute centrifugation at 5000 rpm in a GSA rotor (Sorvall) followed by
ultracentrifugation to remove insoluble cell debris for 1 hour at 30000 rpm in a 50 Ti rotor (Beckman). The supernatant was brought to 40% ammonium sulphate saturation at 4°C. The precipitate was collected by centrifugation (20 minutes at 5000 rpm in a Sorvall GSA rotor) then redissolved in fresh phosphate buffer (10 mM; pH 7.6, containing 10% glycerol, 1 µg/mL aprotinin and 1 µg/mL pepstatin). Ammonium sulphate was removed by dialysis overnight against 2 L fresh buffer.

The protein solution was diluted from 5 mL to approximately 50 mL with 10 mM potassium phosphate buffer (pH 7.6, containing 10% glycerol, 1 µg/mL aprotinin and 1 µg/mL pepstatin) and loaded onto a 20 mL DE-52 (Whatman) column, washed with 50 mL 10 mM potassium phosphate buffer (pH 7.6, containing 10% glycerol, 1 µg/mL aprotinin and 1 µg/mL pepstatin), then eluted with 25 mL 0.4 M NaCl in 10 mM phosphate buffer, pH 7.6 (containing 10% glycerol, 1 µg/mL aprotinin and 1 µg/mL pepstatin). The eluent was concentrated using a Millipore Ultrafree-15 centrifugal filter device with a Biomax-10K NMWL membrane. Excess NaCl was also removed by washing the protein several times with salt-free buffer. The enzyme solution was frozen in liquid nitrogen and stored at -76°C overnight.

The protein sample was filtered through a 0.2 µm syringe filter and about 30 mg were loaded onto a Waters Q-Pak (Q 8HR) HPLC column. The column was eluted with a linear gradient of 0 to 0.4 M NaCl in 10 mM potassium phosphate, pH 7.6 containing 10% glycerol. Flash-frozen samples of purified epimerase were stored in aliquots at -76°C.
Purity was checked by 12% SDS-PAGE on a Bio-Rad Mini-PROTEAN II electrophoresis system. Protein bands were visualized by staining with Coomassie Brilliant Blue.

**Protein Determination**

The extinction coefficient of L-ribulose-5-phosphate 4-epimerase was determined according to the method described by Gill and von Hippel (1989). Various amounts of a protein solution were diluted to 500 μL and mixed with an equal volume of either 6 M guanidinium-HCl, pH 7.7 or potassium phosphate buffer, pH 7.7, and absorbances of these diluted proteins were measured at 280 nm. Using the relationship described in equation 2.1, \( \epsilon_n \) (the extinction coefficient of the native protein) was calculated. \( A_h \) (the absorbance at 280 nm of the native protein) is the measurement in the absence of guanidinium-HCl, \( A_d \) (the absorbance at 280 nm of the denatured protein) is the measurement in the presence of guanidinium-HCl, and \( \epsilon_d \) (the extinction coefficient of the denatured protein) calculated from the number of Trp, Tyr residues and disulphide bonds in the protein as described by Edelhoch (1967): 4 Trp + 9 Tyr + 0 disulphide bonds = 4(5690 cm\(^{-1}\) M\(^{-1}\)) + 9(1280 cm\(^{-1}\) M\(^{-1}\)) = 34280 cm\(^{-1}\) M\(^{-1}\). \( A_d \) (from the guanidinium-HCl containing solutions) and \( A_n \) were measured for each protein concentration. The extinction coefficient determined in this manner, and averaged from three such determinations was \( \epsilon_n = 34314 \) cm\(^{-1}\) M\(^{-1}\) per subunit (1.76 mL mg\(^{-1}\) cm\(^{-1}\)), but was variable, probably due to incomplete denaturation of the protein.
The extinction coefficient at 280 nm was also determined by measuring the absorbance at 280 nm and by analysis of the total amino acid composition. Amino acid analysis on a sample of the purified epimerase (in H$_2$O) to which norleucine (0.1 mM) had been added as an internal standard was performed by Suzanne Perry in the Nucleic Acid and Protein Service at the University of British Columbia. The amount of alanine in the sample was used to calculate the concentration of the protein in the solution of known absorbance. Then, the extinction coefficient was determined using the relationship $A = \varepsilon cl$. This method gave $\varepsilon_n = 33717$ M$^{-1}$ cm$^{-1}$ per subunit, or 1.73 mL mg$^{-1}$ cm$^{-1}$, which is in agreement with the extinction coefficient calculated above. Calculations from the spectrophotometric measurements were performed using the extinction coefficient calculated from amino acid analysis of the protein (1.73 mL mg$^{-1}$ cm$^{-1}$).

**Determination of Metal Ion Content**

The recombinant L-ribulose-5-phosphate 4-epimerase was analysed for metal ion content by Bert Mueller (Dept. of Earth and Ocean Sciences, University of British Columbia) using ICPMS. The enzyme sample was centrifuged briefly (10 minutes) in a Centricon-10 filter unit. The concentration of the resultant enzyme sample was determined. The concentrated enzyme and the filtrate from centrifugation were diluted 1/400 in 10% nitric acid (ultrapure (twice distilled on quartz) from Seastar, Sidney, B.C.). Sc was added to 20 ppb as an internal standard. The samples were digested at 50°C for 12 h prior to analysis. The zinc content of the filtrate was subtracted from that in the enzyme sample.
Chapter II: Cloning, Overexpression and Characterization

**Preparation of Metal-Free Buffers and Glassware**

Metal-free buffer was prepared by passing a solution of 50 mM HEPES (2 L, containing 10% glycerol) through a column of Chelex-100 resin (70 mL Na\(^+\) form, 50-100 mesh, changed to H\(^+\) form by washing with 2 volumes 1 N HCl and rinsing with 5 volumes deionized H\(_2\)O). The pH was adjusted to 7.6 using solid Tris base. All the glassware and plasticware used in this procedure was soaked overnight in 4 N HCl and washed thoroughly with deionized water (17.8 MΩ/cm at 25°C).

**Preparation of Zn-Substituted Enzyme and Analysis of Zinc Content**

Apoenzyme was prepared using a modification of the method described by Deupree and Wood (1972): protein (5 mg) was incubated with 20 mM EDTA (2.5 mL) for 3 h at 25°C. The enzyme was then dialysed against metal-free HEPES/Tris buffer (50 mM, pH 7.6, 2 x 500 mL, containing 10% glycerol) at 4°C. A stopped assay for LRu5P was employed to ensure that the enzyme was inactive (less than 5% remaining activity; see the assay for apoenzyme activity).

The apoenzyme was reconstituted by the addition of 99.99% pure ZnCl\(_2\) (10 equivalents, 250 μL of a 3.9 mM solution) and incubated for 2 h at 25°C. The reconstituted enzyme was dialysed against metal-free HEPES/Tris buffer (50 mM, pH 7.6, 2 x 500 mL) containing 10% glycerol at 4°C and flash frozen in liquid nitrogen.

Samples to be analysed by ICPMS were passed through a size exclusion column (Waters Protein Pak 125), eluted with the same HEPES/Tris buffer, and then diluted 1/400
in 10% nitric acid (ultrapure (twice distilled on quartz) from Seastar, Sidney, B.C.). The column had previously been treated with 200 mM EDTA and then equilibrated with the HEPES/Tris buffer. The protein concentration of the fraction containing the epimerase was determined. Sc was added to 20 ppb as an internal standard. The samples were digested at 50°C for 12 h prior to analysis. These samples were analysed on a VG Elemental Plasma Quad mass spectrometer that had been calibrated with a standard curve of Zn (0, 5, 15, 40 ppb, prepared in 10% ultrapure nitric acid containing a 1/400 dilution of the elution buffer and 20 ppb Sc as an internal standard). The analyses were performed by Bert Mueller in the Department of Oceanography at the University of British Columbia.

**Measurement of Enzyme Activity**

Enzyme activity was measured according to the method of Davis, Lee and Glaser (1972). In this enzyme assay each cuvette contained, in a total volume of 1.00 mL, 25 mM glycylglycine (pH 7.6), 5 mM ribose-5-phosphate, 0.1 mM TPP, 0.15 mM \( \text{NADH} \), \( \text{I} \text{Ru5P} \) between 0.02 and 1 mM, 5 U \( \text{aGDH} \), 50 U TIM, 0.25 U transketolase (Sigma T-6133 from Bakers yeast, dissolved in water). An \( \text{aGDH}/\text{TIM} \) mixture could be purchased in the correct ratio (Sigma G-1881 from rabbit muscle), otherwise an equivalent mixture could be prepared from \( \text{aGDH} \) (Sigma G-6751 from rabbit muscle) and TIM (Sigma T-2391 from rabbit muscle). In either case, these coupling enzymes were in a suspension containing ammonium sulphate and EDTA. Enough of this mixture for ten reactions (24 \( \mu \text{L} \)) was diluted into water (2 mL) and centrifuged in a Millipore 10K NMWL filter unit for 20
minutes prior to use. The resulting enzyme solution (normally 50 μL) was diluted to 500 μL in water. Each cuvette contained 50 μL of the prepared αGDH/TIM mixture. The cuvettes were equilibrated at 37°C and the background rate was monitored for 5 minutes prior to initiation of epimerisation by addition of 0.05 μg of purified L-ribulose-5-phosphate 4-epimerase. The reaction was followed by the decrease in absorbance at 340 nm on a Cary 3E spectrophotometer. The rate of formation of DXu5P was calculated from the initial slopes using an extinction coefficient of NADH of 6220 M⁻¹ cm⁻¹. The background rates were calculated in the same manner from data collected prior to addition of the epimerase. The background rates were subtracted from the rates after addition of epimerase to obtain the rate due to epimerase action. Kinetic parameters were determined from a direct fit of the data to an enzyme kinetics equation using the computer program GraFit (Erithacus Software, UK).

**GAPDH Assay for Enzyme Activity**

An alternate coupled assay for L-ribulose-5-phosphate 4-epimerase activity is described by Wolin, Simpson and Wood (1958). Each cuvette in this assay contained (in a total volume of 1.0 mL): 20 mM glycyglycine (pH 7.4), 0.15 mM NAD⁺, 3.4 mM sodium arsenate (pH 7.5), 1 mM DTT, 0.1 mM TPP, 150 μg crystalline G-3-P dehydrogenase (Sigma G-2267, dissolved in water), 0.50 U transketolase (Sigma T-6133 from Bakers yeast, dissolved in water), 1 mM MgCl₂, 5 mM DR5P, and LRu5P between 0.02 and 1 mM. After the cuvette had reached 37°C, the absorbance was monitored for 5 minutes immediately
prior to initiation by addition of 0.05 µg of the purified 4-epimerase. Rates were determined in exactly the same manner as with the assay described above. The background rates tended to be extremely high.

Assay for Apoenzyme Activity

Apoenzyme at 0.01 mg/mL was incubated at 37°C for 5 minutes in metal-free HEPES/Tris buffer, pH 7.6, containing 9.44 mM lRu5P. Controls were water and untreated (active) enzyme. Aliquots (10 µL) were immediately added to pre-equilibrated cuvettes at 37 °C containing (in 990 µL) 25 µmoles glycylglycine (pH 7.6), 0.15 µmole NADH, 5 µmoles dR5P, 0.1 µmole TPP, 5 U αGDH, 50 U TIM and 0.25 U transketolase. The coupling enzymes were prepared as described for measurement of enzyme activity. The cuvettes were monitored at 340 nm, and the drop in absorbance for the first minute immediately after addition of the enzyme-containing aliquots was measured. The average drop in absorbance per minute before addition of the aliquot was determined and was subtracted from the drop measured after addition of the aliquot. The amount of activity observed in the apoenzyme assayed in this manner was typically about 5% of that observed for the untreated enzyme.
Chapter II: Cloning, Overexpression and Characterisation

**Measurement of Equilibrium**

The Zn$^{2+}$ form of the epimerase (0.5 mg) was incubated at 37°C with 5 mg of the sodium salt of D-xylulose-5-phosphate (Sigma) at pH 7.6 in 1 mL for 4 hours. At this time the enzyme was removed from the equilibrium mixture of the sugars by passage through a Millipore Ultrafree-4 centrifugal filter device with a Biomax-10K NMWL membrane. The filtrate was lyophilized, exchanged with D$_2$O twice and finally dissolved in 500 µL of D$_2$O. This mixture was analysed by quantitative 500 MHz $^1$H-NMR at 25°C. Integration of the peaks arising from the C-4 protons of each epimer provided the relative concentration of each epimer.
Chapter III

Nature of the Metal Ion Ligands of
L-Ribulose-5-Phosphate 4-Epimerase

Introduction

The *Escherichia coli* L-ribulose-5-phosphate 4-epimerase (Deupree & Wood, 1972) and
L-fuculose-1-phosphate aldolase are metalloenzymes with a homologous N-terminal domain
(Dreyer & Schulz, 1996b). Three of the four zinc ligands of the aldolase lie within this
domain and are reported to be conserved in the epimerase (Dreyer & Schulz, 1993; Dreyer
& Schulz, 1996b). A greater understanding about how L-ribulose-5-phosphate 4-epimerase
operates can be gained by characterizing its metal binding site both by metal ion substitution
and by site directed mutagenesis of the putative metal binding ligands. These ligands can be
identified by mutation to residues known to be poorer metal ion ligands. These changes
would result in the subsequent loss of the ability to bind metal ions tightly and reduction in
catalytic activity. Identification of the epimerase’s metal binding ligands as those predicted from homology to the aldolase would serve to establish a structural link between these two enzymes.

**Zinc Ions in Enzyme Catalysis**

It has long been established that many enzymes make use of metal ions in promoting catalysis as well as in forming structural motifs (Mildvan, 1972; Ochiai, 1987). Zinc, which is the second most abundant metal in biological systems, is an important catalytic component in each of the six enzyme classes and across all phyla (Lipscomb & Sträter, 1996; Vallee & Auld, 1990a; Vallee & Auld, 1990b). Zinc is thought to be preferred over other divalent cations for a number of reasons. With a flexible coordination geometry, zinc can have anywhere between four and six ligands. Zinc can undergo fast ligand association and dissociation, which also allows the substrates and products to leave the active site. Zinc can act as a Lewis acid, has intermediate polarizability, is readily available, and can be bound tightly into suitable sites on proteins. Furthermore, zinc does not readily participate in redox reactions (Lipscomb & Sträter, 1996; Maret & Vallee, 1993).

A role for catalytic Zn$^{2+}$ ions in enzymes is to act as electrophilic catalysts. The Zn$^{2+}$ ion acts as a Lewis acid (electrophile), which can coordinate to carbonyl groups. This catalyst stabilizes the developing negative charge during the course of a reaction. Coordination of a carbonyl group to the metal ion polarizes the substrate towards nucleophilic attack (by another group in the enzyme’s active site) or enolization.
A comparison of twelve zinc enzymes whose crystal structures are known has uncovered some common features of the ligands used for binding the catalytically active zinc ion (Vallee & Auld, 1990a; Vallee & Auld, 1990b). In each case, the zinc ions are coordinated by three amino acid residues in the enzyme’s active site and by an activated water molecule. In contrast, structural zinc ions are almost universally coordinated to four cysteine residues in a tetrahedral arrangement.

Zinc forms complexes with nitrogen and oxygen ligands as readily as it does with sulphur ligands. In the catalytic zinc sites, the imidazole group of histidine, the sulphydryl group of cysteine, and the carboxyl groups of glutamic acid and aspartic acid are all possible ligands for zinc. Histidine is, however, the most predominant ligand in catalytic zinc sites, followed by glutamate, aspartate and cysteine. A combination of three of these residues is used to bind zinc tightly in the active site and leave an open coordination sphere. In the free form of the enzyme an “activated” water molecule fills and completes the coordination sphere of the zinc ion. The water molecule can be ionized, polarized, or displaced upon substrate binding.

The first two ligands in catalytic zinc sites are separated by a “short spacer” consisting of one to three amino acid residues. These ligands are separated from the third ligand by a “long spacer” of between about 20 and 120 amino acid residues. The short spacer enables the formation of a primary bidentate zinc complex, whereas the long spacer permits flexibility of coordination number (this can change during the course of a reaction) and geometry. A long spacer also allows for any conformational changes which may occur
during the course of the catalytic reaction, and might be part of the substrate-binding pocket. Structural zinc binding sites have much shorter spacers, which imparts rigidity to the metal centres and is consistent with their role in stabilizing the overall structure of the protein and local conformation (Vallee & Auld, 1990a; Vallee & Auld, 1990b).

Tetradentate zinc-binding sites are inaccessible to solvent. For example, procollagenase has tetradentate zinc coordination, but upon conversion to collagenase loses the fourth amino acid ligand, which is replaced with a water molecule (Vallee & Auld, 1990b). During this process, a structural zinc ion becomes catalytic.

**Characterization of Zinc Binding Sites**

Site directed mutagenesis can be used to alter any amino acid residue of a cloned and overexpressed protein to any other amino acid and can lead to the identification of putative metal-binding ligands. Targets for mutagenesis can be either amino acid residues conserved over a number of related structures or all potential zinc-binding ligands (histidine, glutamate, aspartate or cysteine residues). Recent examples of these kinds of mutations are the identification of histidine ligands in the binuclear metal-ion site of phosphotriesterase (Kuo & Raushel, 1994), identification of histidine zinc ligands of Class II fructose-bis-phosphate aldolase (Berry & Marshall, 1993), confirmation of the zinc-binding site residues of a metallo-β-lactamase (Crowder *et al.*, 1996), and the identification of potential zinc binding ligands in VanX (McCafferty *et al.*, 1997). In addition to ligand identification, site-directed
mutagenesis can be used to probe the structural and functional roles of known metal ion ligands (e.g., Alexander et al., 1993; Kuo et al., 1997).

While histidine, glutamate and aspartate residues are often changed to alanine or valine (and cysteine to serine), mutation of histidine or aspartate to asparagine is another viable alternative (Kuo & Raushel, 1994). This is a more conservative mutation in that there is a possibility that the new amino acid residue may be able to make the same hydrogen bonds that were made by the histidine residue it replaces without itself acting as an effective metal ion ligand. Thus, while catalytic activity and/or the ability to bind the metal ion may be affected, the structural effects should be less drastic than with other mutations. This approach has been used to identify ligands of a binuclear metal centre in phosphotriesterase (Kuo & Raushel, 1994) and the iron ligands of tyrosine hydroxylase (Ramsey et al., 1995), and will also be used in this study.

As long as the mutant enzyme has similar physical characteristics to the wild type enzyme, differences in activity ($k_{cat}$), substrate binding (as approximated by $K_m$) or metal binding ability may be attributed to the amino acid residue that was changed.

A number of other techniques can also provide information about the zinc-binding site, either in combination with studies on mutant enzymes, or alone. Many of these techniques are spectroscopic and require that the Zn$^{2+}$ ion, which is spectroscopically silent, be removed and replaced with other divalent metal ions. These replacements, which can include Co$^{2+}$, Cu$^{2+}$, Ni$^{2+}$, Mn$^{2+}$, or Cd$^{2+}$, often yield active, sometimes even hyperactive
enzymes. Of these divalent metal ions, Co$^{2+}$ is most likely to replace Zn$^{2+}$ efficiently (Lipscomb & Sträter, 1996; Maret & Vallee, 1993).

UV-visible spectra of cobalt(II)-substituted zinc enzymes can provide information on the environment in which the catalytic zinc ion lies. If the coordination geometry of the enzyme's metal-binding site is simple, then it may be possible to determine this geometry from visible spectra, based on the extinction coefficients of absorption features between 500 and 600 nm (Elgren et al., 1994). For most enzymes this is not normally possible; the coordination geometry of the metal ion in the binding site is complex or unsymmetrical. UV-visible spectra can be used to determine the presence and number of thiolate ligands (from cysteine) based on intense bands in the near-ultraviolet. In addition, the absorption maxima of the enzyme-complexed cobalt(II) ion can indicate the type of ligands: in the order of sulphur, nitrogen, and oxygen protein donor ligands, the absorption maxima are progressively shifted to higher energies (shorter wavelengths). In practice this does not allow for distinction between oxygen and nitrogen donor atoms (Maret & Vallee, 1993). Sulphur donor ligands are also known to increase the molar absorptivities of some Co(II) complexes. If thiolate ligands are mutated to nitrogen or oxygen ligands, the spectra of the resulting Co(II) complex may not be readily observable, as was observed with a cysteine to serine mutant of a metallo-β-lactamase (Crowder et al., 1996).

NMR spectroscopy of proteins with a paramagnetic metal centre is a powerful technique for the detection of ligand binding, conformational changes of the protein induced by ligands, or ionizations near the metal site (Bertini et al., 1993; Maret & Vallee, 1993). The
somewhat short electron relaxation time of the cobalt(II) ion leads to relatively well-resolved isotropically shifted resonances in the $^1$H-NMR spectra of metal-coordinating ligands. Proton-NMR of Co$^{2+}$-substituted enzymes can be used to ascertain if the metal ion has any histidine ligands, and often the number of these ligands. The imidazole-NH proton of the histidine metal ligands of Co$^{2+}$-substituted metallo-$\beta$-lactamase from *Bacteroides fragilis* appear downfield between 40 and 60 ppm (Figure 3.1), and as there are three of these resonances, three histidine residues are implicated as metal ion ligands (Crowder *et al.*, 1996).

![Figure 3.1. $^1$H-NMR of the Co$^{2+}$-substituted *B. fragilis* metallo-$\beta$-lactamase (1 mM) at pH 7.4. The three solvent-exchangeable histidinyl NH protons are marked with asterisks. Taken from Crowder *et al.*, 1996.](image)

In addition, if the enzyme is substituted with $^{113}$Cd$^{2+}$, the chemical shift of the metal ion (by $^{113}$Cd-NMR) can indicate the number, type, geometry and dissociation constants of the metal ion's ligands. $^{113}$Cd-NMR can therefore be used to provide insights into protein-
substrate interactions, conformational changes, and metal displacement reactions (Kanaori et al., 1996; Summers, 1988).

Oxidation of Co$^{2+}$-substituted enzymes to Co$^{3+}$-substituted enzymes can be used to investigate the role which the metal ion plays in catalysis (Maret & Vallee, 1993). Co$^{3+}$-substituted enzymes will become inactive if substrate ligation to the metal ion is important to catalysis; Co$^{3+}$ is "substitution-inert". That is, ligands cannot dissociate from Co$^{3+}$ as easily as they might from Co$^{2+}$.

The ligand exchange-inert properties of Co$^{3+}$ may be used in an additional manner. After in situ oxidation of a Co$^{2+}$-substituted enzyme to a Co$^{3+}$-substituted enzyme, limited proteolysis and subsequent purification of the cobalt-bearing peptide can lead to identification of the metal ion ligands (Maret & Vallee, 1993).

A recent study using a Co$^{3+}$-substituted enzyme is that of Hlavaty and Nowak (Hlavaty & Nowak, 1997), who investigated the nature of the phosphoenolpyruvate carboxykinase active site. The Co$^{3+}$-substituted enzyme was found to retain 15 to 25% of the activity of the Co$^{2+}$-substituted enzyme, consistent with ligand exchange not being essential to catalysis in this enzyme. Subsequent proteolytic degradation and purification of the cobalt-containing peptide led to the identification of this peptide, which had two aspartic acid residues as the only feasible metal ion ligands.
The Zinc-Binding Site of L-Fuculose-1-Phosphate Aldolase

Recent X-ray crystal structures of L-fuculose-1-phosphate aldolase in both its normal Zn$^{2+}$-form and the Co$^{2+}$-substituted form have shown unambiguously that this enzyme binds its catalytic zinc ion with four ligands: E73, H92, H94 and H155 (Figure 3.2) (Dreyer & Schulz, 1993; Dreyer & Schulz, 1996b). The spacing of these ligands follows the pattern established by Vallee & Auld (1990a, 1990b): there is a short spacer of one amino acid residue between H92 and H94, and the spacing between these two ligands and E73 is 20 residues. The spacing between these two histidine residues and H155 is 61 residues, which is less than the maximum long spacer length of 120 amino acid residues. Further crystallographic studies on this enzyme and the inhibitor phosphoglycolohydroxamate (which mimics the intermediates in the aldolase’s reaction mechanism) have provided evidence that the E73 zinc ligand rotates out of the ligand sphere upon binding of the substrate or inhibitor (Figure 3.3) and participates in catalysis as an active site base (Dreyer & Schulz, 1996a; Fessner et al., 1996). While this enzyme is unique in that it binds the catalytic zinc ion with four residues, the glutamate ligand appears to be acting like a water molecule in that it becomes displaced from the ligand sphere upon substrate binding. Note, however, that glutamate is charged and bidentate, while water is neutral (and monodentate).
Figure 3.2. Coordination of Zn$^{2+}$ in the Active Site of L-Fuculose-1-Phosphate Aldolase (PDB code: 1FUA). The catalytically active zinc ion is tightly coordinated by three N$^\equiv$ atoms of histidines 92, 94 and 155, and a bidentate contact from glutamic acid residue 73, resulting in a distorted tetrahedral coordination sphere.
Design of Mutant L-Ribulose-5-Phosphate 4-Epimerases

In order to probe the metal binding site of L-ribulose-5-phosphate 4-epimerase it was decided to mutate each of the putative metal ion ligands to an asparagine residue. The N-terminal domain of L-ribulose-5-phosphate 4-epimerase was reported to have significant homology to the N-terminal domain of L-fuculose-1-phosphate aldolase (Dreyer & Schulz, 1993; Dreyer & Schulz, 1996b). Three of the metal-binding residues of the aldolase (E73, H92 and H94), which lie in the homologous domain, were conserved in the amino acid sequence of the 4-epimerase (Dreyer & Schulz, 1993; Dreyer & Schulz, 1996b), and therefore the corresponding residues (D76, H95, and H97) were postulated to be the epimerase’s metal ion ligands. An aspartate residue has replaced the glutamate residue in the
aldolase. Subsequently, amino acid sequence alignments using a different alignment algorithm to that used by Dreyer & Schulz has implicated H171 in the 4-epimerase as the fourth ligand (see the experimental and results sections later in this chapter). A structural link between the epimerase and the aldolase suggests that there may also be a mechanistic link. Therefore, identification of the metal ion ligands of the epimerase predicted from homology to the aldolase implies that the epimerase follows a mechanism similar to the aldolase.

A first step in this study involves the design and construction of asparagine mutants of each of the putative metal ion ligands. Asparagine (Figure 3.4) was chosen for the same reason that Kuo and Raushel (1994) chose it: asparagine can participate in some of the hydrogen bonds in which histidine participates and is smaller in size, yet it is not an efficient metal ion ligand. Asparagine was also chosen as a substitute for the aspartate ligand. Asparagine has a similar shape and size to aspartate, but it is neutral whereas aspartate is negatively charged. Therefore, the mutant L-ribulose-5-phosphate 4-epimerases should better retain the active site structure and the overall folding of the protein should be relatively unaffected.
Site-directed mutagenesis of histidine residues to asparagine involves only a modification of one nucleotide in the DNA sequence: histidine residues are encoded by CAT or CAC, while asparagine residues are encoded by AAT or AAC. Similarly, the change from an aspartic acid (encoded by GAT or GAC) to an asparagine residue is also a single base change in the nucleotide sequence. However, for easy screening of mutant plasmids, it is useful to make more changes and to introduce silent mutations which allow restriction enzymes to cut the mutant enzymes' plasmid DNA but not the plasmid encoding the wild type epimerase. Alternatively, if a unique restriction site exists near the site of the site-directed mutagenesis, this site may be removed by a silent mutation.

A number of different methods of site-directed mutagenesis are currently available. One with the greatest fidelity is the Kunkel method, and requires that the plasmid have an origin of replication recognized by the filamentous phage F1. If this site exists on the plasmid, which is then also called a phagemid, then one of the strands of the plasmid may be encapsulated within the protein coat of a helper phage instead of the phage's own DNA. Growth of this plasmid in a $\textit{dut}^{-}\textit{ung}^{-}$ strain of $\textit{E. coli}$ will produce vector with the occasional uracil in place of thymine in the newly synthesized DNA. This strain of $\textit{E. coli}$ is deficient in dUTPase ($\textit{dut}$) and uracil-N-glycosylase ($\textit{ung}$). Coinfection with a helper phage is required to isolate the single stranded phagemid.

A mutagenic primer (a synthetic single stranded oligodeoxynucleotide containing the desired base changes) is annealed to a single-stranded form of the plasmid isolated from the helper phage. The complementary strand is then synthesized with DNA polymerase (either
T4 or T7) and deoxynucleotides using the single-stranded plasmid as the template. Ligase is used to close the newly formed strand of DNA to the 5'-end of the mutagenic primer. When this double-stranded DNA plasmid is transformed into a host strain of *E. coli* with an active uracil-N-glycosylase, the uracil-containing template strand is destroyed and the new strand containing the mutation becomes the template for production of the mutant plasmid. This mutated plasmid can then be replicated using normal cellular replication machinery (Kunkel, 1992; Zhou *et al.*, 1990).

Removing the *araD* gene from pRE1 (the overexpression plasmid containing the *araD* gene encoding L-ribulose-5-phosphate 4-epimerase, discussed in Chapter II) and cloning new genes into the “empty” plasmid proved to be extremely difficult. Although we tried to solve this problem, it was unclear where the problem lay. Excision of the *araD* gene, ligation of new genetic material and restriction of this DNA in preparation for ligation are all possible sources of difficulty. Therefore, mutagenic methods involving recloning the mutated gene into an overexpression vector were not attractive. When the Kunkel method (as described above) cannot be used, other methods (almost all PCR-based) are available. A mutagenesis strategy involving PCR-amplification of the entire plasmid was designed for this study.

One method which does not require recloning the gene to be mutated, is the “recombinant circle PCR” method of Jones and Winistorfer (1992). In this method (Figure 3.5), the plasmid is amplified in two separate reactions as two “halves”. Each half has an end which is homologous to about 200 bases at one end of the other half. The second end covers the mutated site and also has about 20 bases homologous to the other PCR product.
Figure 3.5. The RCPCR Method of Site-Directed Mutagenesis. Small triangles indicate the site of mutagenesis.

When the products of these reactions are combined, denatured and reannealed, they can form circular DNA products, which after transformation, are extended and ligated by the host cell. Since the PCR products contained the sequences which are required for plasmids to be replicated and retained in the host cell, the plasmids isolated from the resulting bacterial cultures should contain the desired mutation and other features of the plasmid template. In order to avoid large contamination by the template plasmid, it is cut at a unique restriction site in a region which will not be amplified in the original PCR. It may be possible to avoid contamination by using a plasmid template which was isolated from a *dut ung* strain.
of *E. coli* and which therefore contained uracil, as is used in the Kunkel method. Upon transformation of the RCPCR product into a host strain of *E. coli* which had functional *dut* and *ung* gene products, the template DNA should be degraded and replaced by the host cell.

Mutagenesis methods such as that described above require the use of two mutagenic primers. It is possible, however, to carry out site-directed mutagenesis using only one mutagenic primer (Chen & Przybyla, 1994; Sarkar & Sommer, 1990). In this case, a "megaprimer" is created by PCR using the mutagenic primer and another primer. Since the DNA product of this reaction is double stranded, it can be used (purified or not) as the primer in each of two other PCR amplifications (Figure 3.6).

![Figure 3.6. Creation of a Megaprimer Prior to RCPCR. Small triangles indicate the site of mutagenesis.](image)
Results and Discussion

Protein Sequence Alignments

Sequence alignments performed in our laboratory on the L-ribulose-5-phosphate 4-epimerase and Class II L-fuculose-1-phosphate aldolase (FucA) from *E. coli* produced results similar to those reported by Dreyer & Schulz (1993, 1996b). Whereas Dreyer & Schulz report 34% identity in the N-terminal domains, we found that these enzymes share 38% identity and 43% similarity. In addition, alignment using the computer programme SEQSEE (Wishart *et al.*, 1994) showed that H171 of AraD aligned with H155 of FucA and is therefore a likely candidate as the fourth ligand to the metal ion in the epimerase (the other three are D76, H95 and H97). Furthermore, these residues were found to be conserved among all the known sequences in the AraD/FucA family (Figure 3.7). Note that L-rhamnulose-1-phosphate aldolase, which was included in the alignment, did not align well, in agreement with a report by Moralejo *et al.* (1993) that this aldolase has evolved convergendy to FucA.

It was therefore decided to prepare mutants of the epimerase in which each of these residues was independently converted to an asparagine residue by site directed mutagenesis. These mutations should directly perturb the coordination sphere of the bound metal ion and affect both the enzyme's affinity for the metal and its catalytic constants.
Figure 3.7. Amino Acid Sequence Alignment. The conserved metal-binding residues are shown in red. Key: arad, L-ribulose-5-phosphate 4-epimerase; fuca, L-fuculose-1-phosphate aldolase; rhad, L-rhamnulose-1-phosphate aldolase; ecoli, Escherichia coli; salty, Salmonella typhimurium; haein, Haemophilus influenzae.

Site Directed Mutagenesis

Wishing to take advantage of the Kunkel method for site-directed mutagenesis, we investigated whether pRE1 could be used to propagate single stranded DNA. The vector portion of pRE1 is largely from the pBS phagemid, which allows ssDNA to be created and packaged into a phage head. Modifications of pBSTIM (the precursor to pRE1; see Chapter II) may have altered the F1 origin enough to render it useless. Therefore, it was necessary
to determine if a strand of pRE1 could be isolated using helper phage M13K07 before planning a mutagenesis strategy dependent on ssDNA. ssDNA was isolated from *E. coli* DH5αF' transformed with pRE1 and infected with helper phage M13K07 and visualized on an ethidium bromide stained 1.2% agarose gel. Sequencing reactions (in either direction) did not yield readable sequence. We were therefore unsure if a strand of pRE1 had been packaged into the phage since the amount of ssDNA from pRE1 could have been too small to give readable sequence. Therefore, it was decided to use other means of site-directed mutagenesis.

Site-directed mutagenesis was performed using the RCPCR technique on pRE1 isolated from transformed *E. coli* DH5αF', and the mutant plasmids pAJ1, pAJ2 and pAJ3 were constructed, encoding the H95N, H97N and D76N epimerases, respectively. Approximately eight colonies per transformation were checked for the presence of mutant plasmids, 25% of which contained the mutation. In each case the mutant enzymes were overexpressed at levels comparable to the wild-type enzyme. Attempts at preparing the H171N mutant were not successful. Both halves of the plasmid appeared to have been created by PCR, as visualized on an ethidium bromide stained 1.2% agarose gel. It is not clear why mutant plasmid could not be obtained after recombination and transformation. The plasmids containing the genes for the mutant enzymes were sequenced over the *araD* region. Only one error was found: pAJ3 had a base change at position 25 from C to G, which altered nucleotide 303 of the *araD* gene. This error was not from PCR but rather
from an error in designing the primer AJ3. Fortunately, this base change did not cause any change in the amino acid sequence.

**Purification of the Mutant Epimerases**

In order to avoid contamination by endogenous L-ribulose-5-phosphate 4-epimerase, the plasmids encoding the mutant (and wild type) enzymes were transformed into a strain of *E. coli* which does not have a functional 4-epimerase. Therefore, *E. coli* Y1090 (modified), which was used for preparation of the wild type epimerase in Chapter II, was used. As with the wild type enzyme, the mutant enzymes were expressed to a high level, accounting for at least 60% of the total soluble protein in crude cell extracts. After purification, approximately 100 mg of enzyme was obtained per litre of bacterial culture.

The molecular weights of the mutant enzymes were confirmed to be the same as those expected from their gene sequences by electrospray mass spectrometry (Table 3.1).

<table>
<thead>
<tr>
<th>4-Epimerase</th>
<th>Calculated Molecular Mass</th>
<th>Determined Molecular Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>H95N</td>
<td>25495</td>
<td>25499 ± 5</td>
</tr>
<tr>
<td>H97N</td>
<td>25495</td>
<td>25498 ± 5</td>
</tr>
<tr>
<td>D76N</td>
<td>25520</td>
<td>25524 ± 5</td>
</tr>
<tr>
<td>Wild Type</td>
<td>25520</td>
<td>25522 ± 5</td>
</tr>
</tbody>
</table>

*In Da.*
Since zinc is commonly used by enzymes (such as the Class II aldolases) to promote electrophilic catalysis and since the zinc form of the wild type enzyme displayed kinetic constants which agreed with those reported in the literature, it was decided to prepare and use homogeneous samples of the zinc form of the mutant enzymes for further studies. Apoenzymes were prepared by dialysis against EDTA using conditions reported previously (Deupree & Wood, 1972), and they were subsequently reconstituted with excess zinc. The zinc enzymes were extensively dialysed against metal-free buffer to remove excess metal.

**Physical Characterization of the Mutant L-Ribulose-5-Phosphate 4-Epimerases**

It is important to ascertain that the mutations have not drastically altered the structure of the mutant enzymes from that of the wild type enzyme. Kinetic and metal binding results will not be valid if the mutant enzymes do not fold into the proper structure.

The mutant enzymes were analysed using circular dichroism (CD) spectroscopy and were found to possess spectra virtually indistinguishable from that of the wild-type enzyme (Figure 3.8). The differences might be due to small variations in concentration. Intense double minima at 208 and 222 nm are characteristic of α-helical structures, as is an intense maximum at 192 nm (Johnson, 1990). The intense double minima observed in the wild type and mutant enzymes indicates that these enzymes have significant α-helical character. Unfolded protein, which can be approximated by a random coil structure, would be expected to have a small maximum at 220 nm and an intense minimum at 198 nm. This is not seen in the CD spectra of either the wild type or mutant enzymes.
Figure 3.8. CD Spectra of the Wild Type and Mutant 1-Ribulose-5-Phosphate 4-Epimerases.
In addition, the melting temperatures were measured by following the loss of ellipticity at 220 nm as a function of temperature. Each of the enzymes was found to have a melting temperature in the range of 49.5 to 54.5°C (Table 3.2). These observations support the notion that the mutants were fully folded proteins with a tertiary structure and stability comparable to that of the wild-type enzyme. It should be noted that these studies were performed on Zn$^{2+}$ reconstituted enzymes that had been dialysed to remove excess zinc. Subsequent studies showed that apoenzyme can form upon extended dialysis of the mutant enzymes. It would appear that the absence of zinc does not dramatically affect the CD spectra or melting temperatures of the mutant epimerases.

<table>
<thead>
<tr>
<th>4-Epimerase</th>
<th>Melting Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild Type</td>
<td>49.9 ± 0.3</td>
</tr>
<tr>
<td>H95N</td>
<td>49.5 ± 0.3</td>
</tr>
<tr>
<td>H97N</td>
<td>54.4 ± 0.3</td>
</tr>
<tr>
<td>D76N</td>
<td>52.2 ± 0.3</td>
</tr>
</tbody>
</table>

* Determined from the change in CD signal at 220 nm as a function of temperature, and corrected to a calibration curve.

To investigate the quaternary structure of the proteins, size exclusion chromatography was employed. The wild type and mutant enzymes eluted with similar retention times. A comparison to molecular weight standards indicated that they had masses of approximately
127 kDa. This shows that the mutant enzymes had retained the ability to form a quaternary structure similar to that of the wild type enzyme (Figure 3.9). Given a subunit molecular mass of 25.5 kDa (as determined from the amino acid sequence), the native protein appears to be pentameric. Size exclusion chromatography is not the most accurate method for determining native masses; globular proteins are not perfect uniformly packed spheres and exhibit different mobilities based on their shapes and volumes. Ideally the standards used should have the same shape as the sample being analysed. Molecular weight calculated by sedimentation equilibrium at high speed is more accurate because it depends only on density. This technique has been used to show that the native wild type enzyme has a mass of about 105000 ± 2000 Da (Gielow & Lee, 1975), which corresponds to a tetrameric structure.

Activity of the Wild Type and Mutant Enzymes

The kinetics of the zinc-reconstituted enzymes were measured in the L-ribulose-5-phosphate to D-xylulose-5-phosphate direction as described by Davis et al. (1972) and in Chapter II (for the wild type enzyme). As discussed in Chapter II, the $K_M$ value for LRu5P with the wild type enzyme correlates well with those previously published after zinc-reconstitution (Deupree & Wood, 1975; Gielow & Lee, 1975; Lee et al., 1968; Wolin et al., 1958), but not before this treatment. The mutant enzymes had much lower $k_{cat}$ values than the wild type epimerase while their $K_M$ values remained similar (Table 3.3). This might be expected for the alteration of residues which are not directly involved in substrate binding, and may be due to intrinsic changes in the active site or due to loss of the metal ion during
Figure 3.9. Native Structure of L-Ribulose-5-Phosphate 4-Epimerases. PK = pyruvate kinase, Ald = aldolase, LDH = lactate dehydrogenase, BSA = bovine serum albumin, Oval = ovalbumin, ADH = alcohol dehydrogenase, Cyt-C = cytochrome C, RPE = L-ribulose-5-phosphate 4-epimerases (wild type, H97N, H95N, and D76N), Ve = elution volume, Vo = void volume.
dialysis. For this reason, the mutant enzymes were assayed in the presence of excess (0.1 mM) Zn$^{2+}$ (Table 3.4).

**Table 3.3.** Kinetic Parameters of the Zinc-Reconstituted Wild Type and Mutant L-Ribulose-5-Phosphate 4-Epimerases. $^{a,b}$

<table>
<thead>
<tr>
<th>4-Epimerase</th>
<th>$K_M$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_M$ (M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild Type</td>
<td>0.087 ± 0.007</td>
<td>20.4 ± 0.9</td>
<td>$2.3 \times 10^5$</td>
</tr>
<tr>
<td>H95N</td>
<td>0.14 ± 0.01</td>
<td>0.047 ± 0.002</td>
<td>$3.4 \times 10^2$</td>
</tr>
<tr>
<td>H97N</td>
<td>0.10 ± 0.01</td>
<td>0.048 ± 0.001</td>
<td>$4.8 \times 10^2$</td>
</tr>
<tr>
<td>D76N</td>
<td>0.11 ± 0.01</td>
<td>0.073 ± 0.003</td>
<td>$6.6 \times 10^2$</td>
</tr>
</tbody>
</table>

$^a$ No additional Zn$^{2+}$ was added to the reaction cuvettes, nor were any metal ions removed from the solutions used for enzyme assays; transketolase, a vital component of the assay system also requires divalent metal ions.

$^b$ See Appendix D for a graphical representation of these kinetics.

The kinetic constants of the wild type enzyme were relatively unaffected by the presence of additional zinc. This is consistent with previous reports that the metal is bound tightly enough to survive dialysis against metal-free buffer. It also shows that excess zinc does not significantly activate the enzyme. The data obtained with the mutant enzymes were somewhat different: some of the lost activity could be recovered by addition of excess zinc to the cuvettes. Particularly notable are the increases in the $k_{cat}$ values which were observed in the case of the H97N epimerase. The $K_M$ values remained relatively unaffected, however. It would appear that varying amounts of bound metal were lost during dialysis of the mutant enzymes against metal-free buffer resulting in mixtures of the apoenzyme and holoenzyme.
This is consistent with the notion that these mutations have modified a residue involved in metal-ion complexation. Measurements made in the presence of 0.2 mM Zn\(^{2+}\) caused no further changes in the \(k_{\text{cat}}\) values indicating that 0.1 mM Zn\(^{2+}\) was sufficient to activate these mutants fully.

### Table 3.4. Kinetic Parameters of the Zinc-Reconstituted Wild Type and Mutant L-Ribulose-5-Phosphate 4-Epimerases In the Presence of Additional Zn\(^{2+}\). \(^a,b\)

<table>
<thead>
<tr>
<th>4-Epimerase</th>
<th>(K_M) (mM)</th>
<th>(k_{\text{cat}}) (s(^{-1}))</th>
<th>(k_{\text{cat}}/K_M) (M(^{-1}) s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild Type</td>
<td>0.060 ± 0.005</td>
<td>22.3 ± 0.4</td>
<td>3.7 (\times) 10(^5)</td>
</tr>
<tr>
<td>H95N</td>
<td>0.096 ± 0.009</td>
<td>0.099 ± 0.003</td>
<td>1.0 (\times) 10(^3)</td>
</tr>
<tr>
<td>H97N</td>
<td>0.14 ± 0.02</td>
<td>7.3 ± 0.3</td>
<td>5.2 (\times) 10(^4)</td>
</tr>
<tr>
<td>D76N</td>
<td>0.14 ± 0.01</td>
<td>0.162 ± 0.005</td>
<td>1.2 (\times) 10(^3)</td>
</tr>
</tbody>
</table>

\(^a\) 0.1 mM Zn\(^{2+}\) was added to each cuvette.
\(^b\) See Appendix D for a graphical representation of these kinetics.

In each case the \(k_{\text{cat}}\) values obtained with the mutants in the presence of 0.1 mM Zn\(^{2+}\) were lower than that of the wild type enzyme. With the histidine mutants, the lower \(k_{\text{cat}}\) values are consistent with the notion that the ligand sphere about the metal has been altered and that its ability to act as an electrophilic catalyst has been impaired. In the case of the aspartate residue (that is likely displaced from the metal upon substrate binding), the lower \(k_{\text{cat}}\) value may be due to the modification of a catalytic base.
Zinc Ion Content of the Zn\(^{2+}\) Reconstituted Epimerases

The mutant enzymes were analysed for zinc ion content by ICPMS after reconstitution and purification on a size exclusion column. A large decrease in the amount of bound metal was observed for the two histidine residues that were mutated, but not for the aspartate residue (Table 3.5). This observation supports the notion that the histidine residues serve as metal ligands. It is not clear why the zinc-reconstituted, dialysed, H95N and H97N enzymes still have measurable activity in the absence of added zinc. One possible explanation is that the assay solutions probably contain divalent metal ions; transketolase, one of the coupling enzymes, requires a divalent metal ion for activity. The mutant enzymes may be scavenging these metals and may have widely differing binding constants. This could explain why the H97N mutant had a great increase in activity in the presence of additional zinc but the H95N mutant did not.

<table>
<thead>
<tr>
<th>4-Epimerase</th>
<th>Zinc Ions per Subunit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild Type</td>
<td>1.05 ± 0.05</td>
</tr>
<tr>
<td>H95N</td>
<td>&lt;0.01 ± 0.05</td>
</tr>
<tr>
<td>H97N</td>
<td>0.14 ± 0.05</td>
</tr>
<tr>
<td>D76N</td>
<td>0.95 ± 0.05</td>
</tr>
</tbody>
</table>

*As determined by ICPMS.*
Chapter III: Metal Ion Ligands

The metal content of the D76N mutant epimerase remained at about the same level as the wild type enzyme, when purified by size exclusion chromatography. However, it seemed to lose metal ion during extensive dialysis as analysed by kinetics. Presumably the release of the metal ion occurs slowly and little is lost during the column purification, which is relatively rapid. The reduced $k_{\text{cat}}$ of this mutant indicates that D76 does have a role in catalysis. These observations might be consistent with the role of the corresponding amino acid residue (E73) in the aldolase, which rotates away from the zinc ion on binding of the substrate. Therefore, this residue might not be important to metal binding in the aldolase either. It would be interesting to see how the activity and metal binding ability of the aldolase would be affected by mutating this residue.

It would be necessary to analyse the kinetics and thermodynamics of metal complexation before these questions can be answered. $K_d$ values for metal binding could be determined using isothermal titration calorimetry on the apoenzymes.

Activity of the Cobalt-Substituted Epimerases

Zinc is spectroscopically invisible, so in order to study effects of the metal ligands on the metal-binding site, it was decided to prepare samples of the cobalt(II) form of the enzymes for these studies. As in preparation of the zinc forms, apoenzymes were prepared by dialysis against EDTA, and were subsequently reconstituted with excess cobalt, which was removed by extensive dialysis against metal-free buffer.
Kinetic studies of the Co$^{2+}$-substituted wild type enzyme showed that it had a higher $k_{\text{cat}}$ than the Zn$^{2+}$-substituted enzyme (Table 3.6). This is not an unusual effect of cobalt-substitution (Maret & Vallee, 1993), and has also been observed for L-fuculose-1-phosphate aldolase (Dreyer & Schulz, 1996b). Unlike the zinc form of the enzymes, the $K_M$ values of the mutant enzymes were affected by the metal ion substitution and are approximately twice the $K_M$ value of the Co$^{2+}$-substituted wild type enzyme. The origin of the higher $K_M$ values observed for the cobalt-substituted mutant enzymes is not known. The true extent of the effects of Co$^{2+}$ substitution on the kinetics of the wild type and mutant enzymes is not known since the amount of bound Co$^{2+}$ was not determined, nor were the kinetics studied in the presence of additional Co$^{2+}$.

<table>
<thead>
<tr>
<th>4-Epimerase</th>
<th>$K_M$ (mM)</th>
<th>$k_{\text{cat}}$ (s$^{-1}$)</th>
<th>$k_{\text{cat}}/K_M$ (M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild Type</td>
<td>0.093 ± 0.009</td>
<td>26.9 ± 0.8</td>
<td>2.9 × 10$^5$</td>
</tr>
<tr>
<td>H95N</td>
<td>0.28 ± 0.03</td>
<td>0.10 ± 0.004</td>
<td>3.6 × 10$^2$</td>
</tr>
<tr>
<td>H97N</td>
<td>0.20 ± 0.02</td>
<td>0.082 ± 0.003</td>
<td>4.2 × 10$^2$</td>
</tr>
<tr>
<td>D76N</td>
<td>0.23 ± 0.02</td>
<td>0.11 ± 0.003</td>
<td>4.6 × 10$^2$</td>
</tr>
</tbody>
</table>

$^a$ These were determined in the same manner as the Zn$^{2+}$-substituted enzymes without any addition of extra divalent metal ions to the cuvettes.

$^b$ See Appendix D for a graphical representation of these kinetics.
UV-Visible Spectra of the Co(II) Substituted Enzymes

UV-visible difference spectra were collected of the wild type and mutant Co\(^{2+}\)-substituted 4-epimerases. There was no significant absorbance in wild type or mutant enzymes in the 500 to 700 nm range, suggesting that no cysteine residues were acting as metal-ion ligands (none were expected). Unfortunately, little information could be obtained from the UV-visible spectra of the cobalt-substituted enzymes. This is not an unknown event; cobalt(II)-substituted phosphotriesterase had a similar UV-visible spectrum to that of the 4-epimerases until one of its metal ion ligands was mutated to methionine (a sulphur donor) (Kuo et al., 1997).

Co\(^{3+}\)-Modification of the Co\(^{2+}\)-Substituted Wild Type 4-Epimerase

Co(III)-modification of Co\(^{2+}\)-substituted enzymes has been used to identify metal ion ligands (Hlavaty & Nowak, 1997). In an effort to see if this technique could be used to identify the metal ion ligands of L-ribulose-5-phosphate 4-epimerase, the Co\(^{2+}\) form of the wild type enzyme was modified to the Co\(^{3+}\) form.

After hydrogen peroxide oxidation of the Co\(^{2+}\)-substituted 4-epimerase, the mass of the modified epimerase was found to be 25524 Da. The mass of the unmodified epimerase has been experimentally determined to be 25524 Da (Chapter II). There was a small shoulder peak at 25581 Da (Figure 3.10), which corresponds to an addition of 59 Da to the wild type epimerase; this is the mass of cobalt. It was not clear if this peak was indeed from
Figure 3.10. Deconvoluted Electrospray Ionization Mass Analysis of the Co(III) Wild Type 4-Eptimase.
addition of Co(III) to the enzyme or an impurity in the enzyme preparation, or if it was an artifact of hydrogen peroxide oxidation of the enzyme.

Pursuit of the Co(III)-modified wild type epimerase may be worthwhile for a definitive identification of the metal-ion ligands. Better modification may be achieved with a longer hydrogen peroxide incubation; times of up to twelve hours are not unknown (Van Wart, 1988), and thirty minutes might be a little short. In addition, it might be interesting to see how the activity of the epimerase is affected by oxidation of Co(II) to Co(III).

**Conclusion**

This work provides evidence that there is a structural link between the L-ribulose-5-phosphate 4-epimerase from *E. coli* and the Class II L-fuculose-1-phosphate aldolase from *E. coli*. The aldolases have traditionally been characterized as either Class I in which a lysine residue forms a Schiff base with a carbonyl of the nucleophilic substrate or as Class II in which a metal ion serves to stabilize the enediolate intermediate (Horecker *et al.*, 1972; see Chapter I).

Mutation of three of the four putative metal ion ligands of L-ribulose-5-phosphate 4-epimerase caused drastic reductions in the activities of these enzymes. These ligands were postulated to be metal binding based on homology between the 4-epimerase and a bacterial L-fuculose-1-phosphate aldolase, whose crystal structure has been solved (Dreyer & Schulz,
Chapter III: Metal Ion Ligands

Mutagenesis of these residues affected the metal-binding abilities of the enzyme, indicating that these residues are involved in ligation of the divalent metal ion known to be required for catalysis. A reduction in $k_{cat}$ was observed for all three mutant enzymes. In the case of the histidine mutants, this can be explained by having changed the coordination sphere of the metal ion. In the case of the aspartate mutant, this residue, in addition to being a metal ion ligand might also be a catalytic base in the active site. By analogy to X-ray crystal structures of the aldolase (Dreyer & Schulz, 1993; Dreyer & Schulz, 1996a; Dreyer & Schulz, 1996b), this residue might bind the metal ion in the uncomplexed enzyme and then rotate away upon substrate binding, and could then participate in catalysis. Since the active site ligands to the divalent metal ion in the epimerase and the aldolase are conserved, a mechanism very similar to that of the aldolase could be employed by the epimerase provided that the attack on the aldehyde is non-stereospecific and the bound enediolate is protected from any proton source (Path B; see Chapter I).

It is of interest to note that recent additions to the Swiss Protein Data Bank have added an annotation that the D76, H95, H97, and H171 residues of L-ribulose-5-phosphate 4-epimerase are probable zinc-ion ligands based on the similarity of the sequences of the 4-epimerase to L-fuculose-1-phosphate aldolase. Furthermore, with increasing numbers of bacterial genome sequences being reported, new sequences of enzymes are continually being added to the AraD/FucA family. The functions of these proteins are not known, but they are thought to be involved in pentose metabolism pathways. The N-terminal sequences of
each of these proteins share homology in a “catalytic domain”, and in each case the known zinc-ligands of L-fuculose-1-phosphate aldolase are conserved.

Experimental Methods

Data Base Searches and Protein Sequence Alignments

The Swiss Protein Data Bank was searched for protein sequences similar to that of L-ribulose-5-phosphate 4-epimerase. SEQSEE (Wishart et al., 1994) was used to align the protein sequences of the L-ribulose-5-phosphate 4-epimerases of *Escherichia coli*, *Salmonella typhimurium*, and *Haemophilus influenzae*, L-fuculose-1-phosphate aldolases of *E. coli* and *S. typhimurium*, and *E. coli* L-rhamnulose-1-phosphate aldolase.

Bacterial and Phage Strains

*E. coli* DH5αF’ was used for preparation of plasmid DNA for site-directed mutagenesis and sequencing. This strain was also used in the ssDNA rescue experiments. *E. coli* Y1090 (modified) was used (as described in Chapter II) for mutant enzyme overexpression to avoid any contamination from low levels of endogenous L-ribulose-5-phosphate 4-epimerase production. Helper phage M13K07 was used in the ssDNA rescue experiments.
**ssDNA Rescue of pRE1**

LB broth (25 mL) with 50 µg/mL ampicillin was co-inoculated with *E. coli* DH5αF' (transformed with pRE1) and helper phage M13K07, and incubated in a 37°C shaker overnight. Cells were removed from the growth medium by centrifugation at 5000 rpm in a Sorvall GSA rotor for 10 minutes. To the supernatant was added 2.5 mL PEG/NH₄OAc (50:50 vol 50% PEG 8000: 7.5 M NH₄OAc). This mixture was incubated on ice for 30 minutes and then centrifuged for 20 minutes at 10000 rpm in a Sorvall SS34 rotor. The supernatant was removed (and discarded), and the precipitated phage pellet was redissolved in 400 µL sterile distilled water. The phage suspension was extracted twice with 400 µL P/C/I (25:24:1 vol phenol: chloroform: isoamyl alcohol) and then once with 400 µL chloroform. NH₄OAc (20 µL, 3 M) and isopropanol (1 mL) were added to the aqueous phase and mixed. After 15 minutes of incubation at room temperature, ssDNA was collected by a 10 minute centrifugation in a microfuge at 4°C. Recovery of ssDNA was checked by 1.2% agarose gel electrophoresis.

ssDNA was used in sequencing reactions with Sequenase version 2.0 DNA Sequencing Kit and methods suggested by the supplier (United States Biochemical) for dideoxynucleotide chain termination sequencing of single stranded DNA.

**Site-Directed Mutagenesis**

Primers for each mutation (Table 3.6) were designed so as to include a silent mutation that introduces a new restriction site in the araD gene. This allows for quick screening of
Chapter III: Metal Ion Ligands

putative mutants. Figure 3.11 shows the relative primer binding locations and the restriction sites which are present on pRE1 and those which are introduced by mutagenesis.

**H95N.** A megaprimer was formed according to the method of Chen and Przybyla (1994) in a first round of PCR with pRE1 serving as the template and primers MT05 and AJ1. AJ1 contained the desired mutation and a silent mutation introducing a *Hpa* I restriction site. The first round of PCR contained 0.1 mM of each dNTP, 10 mM Tris·HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl₂, 0.5 μM of each primer, ca 10 ng of pRE1 template and 2.5 U *Taq* DNA polymerase (Gibco/BRL) in 100 μL. 100 μL mineral oil was added to prevent evaporation during thermocycling. An initial melt at 95°C for 1 minute was followed by 25 cycles of 95°C (30 sec), 45°C (30 sec), and 72°C (1 min), and ended with a final extension of 7 minutes at 72°C. The megaprimer was purified using Wizard PCR Preps DNA Purification System (Promega), then treated with *Nde* I to remove DNA from MT05 which was not complementary to the plasmid pRE1. The resulting DNA was used as the primer in each of two subsequent rounds of PCR to create two halves of the plasmid, as described by Jones and Winistorfer (1992). Reactions were run with the megaprimer and MT08 on pRE1 (linearized with *Pst* I), and with MT09 (linearized with *EcoR* I). The second round of PCR was similar to the first round except that there were 28 cycles of PCR, 0.5 μM of MT08 or MT09 was used with about 5% of the megaprimer from one PCR, treated with *Nde* I. The two halves were then mixed together, thermally denatured (94°C, 3 min), annealed (50°C, 2 h), and then transformed into competent *E. coli* DH5αF'. Clones were
Table 3.6. Primers Used for Site-Directed Mutagenesis of the araD Gene.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence $^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AJ1 $^b$</td>
<td>5'-GCGAGTGCGTGTTAA4CAATGCGCCGGCAATGGA-3'</td>
</tr>
<tr>
<td>AJ2 $^c$</td>
<td>5'-CCTGGCCGCGCAAATCGATGCACAAATGCGCCG-3'</td>
</tr>
<tr>
<td>AJ3 $^d$</td>
<td>5'-TCTGCATACGA4ATCTCGCGCCACGCGACCA-3'</td>
</tr>
<tr>
<td>AJ4 $^e$</td>
<td>5'-AAAAGCCCTCTCGGA4ACACGCAACTCACCAGG-3'</td>
</tr>
<tr>
<td>AJ5 $^f$</td>
<td>5'-GCTGTCGCTGTTCTGA4AGAGGTTCCTTTGCTGACC-3'</td>
</tr>
<tr>
<td>AJ6 $^g$</td>
<td>5'-TGGTCCATTAGATCCGA4CGGCCCCGTGATTCCAT-3'</td>
</tr>
<tr>
<td>AJ7 $^h$</td>
<td>5'-AACGGGCCCTGCTGA4TGGACCAGACG-3'</td>
</tr>
<tr>
<td>MT08 $^i$</td>
<td>5'-CGACCGAGCGTGACACCACGATGCC-3'</td>
</tr>
<tr>
<td>MT09 $^j$</td>
<td>5'-GCAGAGCGAGGTATGTCGTAGGCGGTCG-3'</td>
</tr>
</tbody>
</table>

$^a$ Mismatches are shown in bold, new restriction sites in italics.

$^b$ This primer, used to make the H95N mutation found in pAJ1, is complementary to the sense strand of araD between nt 265 and 295, and introduces a Hpa I restriction site.

$^c$ AJ2 is complementary to the sense strand of araD between nt 273 and 301, and introduces an EcoR I restriction site as well as the H97N mutation found in pAJ2.

$^d$ This primer was used in construction of the H97N mutation of pAJ2. It is identical to the sense strand of araD between nt 279 and 307, and, like AJ2, introduces an EcoR I restriction site and the H97N mutation.

$^e$ Identical to the sense strand of araD between nt 212 and 243, this primer introduces the D76N mutation and a Csp45 I restriction site found in pAJ3.

$^f$ This primer, which binds to the sense strand between nt 205 and 247, was used to create pAJ3. This plasmid encodes the D76N mutation of L-ribulose-5-phosphate 4-epimerase, and has a new Csp45 I restriction site.

$^g$ AJ6 binds to the antisense strand of araD between nt 526 and 500, and introduces an H171N mutation along with a new Csp45 I restriction site.

$^h$ Binding to the sense strand of araD between nt 495 and 521, AJ7 introduces a H171N mutation and a new Csp45 I restriction site.

$^i$ MT08 binds to the + strand of pRE1 (contains the anti-sense strand of araD) between nt 3071 and 3095 (see Appendix A).

$^j$ MT09 binds to the - strand of pRE1 (contains the sense strand of araD) between nt 2365 and 2388 (see Appendix A).
Figure 3.11. A Schematic Diagram of pRE1 Indicating Relative Primer Binding Sites and Locations of Restriction Sites on pRE1 and Restriction Sites Introduced by Site Directed Mutagenesis.
poured and grown up in 5 mL cultures of LB with 50 μg/mL ampicillin at 37°C. Plasmids were purified and screened for the desired mutation by digestion with both *Hpa* I and *EcoR* I. The construct pAJ1 was found to contain the desired mutation by sequencing of the mutant *araD* gene.

**H97N.** A plasmid encoding the H97N mutant was not readily made using AJ2 and MT05 following the same method as for H95N. Instead, two primers (AJ2 and AJ3) complementary to each other and both encoding the desired mutation and a silent mutation introducing a new *EcoR* I restriction site were used. These primers were used in Jones and Winistorfer’s (1992) recombinant circle PCR (RCPCR) technique. To obtain one half of the plasmid, pRE1 (linearized with *Pst* I) was amplified by PCR using primers AJ2 and MT08, while the other half was obtained by PCR amplification of pRE1 (linearized with *EcoR* I) and primers AJ3 and MT09. Each amplification went through 28 cycles of PCR as described for formation of the H95N mutant, and each reaction contained in 100 μL: 0.5 μM of each primer, 0.1 mM each dNTP, ca. 10 ng of template DNA, 10 mM Tris·HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl₂, and 2.5 U Taq DNA polymerase. 100 μL of mineral oil was used to prevent evaporation during thermocycling. As with H95N, the two plasmid halves were annealed and transformed into *E. coli* DH5αF'. Plasmids isolated from saturated (overnight) cultures of resulting colonies were screened for the desired mutation by digestion with *EcoR* I. pAJ2 contained the desired mutation, and its sequence was confirmed by DNA sequencing.
**D76N.** The construct pAJ3, which contained the *araD* gene encoding the desired mutation was created using the same RCPCR method as was used for H97N. In this case, however, a silent mutation was used which introduced a *Csp45* I restriction site. The two plasmid halves were created by PCR amplification of pRE1 (linearized with *EcoR* I) and primers AJ4 and MT09, and of pRE1 (linearized with *Pst* I) and primers AJ5 and MT08. Plasmids were screened by restriction digestion with both *Csp45* I and *EcoR* I. The sequence of *araD* was verified by DNA sequencing.

**H171N.** Primers encoding this mutation also included a silent mutation introducing a *Csp45* I restriction site, and were used in the same RCPCR procedure that was used for creating H97N and D76N. In this case primers AJ6 and MT09 were used for PCR amplification of pRE1 (linearized with *EcoR* I), and primers AJ7 and MT08 were used to amplify pRE1 (linearized with *Pst* I). Colonies resulting from the transformation of *E. coli* DH5αF' with annealed plasmid halves (the two PCR products) were used to inoculate overnight cultures of LB broth containing 50 μg/mL ampicillin and grown overnight in a 37°C shaker. Plasmids isolated from these cultures were screened for restriction by both *EcoR* I and *Csp45* I.
DNA sequencing

Sequencing reactions were performed on the ssDNA (isolated as described above) using primers MT12 and MT15 to determine if single stranded pRE1 (containing the araD gene) was among the ssDNA isolated. These reactions were performed in the same manner as those performed on plasmid DNA except that there was no denaturation step. The araD gene of each mutant plasmid was sequenced in entirety using the Sequenase version 2.0 DNA sequencing kit (United States Biochemical) to ensure that no spurious mutations had occurred during site-directed mutagenesis. Double-stranded plasmid DNA was isolated for this purpose using the Wizard Minipreps DNA Purification System (Promega). The primers used were the same as those used to sequence the wild-type araD gene of pRE1 as described earlier. In addition, the araD gene of pAJ3 was sequenced by Debbie Neufeld in the Nucleic Acid and Protein Service at the University of British Columbia using primer AJ10 (5'-CGAACACGCCCAACTCACC-3') and Applied Biosystems Inc. PRISM technology.

Enzyme Purification

Mutant L-ribulose-5-phosphate 4-epimerases were purified using a modification of the method described by Lee et al. (Gielow & Lee, 1975; Lee et al., 1968), and in the same manner as the wild-type enzyme (Chapter II).

For the purification of the mutant enzymes, pAJ1 (for H95N), pAJ2 (for H97N), and pAJ3 (for D76N) were individually transformed into E. coli Y1090 (modified), which lacks a functional 4-epimerase. The ammonium sulphate precipitation step was omitted with the
H97N mutant as this step led to large losses of this mutant. Otherwise, the purification procedure used is described in detail in Chapter II.

Protein determination was performed by measuring the absorbance of protein samples at 280 nm and using the extinction coefficient determined for the wild type enzyme (Chapter II), \( \varepsilon = 1.73 \text{ mL mg}^{-1} \text{ cm}^{-1} \), or \( \varepsilon = 33717 \text{ M}^{-1} \text{ cm}^{-1} \) per subunit.

**Calculation and Determination of Mutant Enzymes' Subunit Masses**

As with the wild-type enzyme, the masses of the mutant enzymes were calculated based on their amino acid content. The masses were determined experimentally by electrospray ionization mass spectrometry (performed by David Chow) as described in Chapter II.

**Determination of Tetrameric Structure of the Native Enzymes**

The tetrameric nature of the mutant enzymes was established by gel filtration chromatography with a Waters HPLC and a Protein-Pak 300 SW column (Waters). A gel filtration calibration curve was constructed using pyruvate kinase (237 kDa tetramer), aldolase (158 kDa tetramer), rabbit muscle lactate dehydrogenase (140 kDa monomer, bovine serum albumin (66 kDa monomer), and ovalbumin (43 kDa monomer). Blue Dextran was used to determine the void volume of the column. Standard solutions were made in 100 mM phosphate buffer (pH 7.0) at concentration of 1 mg/mL. Wild type and mutant L-ribulose-5-phosphate 4-epimerase were also prepared in 100 mM phosphate buffer.
Chapter III: Metal Ion Ligands

(pH 7.0) at 1 mg/mL. Injections of 50 μL of each sample and standard were made at a flow rate of 1 mL/min.

Preparation of Metal-Free Buffers and Glassware

Metal-free buffer was prepared by passing a solution of 50 mM HEPES (2 L, containing 10% glycerol) through a column of Chelex-100 resin (70 mL, Na⁺ form, 50-100 mesh, changed to H⁺ form by washing with 2 volumes 1 N HCl and rinsing with 5 volumes deionized H₂O). The pH was adjusted to 7.6 using solid Tris base. Adventitious metal ions were removed from glassware and plasticware used in this procedure by soaking them overnight in 4 N HCl and then washing them thoroughly with deionized water (17.8 MΩ/cm at 25°C).

Preparation of Apoenzyme and Reactivation with Zn(II) or Co(II)

Apoenzymes were prepared in the same manner as described for the wild type enzyme in Chapter II. 5 mg of the purified mutant (and wild type) enzymes were incubated with 20 mM EDTA in a total volume of 2.5 mL at room temperature for 3 hours. EDTA and any metal ions chelated to it were removed by dialysis against two 500 mL portions of metal-free HEPES/Tris buffer (pH 7.6) containing 10% glycerol. The apoenzymes were reconstituted by addition of 10 equivalents of ZnCl₂ (99.99%) to a concentration of 0.4 mM and incubating at room temperature for about 2 hours. Excess metal ions were removed by dialysis against two 500 mL portions of metal-free HEPES/Tris buffer.
Alternatively, the apoenzymes (of the wild type and mutant enzymes) at a concentration of between 2 and 2.5 mg/mL were reconstituted with the addition of 10 equivalents of a solution of CoCl$_2$ (99.999%, 250 μL of a 4 mM solution) to a concentration of 0.4 mM and incubating the mixture at room temperature for about 2 hours. Excess metal ions were removed by dialysis against two 500 mL portions of metal-free 50 mM HEPES/Tris (pH 7.6) buffer.

**Circular Dichroism and Thermal Stability**

Solutions of the zinc-substituted enzymes were exchanged into 10 mM potassium phosphate buffer, pH 7.6 using Amicon centricon-10 concentrators. These solutions were diluted to 0.17 mg/mL and were scanned in the far UV (300 to 190 nm) on a Jasco J-720 spectropolarimeter.

The thermal stabilities of the wild-type and mutant epimerases were determined by observing the change in CD signal at 220 nm in 50 mM HEPES/Tris buffer, pH 7.6 as a function of temperature. The temperature of the sample was raised from 30°C to 75°C at a rate of 50°C/hour using a NESLAB M-RS-232 bath/computer interface, a NESLAB RTE-111 waterbath and a NESLAB RS-2 remote sensor. These experiments were performed in the laboratory of Dr. Grant Mauk in the Department of Biochemistry at the University of British Columbia. The data were analysed by fitting them to an IC$_{50}$ curve using the computer programme GraFit.
Chapter III: Metal Ion Ligands

**Determination of Zinc(II) Content**

As described for the wild type enzyme in Chapter II, samples of the dialysed Zn-reconstituted enzymes were passed through a size exclusion column (Waters Protein Pak 125) at 0.5 mL/min to remove any excess metals remaining after dialysis. These samples were diluted 1/400 in 10% nitric acid containing 20 ppb Sc as an internal standard, and allowed to sit at room temperature at least overnight. A sample of the column buffer was also treated in this way. These samples were analysed for Zn content by ICPMS on a VG Elemental Plasma Quad mass spectrometer which had been calibrated with a standard curve of Zn (0, 5, 15, 40 ppb, prepared in 10% nitric acid containing a 1/400 dilution of the elution buffer and 20 ppb Sc as an internal standard) by Bert Mueller in the Department of Oceanography at the University of British Columbia.

**Assay for L-ribulose-5-phosphate 4-epimerase activity**

The kinetic parameters for each of the Zn(II)- and Co(II)-reconstituted enzymes were measured by monitoring the absorbance at 340 nm on a Cary 3E spectrophotometer at 37°C, as described in Chapter II, but with some minor modifications. In these assays, performed in 0.5 mL quartz cuvettes, each reaction contained, in a total volume of 500 μL, 25 mM glycylglycine (pH 7.6), 5 mM ribose-5-phosphate, 0.1 mM TPP, 0.15 mM NADH, LRu5P between 0.02 and 1 mM, 2.5 U αGDH, 25 U TIM, and 0.125 U transketolase. The reactions were initiated by addition of 0.025 μg of wild type or 0.25 μg of mutant 4-epimerase.
Additionally, the kinetic parameters were determined for the Zn(II) reconstituted enzymes in the presence of 0.1 mM Zn$^{2+}$ in each 0.5 mL cuvette. In this case, each reaction was initiated by the addition of 0.025 μg of wild type epimerase, 0.25 μg of the H95N or D76N epimerases, or 0.050 μg or the H97N epimerase.

Kinetic parameters were determined from a direct fit of the data to an enzyme kinetics equation using the computer program GraFit (Erithacus Software, UK). Errors were obtained using GraFit, which performs a non-linear regression using the method of Marquart (1963).

**Assay for Apoenzyme Activity**

Apoenzyme at 1 mg/mL (mutant epimerases) or 0.01 mg/mL (wild type epimerase) was incubated at 37°C for 5 minutes in metal-free HEPES/Tris buffer, pH 7.6, containing 9.44 mM LRu5P. Controls were water and untreated enzyme. 10 μL aliquots were immediately added to pre-equilibrated cuvettes containing (in 990 μL) 25 μmoles glycylglycine (pH 7.6), 0.15 μmole NADH, 5 μmoles DR5P, 0.1 μmole TPP, 5 U αGDH, 50 U TIM and 0.25 U transketolase. The activity of the apoenzyme was compared to the activity of the untreated enzyme and water control over one minute of reaction in the cuvette, as described in Chapter II.
UV-Visible Spectra of the Co(II)-Substituted 4-Epimerases

UV-visible spectra between 350 and 700 nm of the wild type and mutant enzymes were collected on a Cary 3E UV-Vis spectrophotometer at 25°C. Difference spectra were obtained by subtracting the UV-Vis spectrum of the Zn(II) containing epimerase from that of the Co(II)-substituted epimerase. In addition, spectra were buffer subtracted. The protein solutions used were all at 200 μM (based on subunits) in 50 mM HEPES/Tris buffer, pH 7.6.

Co(III) Modification of the Co(II)-Substituted Wild Type L-Ribulose-5-Phosphate 4-Epimerase

This procedure was a variation of the Co(III) modification of phosphoenolpyruvate carboxykinase reported by Hlavaty & Nowak (1997). Hydrogen peroxide (3%; 22 μL) was added to Co(II) substituted wild type epimerase at 61.5 μM (1.57 mg/mL; 98 μL) in 50 mM HEPES/Tris buffer (pH 7.6, containing 10% glycerol) to give a final concentration of 20 mM H₂O₂. This solution was incubated on ice for 30 minutes with occasional agitation. After incubation, the reaction was diluted to 2 mL with 50 mM HEPES/Tris buffer (pH 7.6), and excess hydrogen peroxide was removed by passage through an Amicon centrificon-10 filter unit at 5000 rpm in a Sorvall GSA rotor for 1 h. The final enzyme concentration, as determined by the absorbance at 280 nm was 2.6 mg/mL. The mass of the modified enzyme was determined by electrospray ionization mass spectrometry performed by Mr. Shouming He.
Chapter IV

L-Ribulose-5-Phosphate 4-Epimerase Follows a Retroaldol/Aldol Mechanism

Introduction

A structural relationship between *Escherichia coli* L-ribulose-5-phosphate 4-epimerase and the Class II L-fuculose-1-phosphate aldolase of *E. coli* was established in Chapter III. In particular, the active site zinc-binding ligands were found to be conserved in these two enzymes. Given this similarity, it is reasonable to suspect that these enzymes also share a related mechanistic strategy. L-Fuculose-1-phosphate aldolase uses an aldol (or retroaldol in the reverse direction) reaction to form L-fuculose-1-phosphate from dihydroxyacetone phosphate and L-lactaldehyde (Figure 4.1A). A similar mechanism can be drawn for the epimerase (Figure 4.1B) in which an enzyme-catalysed retroaldol reaction breaks the carbon-carbon bond of L-ribulose-5-phosphate between C-3 and C-4, creating the
Figure 4.1. Class II L-Fuculose-1-Phosphate Aldolase (A) and L-Ribulose-5-Phosphate 4-Epimerase (B) Can Utilise Similar Mechanistic Strategies.
enediolate of dihydroxyacetone and glycolaldehyde phosphate. If the enediolate adds back to the opposite face of the aldehyde in an aldol condensation, D-xylulose-5-phosphate is formed.

The best proof of an enzymatic pathway is identification of the reaction intermediates (if there are any). Unfortunately it is not always possible to isolate and characterize the intermediates: they may not be released from the enzyme’s active site; they may be unstable outside the enzyme’s active site; or, if they are released, the equilibrium constant of their release may be sufficiently unfavourable that only very small quantities are ever present in solution. An alternative approach is to supply the proposed intermediates in solution to the enzyme. If the enzyme is able to bind these molecules into its active site, they may be treated by the enzyme in exactly the same manner as the regularly formed intermediates and equilibrated into the normal ratio of substrate and product.

In the case of L-ribulose-5-phosphate 4-epimerase, release of the intermediates formed during an aldol mechanism would result in the production of dihydroxyacetone and glycolaldehyde phosphate. Dihydroxyacetone is the protonated form of the enediolate

![Figure 4.2. Reduction of Dihydroxyacetone to Glycerol by Glycerol Dehydrogenase.](image)
intermediate, and can be detected by coupling the release reaction to glycerol dehydrogenase, an enzyme which reversibly reduces dihydroxyacetone to glycerol with concomitant oxidation of NADH to NAD$^+$ (Burton & Kaplan, 1953; Spencer et al., 1989; Scharschmidt et al., 1983) (Figure 4.2). It is feasible that a small fraction of the time the bound enediolate could accept a proton and the resulting dihydroxyacetone could be displaced from the catalytic Zn$^{2+}$ ion and released from the enzyme.

Alternatively, the coupling of dihydroxyacetone and glycolaldehyde phosphate would produce the epimeric pentose phosphates. The formation of these sugars could be detected using a stopped form of the normal epimerase assay (described in Chapter II). In coupling dihydroxyacetone and glycolaldehyde phosphate to form an epimeric mixture of the ketopentose phosphates, the enzyme must overcome a number of obstacles. The aldehyde in solution is primarily in its hydrated form, and would have to be dehydrated before an aldol condensation could occur. A large proportion of dihydroxyacetone (60 to 80%) is also present in aqueous solution as the hydrate. Furthermore, the enzyme must have a means of deprotonating dihydroxyacetone to form the proposed enediolate intermediate.
Results and Discussion

*Intermediate Identification with the Wild Type L-Ribulose-5-Phosphate 4-Epimerase*

Initial studies were directed toward investigating whether or not the wild-type epimerase was able to promote carbon-carbon bond cleavage by observing low levels of aldolase activity. To begin with, $^1$H-NMR (200 MHz) was used to detect formation of LRu5P and DXu5P from dihydroxyacetone and glycolaldehyde phosphate. The very small amounts of sugar formed in the coupling reaction, relative to the amounts of dihydroxyacetone and glycolaldehyde phosphate, made it very difficult to be certain that coupling had occurred. Additionally, these coupling reactions were performed in a phosphate buffer, and it appeared that the phosphate buffer itself may have catalysed the aldol reaction to some extent. $^1$H-NMR was also used to observe the release of dihydroxyacetone and glycolaldehyde phosphate from the enzyme with an equilibrating pool of LRu5P and DXu5P. Again, it was not certain that dihydroxyacetone and glycolaldehyde phosphate were observed. It was decided that a coupled assay for release of dihydroxyacetone or glycolaldehyde phosphate might be more sensitive. The chance observation that glycolaldehyde phosphate could act as a substrate (albeit a poor one) for $\alpha$-glyceraldehyde phosphate dehydrogenase ($\alpha$ GDH) provided a means of observing the release of this proposed intermediate. These reactions were initially followed for a short time by observing a change in the absorbance at 340 nm, but no dihydroxyacetone or glycolaldehyde phosphate
could be detected. Further incubation, periodically measuring the absorbance at 340 nm, still
did not detect formation of dihydroxyacetone or glycolaldehyde phosphate. Following the
release of dihydroxyacetone or glycolaldehyde phosphate continuously for up to 12 hours
still did not provide any evidence for intermediate release.

Eventually it was decided that stopped assays to follow the aldol coupling of
dihydroxyacetone and glycolaldehyde phosphate over a long time (up to 72 hours) might
yield better results. The wild type enzyme was therefore assayed for aldolase activity
(carbon-carbon bond formation) by extensive incubation of a large amount of enzyme (0.5
mg/mL) with dihydroxyacetone (20 and 50 mM), glycolaldehyde phosphate (5 mM) and
ZnCl₂ (0.1 mM) at pH 7.6. At timed intervals, aliquots were removed and the total amount
of lRu5P and dXu5P was determined using a stopped assay for dXu5P using the coupled
system described for epimerase kinetics in Chapter II. This assay normally detects only
dXu5P; however, the large amounts of epimerase present ensured rapid equilibration
between the epimers and allowed determination of both epimers. This assay could not be
run continuously since very long incubation times were employed and a small background
rate was observed when glycolaldehyde phosphate itself was incubated in the presence of the
coupling enzymes and NADH. The results showed that low levels of the pentose phosphate
epimers were produced under these assay conditions; however, it was difficult to
demonstrate unambiguously that they were not a result of a background buffer-catalysed
reaction.
Equilibrium constants observed for related aldolase-catalysed reactions, where $K_{eq} = [\text{DHAP}] [\text{aldehyde}] / [\text{sugar}]$, are $1 \times 10^{-4}$ M for fructose-1,6-bisphosphate aldolase (Wong et al., 1995), $4.6 \times 10^{-4}$ M for FucA (Ghalambor & Heath, 1962; Ghalambor & Heath, 1966), and $1 \times 10^{-3}$ M for tagatose-1,6-bisphosphate aldolase (Fessner & Eyrisch, 1992). If the equilibrium constant for the aldolase activity of L-ribulose-5-phosphate 4-epimerase lies in this range, then far more than 0.4 mM ketopentose phosphate should have been formed from 50 mM dihydroxyacetone and 5 mM glycolaldehyde phosphate. It was therefore of interest to determine if the low conversion of dihydroxyacetone and glycolaldehyde phosphate to LRu5P and DXu5P could be a result of "product" inhibition. The aldol substrates did not appear to bind well to the epimerase since the epimerisation of L-ribulose-5-phosphate (0.094 mM) was not significantly inhibited in the presence of 0.1 mM Zn$^{2+}$ by either dihydroxyacetone (50 mM) or glycolaldehyde phosphate (5 mM). The rate was within 5% of the rate of the enzyme assayed in the absence of dihydroxyacetone and glycolaldehyde phosphate. In addition, a combination of these molecules (at 50 and 5 mM, respectively) did not cause any synergistic inhibition of the epimerase in the presence of 0.094 mM L-ribulose-5-phosphate and 0.1 mM Zn$^{2+}$. The enzyme therefore has relatively high $K_M$ values for dihydroxyacetone and glycolaldehyde phosphate. The aldol products, Ru5P and Xu5P, are therefore good competitive inhibitors of this reaction and prevent the equilibrium from being reached within 48 hours. In addition, the enzyme was found to lose activity with a half life of about 30 h under these conditions.
Although L-ribulose-5-phosphate 4-epimerase and L-fuculose-1-phosphate aldolase were found to share a common metal binding motif, they do not necessarily implement similar mechanisms. The observation of any aldolase activity with the epimerase would provide evidence in support of the idea that this enzyme operates by catalysing carbon-carbon bond breaking and making reactions. The wild type enzyme did exhibit aldolase activity, but at levels only slightly greater than the background rates. If the enediolate of dihydroxyacetone and glycolaldehyde phosphate are intermediates in this reaction, the enzyme has evolved to keep them tightly bound in its active site. Otherwise, aldolase activity might have been the predominant function of this enzyme. It is not common for enzymes readily to release their intermediates (a notable exception is the E. coli UDP-N-acetylglucosamine 2-epimerase (Morgan et al., 1997; Sala et al., 1996), and L-ribulose-5-phosphate 4-epimerase is no exception to this rule.

**Intermediate Identification Using Mutant 4-Epimerases**

The observation that the epimerase mutants still retained significant activity (Chapter III) despite having an altered metal binding site led to their being tested for aldolase activity as well. It was thought that a mutant enzyme might be able to permit protonation of the bound enediolate (or conversely deprotonation of dihydroxyacetone), which could result in the release of dihydroxyacetone and glycolaldehyde phosphate (or conversely, of their consumption in the formation of LRu5P and DXu5P). Each of the mutants was tested for
aldolase activity in the same manner as described for the wild type enzyme, always in the presence of 0.1 mM Zn$^{2+}$ to ensure that they were fully active (Chapter III).

The H97N mutant clearly did catalyse the formation of LRu5P and DXu5P from dihydroxyacetone (20 or 50 mM) and glycolaldehyde phosphate (5 mM) in the presence of 0.1 mM Zn$^{2+}$. The aldolase activity exhibited by this mutant was quite low (approximately 0.004 units/mg) but was greater than that observed with the wild type epimerase. The reaction was clearly catalytic and, in the case of the 50 mM dihydroxyacetone reaction, was followed until 30% of the glycolaldehyde phosphate was consumed (Figure 4.3A). When the reaction contained 20 mM dihydroxyacetone, 15% of the glycolaldehyde phosphate was consumed (Figure 4.3B). In both cases, formation of ketopentose phosphates consumed about 2% of the dihydroxyacetone. If the equilibrium is similar to that for aldolases catalysing reactions with similar substrates and favours sugar formation with an equilibrium constant in the range of $10^3$ to $10^4$ M$^{-1}$ (for $K_{eq} = [\text{sugar}] / [\text{dihydroxyacetone}][\text{aldehyde}]$) (Fessner & Eyrisch, 1992; Ghalambor & Heath, 1962; Ghalambor & Heath, 1966; Wong et al., 1995), then one would expect at least 4.5 mM sugar to be formed in an analogous situation. Since this enzyme forms two isomers of approximately equal energy, then entropy should favour carbon-carbon bond formation even more than usual. Equilibrium was presumably not attained by the H97N catalysed condensation reactions because the extremely long incubation times resulted in a loss of enzyme activity ($t_{1/2} = 30$ h).
Figure 4.3. Formation of An Epimeric Mixture of L-Ribulose-5-Phosphate and D-Xylulose-5-Phosphate From Dihydroxyacetone and Glycolaldehyde Phosphate (5 mM). A: 50 mM dihydroxyacetone; B: 20 mM dihydroxyacetone. Closed squares, wild type enzyme; closed circles, H97N; open squares, D76N; open circles, H95N; open triangles, supernatant from heat-inactivated H97N.
Control reactions performed with the supernatant obtained from heat denatured H97N enzyme samples showed no detectable aldol products above background, indicating that the reaction was not a result of a non-protein impurity in the enzyme preparation. As expected, a control reaction run with water in place of enzyme also had no detectable activity (data not shown); therefore, Zn$^{2+}$ itself cannot catalyse the aldol reaction to any appreciable extent. In addition neither the wild type enzyme nor the other two mutants showed comparable activity (all were isolated in an identical manner and assayed in the presence of 0.1 mM Zn$^{2+}$). Non-enzymatic condensation of glycolaldehyde phosphate and dihydroxyacetone activity has been observed to occur on strongly basic anion exchange resin (Morgenlie, 1987), so the aldol condensation could possibly occur slowly at pH 7.6 on the surface of the enzyme. The very low levels of coupling seen in solutions containing the (active) H95N and D76N 4-epimerases indicate that only insignificant amounts of LRu5P and DXu5P could be formed in this manner.

The lack of aldolase activity observed in the D76N mutant epimerase is not completely unexpected. The aspartate residue, by analogy to the L-fuculose-1-phosphate aldolase, could act as a catalytic base in the aldolase reaction (as well as a metal ion ligand), removing the proton from C-3 of dihydroxyacetone to create the reactive anionic species. Therefore, mutation of this residue to a non-reactive, neutral asparagine residue would eliminate this activity entirely. The H95N mutant also did not exhibit any appreciable aldolase activity. This mutant's low aldolase activity could simply be a reflection of its low epimerase activity.
In order to identify the products unambiguously as the epimeric pentose phosphates, 1.0 mL of a reaction mixture containing 5 mM glycolaldehyde phosphate, 50 mM dihydroxyacetone and 0.1 mM ZnCl₂ was lyophilised after 48 h of incubation at 37°C with 0.5 mg of the H97N epimerase, dissolved in D₂O and analysed by ¹H-NMR spectroscopy. The resulting spectrum showed signals identical to those of an authentic sample of the epimeric mixture that had been independently prepared using the wild type epimerase and D-xylulose-5-phosphate (Figure 4.4). A control reaction, which had no enzyme in it, had no detectable peaks due to either LRu5P or DXu5P.

The observation of aldolase activity indicates that the mutant enzyme is able to bind the cleaved products from solution. In fact, glycolaldehyde phosphate was found to act as a competitive inhibitor (Figure 4.5), with $K_I = 0.37$ mM (Figure 4.6) of the H97N enzyme in an assay for L-ribulose-5-phosphate epimerisation in the presence of 0.1 mM Zn²⁺. Dihydroxyacetone (50 mM) did not significantly inhibit the epimerisation, nor was there any synergistic inhibition with dihydroxyacetone (50 mM) and glycolaldehyde phosphate (5 mM). The other two mutant epimerases were not tested for inhibition since they did not exhibit any aldolase activity.
Figure 4.4. $^1$H-NMR (500 MHz, D$_2$O, solvent suppressed) of the H97N Catalysed Coupling Reaction Products at 72 h. A, non-enzymatic control reaction; B, H97N catalysed reaction; C, equilibrium mixture of LRu5P and DXu5P. * indicates spinning side bands from the dihydroxyacetone peak.
The observation that the H97N epimerase exhibits significant aldolase activity provides more insight into the role of the divalent metal ion in the epimerisation reaction. This activity indicates that the active site of the epimerase is capable of promoting carbon-carbon bond cleavage. The H97N mutation has probably lowered the rate of the carbon-carbon cleavage reaction. In addition, this mutation has disrupted the ability of the enzyme to keep the enediolate intermediate from being protonated. It seems odd that glycolaldehyde phosphate is a competitive inhibitor of the mutant enzyme but not of the wild type enzyme. This could be a kinetic effect if the form of the enzyme which accepts the substrates is different from the form which sequesters the intermediates, and binding of the intermediates from solution by the wild type enzyme occurs very slowly. Alternatively, the mutations may simply have altered the enzyme's active site structure and the coordination properties of the divalent metal ion, which could allow the mutant enzyme to bind the aldehyde more tightly. It should be noted that the free aldehyde exists primarily in the hydrated form (Müller et al., 1990) and this may be the species causing the inhibition. If this is the case, perhaps the wild type enzyme is unable to bind the hydrated aldehyde.
Figure 4.5. Glycolaldehyde Phosphate is a Competitive Inhibitor of Epimerisation by H97N L-Ribulose-5-Phosphate 4-Epimerase. Closed circles, 0.1 mM inhibitor; open squares, 0.2 mM inhibitor; closed squares, 0.4 mM inhibitor; open triangles, 0.5 mM inhibitor; closed triangles, 1 mM inhibitor; open inverted triangles, 2 mM inhibitor.
Figure 4.6. $K_i$ Determination for H97N L-Ribulose-5-Phosphate 4-epimerase Inhibition by Glycolaldehyde Phosphate. The data points are the $K_M$ (apparent) values from Figure 4.5.
Detection of Intermediate Release

A continuous assay to detect release of either dihydroxyacetone (using glycerol dehydrogenase) or glycolaldehyde phosphate (using α-glycerophosphate dehydrogenase) by the wild type and mutant enzymes did not demonstrate any significant activity. Presumably, release of the intermediates was sufficiently slow during the time frame of the continuous assay (4 hours). Therefore a stopped assay was employed to observe release of dihydroxyacetone over 72 hours. The wild type and H97N epimerases were assayed for release in the presence of 10 mM of the equilibrating ketopentose phosphates and 0.1 mM Zn$^{2+}$ using a stopped assay for dihydroxyacetone. At timed intervals, aliquots were removed and the amount of dihydroxyacetone present was determined. The H97N epimerase catalysed the reaction slightly more efficiently than the wild type epimerase (Figure 4.7), and the wild type epimerase catalysed the reaction at a rate greater than the background. In this case the background was determined using the supernatant from a heat-inactivated sample of the H97N epimerase. There are a few points which indicate negative amounts of dihydroxyacetone as a result of random error. The method for determining the amount of dihydroxyacetone is limited by the small amounts released. Therefore, this experiment served only to check the results observed from the coupling experiment and gave only qualitative results. The rate of dihydroxyacetone formation by this mutant was too low to measure accurately, however, it was clear that the aldolase products were formed. The release was clearly catalytic with the H97N epimerase, since approximately three equivalents
Figure 4.7. Dihydroxyacetone Release by H97N and Wild Type L-Ribulose-5-Phosphate 4-Epimerases Equilibrating 10 mM L-Ribulose-5-Phosphate and D-Xylulose-5-Phosphate. Closed circles, H97N; closed squares, wild type; open triangles, supernatant from heat-inactivated H97N.
of dihydroxyacetone were produced over 24 hours. Approximately 2.5 equivalents of
dihydroxyacetone were released over 48 hours with the wild type epimerase. These reactions
level off after about 40 hours, at dihydroxyacetone concentrations well below expected
equilibrium values, presumably as a result of the loss of enzyme activity.

**An Aldol-Like Mechanism for \( \text{L-Ribulose-5-Phosphate 4-Epimerase} \)**

The observed aldol reaction could occur during epimerisation by an infrequent
protonation of the bound enediolate by an enzymatic acid/base residue. In the reverse of
the aldol reaction, this acid/base residue would deprotonate dihydroxyacetone. An active
site residue which could act as the acid/base catalyst in the aldolase reaction is the displaced
D76 ligand, by analogy to the glutamate residue (E73) of FucA. Alternatively, the mutant
epimerase may simply be much “leakier” than the wild type enzyme and able to release the
enediolate into solution. In this case the enzyme would have to bind the enediolate or
perhaps even a zinc-enediolate complex directly from solution in the reverse direction.

The retroaldol/aldol mechanism requires that for stereochemical inversion to occur,
the enediolate must be able to add to either face of the bound aldehyde. Although the
aldolases are usually thought to exhibit high stereoselectivity, this is not always the case. For
example, the Class II tagatose-1,6-bisphosphate aldolase is able to cleave the C-4 epimer (\( \text{D-fructose-1,6-bisphosphate} \)) of its natural substrate, endowing it with what is effectively an
epimerase activity (Fessner & Eyrisch, 1992).
The aldolase activity of the H97N mutant epimerase supports a mechanism for the reaction catalysed by L-ribulose-5-phosphate 4-epimerase involving carbon-carbon bond cleavage and strengthens the notion that the enzyme is evolutionarily linked to the Class II aldolases. One noticeable difference is that FucA and the related aldolases (rhamnulose-1-phosphate, fructose-1,6-bisphosphate and tagatose-1,6-bisphosphate aldolases) are highly specific for dihydroxyacetone phosphate as a donor substrate (Gijsen et al., 1996; Wong et al., 1995). In the case of the epimerase, dihydroxyacetone must play this role. The three-dimensional structure of FucA shows that the phosphate of the intermediate analogue lies in a pocket lined with neutral residues that participate in hydrogen bonds. These residues are conserved in the epimerase sequence even though there is no phosphate group at C-1 of LRu5P and DXu5P. A possible explanation for this is that the residues have been conserved in order to bind to the displaced D76 residue and prevent it from protonating the bound enolate. Alternatively, this pocket could bind the C-5 phosphate group of LRu5P and DXu5P. Another difference is that sequence alignments place a threonine residue of the epimerase at the position of the FucA tyrosine 113. Threonine residues are not generally thought of as catalytic residues. This tyrosine is thought to act as a general base during the carbon-carbon bond cleavage step (Figure 4.1A) in the reaction catalysed by L-fuculose-1-phosphate aldolase. In the case of the epimerase, it is likely that two spatially distinct bases (in place of the aldolase's tyrosine residue) would be required in order to promote the cleavage of both epimers (the aldolase cleaves only one epimer). A number of epimerases require two bases for epimerisation (Tanner & Kenyon, 1998). The changes required to
permit this may have caused significant restructuring in this portion of the active site. A possible residue which may serve as a catalytic base in the epimerisation is the displaced D76, although the side chain of this residue may be too short for it to reach and deprotonate the C-4 hydroxyl group of either L\textsubscript{R}u\textsubscript{5}P or D\textsubscript{X}u\textsubscript{5}P. It is more likely that this residue lies near C-3 of the sugar substrate, where it could occasionally act in the same manner as does E73 in the aldolase (and give the epimerase some aldolase activity), deprotonating dihydroxyacetone or conversely, protonating its enediolate.

**Conclusion**

The *E. coli* L-ribulose-5-phosphate 4-epimerase (and especially the H97N mutant) has been shown to catalyse an aldol addition of dihydroxyacetone to glycolaldehyde phosphate. This work strongly supports a retroaldol/aldol mechanism, similar to the mechanism used by the *E. coli* Class II L-fuculose-1-phosphate aldolase, as the mechanism used by L-ribulose-5-phosphate-4-epimerase. Therefore, the similarities between these two enzymes extend beyond amino acid sequence homology and a common metal-binding motif.
Experimental Methods

Preparation of Glycolaldehyde Phosphate

Glycolaldehyde phosphate was prepared according to a modification of the method of Müller et al. (1990). In this procedure (outlined in Figure 4.8), allyl alcohol is phosphorylated and isolated as bis(cyclohexylammonium) allyl phosphate, which is then converted by ozonolysis and exchange of the counterion to calcium glycolaldehyde phosphate. Because this compound is not very soluble in aqueous solutions at neutral pH, calcium was exchanged for sodium, creating a hygroscopic solid.

Figure 4.8. Synthetic Route to Glycolaldehyde Phosphate from Allyl Alcohol.
A) *Bis(cyclohexylammonium) allyl phosphate*

To a mixture of 1.90 g (19.4 mmol) \( \text{H}_3\text{PO}_4 \) (crystalline) and 4.00 g (39.6 mmol) triethylamine in 30.00 g (517 mmol) allyl alcohol was added 14.40 g (100 mmol) trichloroacetonitrile dropwise, and stirred for 4 hours at 75°C. After distilling off the trichloroacetonitrile (75°C, ca. 15 mm Hg) the mixture was concentrated to 15 mL by rotary evaporation at room temperature, and then mixed with 200 mL water. The aqueous phase was extracted twice with 150 mL diethyl ether and then mixed with 15 mL (130 mmol) cyclohexylamine (Chx) and evaporated at room temperature. The white powder was dissolved in 50 mL water and mixed with acetone until it became turbid. After sitting overnight at 4°C, the product precipitated as a white amorphous powder. The precipitate was dried under high vacuum (0.3 mm Hg) at ambient temperature. The \(^1\)H-, \(^13\)C-, and \(^31\)P-NMR spectra agreed with those reported by Müller *et al.* (1990). \(^1\)H-NMR (200 MHz, D\(_2\)O): \( \delta \) 1.07-1.87 (cluster of peaks, 20 H, ChxNH\(_3^+\)); \( \delta \) 2.99 (m, 2 H, ChxNH\(_3^+\)); \( \delta \) 4.11 (tt, \( \text{J} = 4.7, 1.4 \text{ Hz}, 2 \text{ H}, \text{H-C1} \)); \( \delta \) 5.02 (dd, \( \text{J} = 10.4, 1.5 \text{ Hz}, 1 \text{ H}, \text{H-C3} \)); \( \delta \) 5.19 (dd, \( \text{J} = 17.2, 1.8 \text{ Hz}, 1 \text{ H}, \text{H-C3} \)); \( \delta \) 5.84 (m, 1 H, H-C2); see Appendix B. \(^13\)C-NMR (50 MHz, D\(_2\)O): \( \delta \) 24.50 (CH\(_2\), ChxNH\(_3^+\)); \( \delta \) 25.00 (CH\(_2\), ChxNH\(_3^+\)); \( \delta \) 31.05 (CH\(_2\), ChxNH\(_3^+\)); \( \delta \) 51.00 (CH, ChxNH\(_3^+\)); \( \delta \) 65.92 (CH\(_2\), C1); \( \delta \) 116.71 (CH\(_2\), C3); \( \delta \) 136.20 (CH, C2). \(^31\)P-NMR (51 MHz, D\(_2\)O): \( \delta \) 3.84 (s, 1P).
B) Sodium Glycolaldehyde Phosphate

To convert bis(cyclohexylammonium) allyl phosphate to triethylammonium allyl phosphate, 5.4 g (16 mmol) of the former was suspended in 50 mL water and shaken for 30 minutes with 7.5 g washed Dowex AG50W (H\(^+\)). The resin was filtered out and washed with 150 mL water. The combined filtrates were mixed with 250 mL diethyl ether and 20 mL (144 mmol) triethylamine. The aqueous phase was extracted and evaporated at room temperature and dried under high vacuum (0.3 mm Hg). The resulting oil was ozonolyzed in 130 mL of methanol at -78°C until the solution was an intense blue colour. After removal of excess O\(_3\) with an argon stream, 5.5 mL methyl sulphide (75 mmol; 5 equivalents) was added, and the colourless mixture was left at -20°C for 24 h. This mixture was then mixed with 100 mL ice water and 33 g washed Dowex AG50W (H\(^+\)) and stirred for 30 minutes. The ion exchanger was filtered out and washed with 400 mL ice water. The combined filtrates were concentrated at room temperature to 50 mL. The resulting mixture was then mixed with a solution of 2.5 g Ca(OAc)\(_2\) (15.0 mmol, 1 equivalent) in 25 mL water. Acetone (50 mL) was dropped slowly into this solution at 4°C. This suspension was kept at 4°C for at least 14 h, at which time the precipitate was collected by centrifugation. Additional precipitate was obtained by mixing the mother liquor with 30 mL of acetone. The combined precipitates were suspended in water and lyophilized. The NMR data were consistent with those presented in Müller et al. (1990). \(^1\)H-NMR (200 MHz, D\(_2\)O/drop HCl): δ 3.65 (dd, J= 4.90, 6.87 Hz, 2 H, H-C2), δ 4.97 (t, J= 4.83 Hz, 1 H, H-C1); see Appendix B. \(^13\)C-NMR
(50 MHz, D$_2$O/drop HCl): $\delta$ 68.84 (d, J = 6.74 Hz, CH$_2$, C2), $\delta$ 88.81 (d, J = 9.00 Hz, CH, C1). $^{31}$P-NMR (51 MHz, D$_2$O/drop HCl): $\delta$ 0.00 (s, 1P).

Calcium glycolaldehyde phosphate (1.0 g) was stirred with a weakly acidic cation exchanger (Amberlite DP-1, previously washed with distilled water) in Na$^+$ form. The resin was filtered out and washed with distilled water. This solution was lyophilized under high vacuum (0.3 mm Hg) to yield a hygroscopic pale yellow solid, and stored at -20°C over desiccant. In this form the aldehyde was stable for at least two years as judged by $^1$H-NMR.

**Continuous Assay for Dihydroxyacetone Release**

Release of dihydroxyacetone was monitored at 37°C in a coupled assay containing 50 mM glycylglycine (pH 7.6), 5 U glycerol dehydrogenase (Sigma G-6267 from *Bacillus megaterium*, dissolved in water), 0.15 mM NADH, 0.47 mM LRu5P, 0.1 mM ZnCl$_2$, and 1.25 mg of 4-epimerase (Zn$^{2+}$-reconstituted) in a total of 1 mL. The absorbance of the cuvette at 340 nm was monitored as a function of time.

**Glycolaldehyde Phosphate Release (A Continuous Assay)**

Release of glycolaldehyde phosphate was monitored at 37°C in a coupled assay containing 50 mM glycylglycine (pH 7.6), 5 U $\alpha$GDH (Sigma G-6751 from rabbit muscle, in an ammonium sulphate and EDTA suspension, diluted with water and concentrated in a Millipore Ultrafree Centrifugal Device (10K NMWL) to remove ammonium sulphate and EDTA immediately prior to use), 0.15 mM NADH, 0.47 mM LRu5P, 0.1 mM ZnCl$_2$, and
1.25 mg of 4-epimerase (Zn\(^{2+}\)-reconstituted) in a total of 1 mL. The absorbance of the cuvette at 340 nm was monitored as a function of time.

**Enzymatic Coupling of Dihydroxyacetone and Glycolaldehyde Phosphate**

Glycolaldehyde phosphate (5 mM, pH 7.6), dihydroxyacetone (either 20 or 50 mM) and 0.1 mM ZnCl\(_2\) were incubated at 37°C in a total of 1 mL with 0.5 mg of 4-epimerase (Zn\(^{2+}\)-reconstituted). Aliquots (60 μL) were taken periodically over the course of several days to assay for sugar content.

**Assay for Formation of Ketopentose Phosphates in Coupling Experiments**

Sugar content was determined by putting a 50 μL aliquot of the coupling reaction into a cuvette (preincubated and monitored at 340 nm for 5 minutes at 37°C) containing the normal assay components as described in Chapter II. The absorbance at 340 nm was monitored, and the rapid change in absorbance used to calculate the concentration of sugar in the coupling reaction.

The time (T) after addition of each aliquot was noted (see Figure 4.9). The background rate (collected during the five minute pre-incubation) was used to extrapolate the absorbance (A) at time T. The observed decrease in absorption upon addition of 50 μL of water to cuvettes containing all of the assay components (after a 5 minute incubation at 37°C) was subtracted from A to account for a small dilution of the cuvette contents: this was used as the absorbance (B) immediately after addition of the aliquot. The rate between
16 and 18 minutes was used to extrapolate the absorbance (C) at time T. The difference in absorption between B and C was proportional to the decrease in NADH concentration as a result of LRu5P and DXu5P in the aliquot. Therefore this difference could be used to calculate the concentration of these sugars in the assay and therefore also in the aliquot, using the extinction coefficient $\varepsilon = 6220 \, \text{M}^{-1} \, \text{cm}^{-1}$ for NADH at 340 nm.

Figure 4.9. Determination of Sugar Content.
Identification of Ketopentose Phosphates in Coupling Experiments

Glycolaldehyde phosphate (5 mM, pH 7.6), 50 mM dihydroxyacetone and 0.1 mM ZnCl$_2$ were incubated at 37°C in a total of 1 mL with 0.5 mg of H97N 4-epimerase (Zn$^{2+}$-reconstituted). A control containing no enzyme was used. After 48 hours of incubation, the reactions were passed through Millipore Ultrafree Centrifugal Devices (10K NMWL) and the filtrate was lyophilized. The residues were dissolved in 500 µL D$_2$O and lyophilized twice, and finally redissolved in 500 µL of D$_2$O. These samples were analysed by 500 MHz $^1$H-NMR. The sample which had been catalysed with the H97N mutant epimerase was spiked with an authentic sample of the equilibrium mixture (prepared as described in Chapter II) to see if the peaks due to the sugars formed in the aldol addition would increase.

Stopped Assays for Dihydroxyacetone Release

LRu$_5$P (10 mM, pH 7.6) and 0.1 mM ZnCl$_2$ were incubated at 37°C in a total of 1.0 mL with 0.5 mg of wild type or H97N 4-epimerase (Zn$^{2+}$ reconstituted). Aliquots (50 µL) were taken periodically over the course of several days to assay for dihydroxyacetone release. A negative control consisting of the supernatant of the heat-inactivated H97N 4-epimerase was also run.

The amount of dihydroxyacetone was determined spectrophotometrically. The 50 µL aliquots were incubated at 37°C in cuvettes containing 25 mM glycylglycine (pH 7.6), and 0.15 mM NADH. Addition of glycerol dehydrogenase (1 U; Sigma G-6267 from Bacillus megaterium, dissolved in water for use) to the cuvette initiated the reaction. The absorbance
at 340 nm was monitored, and the change in absorbance was used to calculate the concentration of dihydroxyacetone in the equilibration of L-ribulose-5-phosphate and D-xylulose-5-phosphate in a similar manner to that used in determination of the ketopentose phosphate content of the coupling reactions. The time (T) after addition of each aliquot was noted. The background rate (collected during the five minute preincubation) was used to extrapolate the absorbance (A) at time T. The observed decrease in absorption upon addition of 10 μL of water to cuvettes containing all the assay components (after a 5 minute incubation at 37°C) was subtracted from A to account for a small dilution of the cuvette contents: this was used as the absorbance (B) immediately after addition of the aliquot. The rate between 18 and 20 minutes was used to extrapolate the absorbance (C) at time T. The difference between B and C was proportional to the decrease in NADH concentration as a result of oxidation of the dihydroxyacetone in the aliquot. This difference was therefore used to calculate the concentration of dihydroxyacetone in the assay cuvette and also in the aliquot using the extinction coefficient for NADH ($\varepsilon = 6220 \text{ M}^{-1} \text{ cm}^{-1}$) at 340 nm.

**Assay for Enzyme Activity**

The activities of the enzymes in the intermediate release incubation mixtures were checked periodically and were determined by monitoring the decrease in NADH concentration (at $A_{340\text{ nm}}$) in a coupled reaction in a Cary 3E UV-VIS spectrophotometer at 37°C. In this assay, each cuvette contained, in a total volume of 0.5 mL, the components of the normal assay reaction with 0.1 mM ZnCl$_2$, as described in Chapter III, and 0.24 mM
L-ribulose-5-phosphate. The reaction was initiated by addition of L-ribulose-5-phosphate 4-epimerase five minutes after beginning observation of the absorbance. The amounts of enzyme used in this assay were 0.025 µg wild type epimerase and 0.05 µg H97N epimerase.

**Assay for Inhibition and Determination of $K_i$**

Zinc-reconstituted wild type and H97N enzyme activity was monitored in the usual manner in the presence of 0.1 mM Zn$^{2+}$ at 0.0944 mM LRu5P both with and without 5 mM glycolaldehyde phosphate, and both with and without 50 mM dihydroxyacetone.

In determining $K_i$ for glycolaldehyde phosphate inhibition of the H97N enzyme, the substrate concentration varied from 0.038 to 0.94 mM; while the glycolaldehyde phosphate (inhibitor) concentration varied from 0 to 2.0 mM. All other assay conditions were the same as those described for determining the enzyme's activity. Data were analysed using a Lineweaver-Burk plot, and the apparent $K_m$ values replotted as a function of inhibitor concentration. The resulting x-intercept provided the $K_i$ value. The slopes of the lines in the Lineweaver-Burk plot were also plotted as a function of inhibitor concentration, the resulting x-intercept again providing the $K_i$ value. Both $K_i$ values were identical.
Chapter V

Proposed Exploration of the Active Site of
L-Ribulose-5-Phosphate 4-Epimerase

Rationale

In Chapters III and IV, the Escherichia coli L-ribulose-5-phosphate 4-epimerase was shown to have both a metal-binding motif and a mechanistic strategy in common with the E. coli L-fuculose-1-phosphate aldolase, with which it shares significant amino acid sequence homology (Dreyer & Schulz, 1993; Dreyer & Schulz, 1996b). The similarities between these two enzymes are intriguing because this is the only epimerase currently known to use an aldolase-like mechanism. It is therefore of interest to explore the nature of the active site of the 4-epimerase, and to generate an understanding of how the active site of the 4-epimerase is different from the active site of the aldolase.
Two approaches to this understanding are through use of 1) substrate and intermediate analogues, which may act as alternative substrates, as inhibitors or as active site labelling agents, and 2) site-directed mutagenesis studies aimed at altering the activity and specificity of the epimerase.

**Small Molecule Analogues**

Stable molecules which mimic the transition state(s) and/or intermediates of a reaction pathway, and for which the enzyme has great affinity, are often bound tightly by the enzyme.

Hydroxamic acid derivatives are potent inhibitors of a number of enzymes using a catalytic Zn$^{2+}$ ion, including carbonic anhydrase II (Scolnick et al., 1997), thermolysin (Holmes & Matthews, 1981; Izquierdo-Martin & Stein, 1992), and matrilysin (Browner et al., 1995). Hydroxamates are stable analogues of enediolate intermediates of a number of enzymes (Collins, 1974), and are known to chelate metal ions (Bauer & Exner, 1974). A potent inhibitor of the Class II dihydroxyacetone-dependent aldolases is phosphoglycolohydroxamate (Fessner et al., 1996), which mimics the enediolate intermediates of dihydroxyacetone phosphate and bidentately chelates the Zn$^{2+}$ ion in their active sites. A similar molecule, glycolohydroxamate, could inhibit the 4-epimerase and provide further evidence that this epimerase follows a retroaldol/aldol reaction mechanism.
A series of hydroxamate derivatives (Figure 5.1) could be designed to determine the requirements on substrates to fit into the epimerase’s active site. Phosphoglycolohydroxamate itself may inhibit the epimerase.

Glycolohydroxamate and acetohydroxamate might exhibit synergistic inhibition in combination with glycolaldehyde phosphate; this aldehyde has been observed to inhibit the H97N epimerase competitively. Glycolaldehyde phosphate was found to inhibit the H97N mutant epimerase but not the wild type epimerase (Chapter IV).

A consideration in testing the hydroxamate derivatives as inhibitors of L-ribulose-5-phosphate 4-epimerase is that phosphoglycolohydroxamate is an inhibitor of triose
phosphate isomerase (TIM) (Collins, 1974), an enzyme used in assays of epimerase activity, coupling the epimerization reaction to NADH oxidation.

Another potential inhibitor of the epimerase is L-ribitol-5-phosphate. This compound can be prepared as a mixture of L-ribitol-5-phosphate and L-arabitol-5-phosphate by NaBH₄ reduction of L-ribulose-5-phosphate in a buffered solution (Figure 5.2) (London & Hausman, 1982). Alternatively, L-ribitol-5-phosphate may be prepared by

Glycolic acid or glycolate may inhibit the epimerase. It might also be able to add to glycidosid phosphate to form an ester (Figure 5.3). Synergistic inhibition is possible and should be investigated.

**Figure 5.3.** Addition of Glycolate to Glycidol Phosphate.

**Inactivating Agents**

Irreversible inhibitors are active site labelling agents which react with catalytic residues. Therefore they are useful in combination with proteolytic degradation and peptide purification in identifying which amino acid residue catalyses the reaction. The epimerase
may be inhibited or inactivated by glycidol phosphate. This molecule is an active site labelling agent for TIM (Rose & O’Connell, 1969; Schray et al., 1973). Haloacetol phosphates (Figure 5.4) are electrophiles which are known to inhibit TIM irreversibly (Coulson et al., 1970; de la Mare et al., 1971; Hartman, 1971; Norton & Hartman, 1972) in addition to inactivating yeast aldolase (a Class II aldolase) (Hartman, 1970; Lin et al., 1971).

Another potential inactivator of L-ribulose-5-phosphate 4-epimerase is phosphoglycolic acid chloride, which could be synergistic with dihydroxyacetone or glycolate. The anhydride shown in (Figure 5.4) is another potential active site labelling agent.

![Figure 5.4. Potential Active Site Labelling Agents.](image)
Epimerase Structure and Site-Directed Mutagenesis Studies

In order to make rational decisions about which residues to alter and with what to replace them, an X-ray crystal structure of the epimerase is required. The epimerase has been crystallized and found to diffract X-rays (Andersson et al., 1995), the preliminary steps toward obtaining a crystal structure. In addition, a collaboration with Dr. Natalie Strynadka (Department of Biochemistry, University of British Columbia) has been established to determine the crystal structure of the wild type L-ribulose-5-phosphate 4-epimerase and eventually of the mutant 4-epimerases. In addition, the free and inhibitor-complexed structures could provide information about the residues involved in epimerisation and their roles.

The D76N 4-epimerase was observed to retain the ability to bind the metal ion (to a similar degree as the wild type enzyme; see Chapter III) and also had some epimerase activity. It would be interesting to find out what the fourth ligand to the metal ion is in this mutant (instead of aspartate). If there is sufficient room in the active site, the activity of this mutant could be “rescued” by inclusion of formate in the assay buffer. The corresponding residue in L-fuculose-1-phosphate aldolase is glutamate. A D76E mutation may make the epimerase more like the aldolase, especially in combination with a H97N mutation.

The ultimate goal of these studies is to determine which residues cause the active sites of L-ribulose-5-phosphate 4-epimerase and L-fuculose-1-phosphate aldolase to catalyse different reactions. Eventually this could lead to the ability to alter enough residues in the
epimerase's active site that aldolase activity can be optimized. A comparison of the crystal structures of the aldolase (this is known: Dreyer & Schulz, 1993; Dreyer & Schulz, 1996a; Dreyer & Schulz, 1996b) and of the epimerase should unambiguously reveal which residues define the active site specificity and chemistry.

**Related Enzymes**

Recent bacterial genome sequencing projects have uncovered a number of sequences which have amino acid sequence homology to both the *E. coli* L-ribulose-5-phosphate 4-epimerase and the *E. coli* L-fuculose-1-phosphate aldolase (see Appendix C). These enzymes appear to have the same divalent metal-binding motif as the epimerase and aldolase, and likely use a similar mechanistic strategy. Neither their substates nor their metabolic roles are known. Studies on these enzymes to identify their metal binding ligands and to discover their mechanisms of catalysis could strengthen and provide further insight into the relationship between L-ribulose-5-phosphate 4-epimerase and L-fuculose-1-phosphate aldolase.
Conclusion

The catalytic bases used by L-ribulose-5-phosphate 4-epimerase to catalyse the interconversion of L-ribulose-5-phosphate and D-xylulose-5-phosphate are not known. Possible future studies could be performed and are outlined in this chapter. A combination of the active-site labelling agents and site-directed mutagenesis studies presented here will be able to identify these bases. X-ray crystallography on the epimerase should verify the identity of these residues as well as show the similarities and differences between the active sites of L-fuculose-1-phosphate aldolase and L-ribulose-5-phosphate 4-epimerase.
Appendix A

DNA Sequences

AraD of Escherichia coli K12

| M | L | E | D | L | K | R | Q | V | L | E | A | N | L | A | L | E | P | K | H | N |
| 20 |
| T | A | G | G | T | A | T | C | C | A | A | G | C | G | A | C | G | A | T | G | T | T | G | T | G |
| T | A | G | T | A | T | C | C | A | A | G | C | G | A | C | G | A | T | G | T | T | G | T | G |
| 60 |
| T | A | C | C | G | A | G | T | C | A | G | C | C | G | A | C | G | A | T | G | T | T | G | T | G |
| 120 |
| L | V | T | L | T | W | G | N | V | S | A | V | D | R | E | R | G | V | F | V | 40 |
| G | C | A | G | C | A | G | T | C | A | G | C | C | G | A | C | G | A | T | G | T | T | G | T | G |
| 180 |
| A | T | G | A | A | A | C | T | G | C | T | C | G | C | G | A | G | A | T | G | T | G | T | G |
| 240 |
| T | A | T | C | C | G | A | G | T | C | A | G | C | C | G | A | C | G | A | T | G | T | T | G | T | 80 |
| 300 |
| R | L | L | Y | Y | F | T | I | P | C | T | R | K | M | T | D | A | E | I | N | G | E |
| 360 |
| G | C | C | A | T | C | C | G | A | G | T | C | A | G | C | C | G | A | T | G | T | T | G | T | G |
| 420 |
| Y | F | Y | F | T | I | P | C | T | R | K | M | T | D | A | E | I | N | G | E |
| 480 |
| A | T | G | A | A | A | C | T | G | C | T | C | G | C | G | A | G | A | T | G | T | T | G | T | G |
| 540 |
| A | T | G | A | A | A | C | T | G | C | T | C | G | C | G | A | G | A | T | G | T | T | G | T | G |
| 600 |
| Y | E | W | E | T | G | N | V | I | V | E | T | T | F | E | K | Q | G | I | D | A |
| 660 |
| G | A | G | A | A | C | T | G | C | T | C | G | C | G | A | G | A | T | G | T | T | G | T | G |
| 720 |
| E | D | A | V | H | N | A | I | V | L | E | E | E | V | A | Y | M | G | I | F | C |
| 780 |
| G | A | G | A | A | C | T | G | C | T | C | G | C | G | A | G | A | T | G | T | T | G | T | G |
| 840 |
| R | Q | L | A | P | L | F | D | N | Q | Q | T | L | L | D | K | H | Y | L |
| 900 |
| G | C | A | G | A | A | C | T | G | C | A | T | G | C | G | A | G | A | T | G | T | T | G | T | G |
| 960 |
| R | K | H | G | A | K | A | Y | Y | G | Q | E | N | D |
| 1020 |
Appendix A: DNA Sequences

AraB of Escherichia coli

ATGGCGATTG CAATGGCTCTG CATTTGCTGG TGCTTCTGAG ACGGCTCAGT 60
TACCCGCTAC GTAGCATCAG TGGATCAGCT GAGCCAAGTG 120
AATCGCGGCTG GCTGGGCTTC GAAATCAGCT GCTGAGCTTG 180
ATGAAAGCCT GCCCGAGCTT CTTTCGAGCT GAGGATTCTG 240
TCGCGACGCT GCTGGGCTTC GAAATCAGCT GCTGAGCTTG 300
GACGACGGCT GCTGGCTGGC GAAATCAGCT GCTGAGCTTG 360
CTGTGCTGGT CCAGACGGCT TGGGCTGCAG GCAGCAAGCT 420
GCTGGGCTTC GAAATCAGCT GCTGAGCTTG 480
CACTGCTGCT TGGGCTGCAG GCAGCAAGCT 540
GCTGGGCTTC GAAATCAGCT GCTGAGCTTG 600
ATGAAAGCCT GCCCGAGCTT CTTTCGAGCT GAGGATTCTG 660
TCGCGACGCT GCTGGGCTTC GAAATCAGCT GCTGAGCTTG 720
GACGACGGCT GCTGGCTGGT CCAGACGGCT TGGGCTGCAG GCAGCAAGCT 780
GCTGGGCTTC GAAATCAGCT GCTGAGCTTG 840
ATGAAAGCCT GCCCGAGCTT CTTTCGAGCT GAGGATTCTG 900
TCGCGACGCT GCTGGGCTTC GAAATCAGCT GCTGAGCTTG 960
GACGACGGCT GCTGGCTGGT CCAGACGGCT TGGGCTGCAG GCAGCAAGCT 1020
GCTGGGCTTC GAAATCAGCT GCTGAGCTTG 1080

Nde I

GGCCATTTCTG AATGGGCTGC GCGGGTGTTG TGCGAGCTGT GGCAGGGCTG 840
CCCGAGGCTG TTTTGCACCA CCAGAGGCCG GCGGGTGTTG TGCGAGCTGT GGCAGGGCTG 900
GGGCTCGGCTG CGGCGTTCTG GCGGCTCTGTT TGCGAGCTGT GGCAGGGCTG 960
CTGGATGCTGC ACAAAAGCAG CGTGGGCGAG GCAGCGTCTC TGGGCTGCAG GCAGCAAGCT 1020
GATGGGCGAG CTGGTGGCTG TAGGATGCTG TGGGCTGCAG GCAGCAAGCT 1080
Appendix A: DNA Sequences

<table>
<thead>
<tr>
<th>DNA Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCGGACTGA AAGCGCAAT CAACGCCAGC CAGAAACAAAC TGCTTCCGGC GCTGACCGAA</td>
<td>1140</td>
</tr>
<tr>
<td>GGCCTTGACT TACGCGGTTTA GTTCCGCTGCG GTCTTTGTG GACGAAGGCCG CGACTGGGCTT</td>
<td>1200</td>
</tr>
<tr>
<td>GCAAGGCGCA AAAATCGTAC TCTGCTGACTAC TGCTCGAGTG TTCTACGAGTGCTT</td>
<td>1260</td>
</tr>
<tr>
<td>CGTACCCGGT TTCTGGGATT GGTGCGGGACT TTTCCCACACT ATTGGCTAGA ATGGGAGCGA</td>
<td>1320</td>
</tr>
<tr>
<td>ACCGACGCTC CGCTGCGTT TGGGCTTGGG ATGCTGCGCA CGCAGCGGCA CGCACCGGCA</td>
<td>1380</td>
</tr>
<tr>
<td>ATCATGAGAT GTTTCACCGA TCAGGGGACT GCCGTCAATA AGCTGATGCA GCTGCGGCA</td>
<td>1440</td>
</tr>
<tr>
<td>ATCCGCGCGA AAAACCAGTG CATTATCGACT GGTCTGCTGCG ATGCTGCGGCA CGACGCGAC</td>
<td>1500</td>
</tr>
<tr>
<td>CAAATGGTGTG CTCCTAGGCA GTCTGTACGC TCCGTGCGGCG CAGTGTCTGCGCGGACGAC</td>
<td>1560</td>
</tr>
<tr>
<td>CGCAAAAGGTGC AGGCAGACAT CCCATTGGCC CAGCAAAAAA TGGGCAGGCG ATGAGGAC</td>
<td>1620</td>
</tr>
<tr>
<td>ACCGTCAAGCA CACGCGCAGC AGCTTGCAGAC GTCTGCTTGAG TGGCTGGTCGGCG</td>
<td>1680</td>
</tr>
<tr>
<td>CACTGCTGTTT GGTTGCTGTTT TGGTGGCTTGG ATGCTGCGTACG</td>
<td>1701</td>
</tr>
</tbody>
</table>
## Appendix A: DNA Sequences

### Sequence of pRE

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCGCGCGTTT CGGTGATGAC GGTGAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA</td>
<td>60</td>
</tr>
<tr>
<td>AGGCGCGAAA GCCACTACTG CCACTTTTGG ATGACTGTA CGTCCAGGTC CTCTGCCCCG</td>
<td>120</td>
</tr>
<tr>
<td>CAGCTTTGGCT GATAAGGGAG GACGACGGCC TCAAGGGCGGC TCAGCGGCTAT AACCACCACA</td>
<td>180</td>
</tr>
<tr>
<td>TTTGCAGGCGG TGCTGGCTGG CTTAATAACTG CGGCATCAGA GCAGATTGTA CTGCGACTGC</td>
<td>240</td>
</tr>
<tr>
<td>ACCATGAGCC GTGTGAAATA CGCAGAGGAG ATAAAAAGGAT AAAAAAGGAT GACGAGGCA</td>
<td>300</td>
</tr>
<tr>
<td>TTAATTGTTT TGGTGGACTG TTTTGAGTCA ATTTTCTTGG TGTTTGCGAT</td>
<td>360</td>
</tr>
<tr>
<td>GGGTTTGGAG TTTCTCGCTG TTTGACGACT GTAGTCTGCC ATGAAAGCTG ATGAGCGC</td>
<td>420</td>
</tr>
<tr>
<td>GCCACTCAG AACCCGTCTG TCCGCTCTTA GCAGTTACTG GCCGATGAA</td>
<td>480</td>
</tr>
<tr>
<td>TCAAGCTTAAT TGGGCTGGGA GACATTATAC AGAACCTTAA AGGGACCGCC</td>
<td>540</td>
</tr>
<tr>
<td>CGATTGAGGG CTTGAGCCGG AAAACGCCGG AAGTGGCGGA GAAAGGAGGG</td>
<td>600</td>
</tr>
<tr>
<td>GTAAATACCT GAAGCTCCGG TTCCGCCCGT TATCCGTTATG ATGACTTACG ATCCCTCCG</td>
<td>660</td>
</tr>
<tr>
<td>AAAAGCGCGGC GCCTAGGGGC GCTGAGATCG GGACTGACCC ATGAGATGCT CGTTTGGTGG</td>
<td>720</td>
</tr>
<tr>
<td>CCGCCGCGCC TTAACTGAGG GTTACGAGGC GGTCGCGGCT ATCGGCCTTT TCGCTCGGGG</td>
<td>780</td>
</tr>
<tr>
<td>AACTTGTTGGG AAGGGCGGATC GGTCGGCGGC TCTTCGCTAT AGCGCAGCGT TACGTTACTG</td>
<td>840</td>
</tr>
<tr>
<td>TGGCACAACC TTCCCGCTAG CCACGGCGGC AAGAGCGATA ATGCGGGCTCG</td>
<td>900</td>
</tr>
<tr>
<td>CCGGATTATA CGTGAAGCCG CTAGCACTCG CAGACTACGT TTATCGCAAG GGTCGGTCTG</td>
<td>960</td>
</tr>
<tr>
<td>GTAATCGCGT GCAGGCGGCC CTTGACGCTC GTGTAGTGTC ATCTTACTGA ATGAGCCGAC</td>
<td>1020</td>
</tr>
<tr>
<td>CAGGTGACTG GCGTATGGCA CGCAGTCTCC GCCAATTGTT AATTCGCTTT GCAACCGGAC</td>
<td>1080</td>
</tr>
</tbody>
</table>

*Pst I*
### Appendix A: DNA Sequences

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nde I</strong></td>
<td></td>
</tr>
<tr>
<td>TAATTACCTGG CGTTTGAAGAT CTCTTAAATAT ATGCTGTTTCT CTGTTTAGAA CTGTTATCCG</td>
<td>1500</td>
</tr>
<tr>
<td><strong>Pribnow (trc promoter) -35</strong></td>
<td></td>
</tr>
<tr>
<td>CTACCACATTC CACCACATTAT ACAGACTCTGA TGGATTTAGT TCAACACTCT ACATATTTCG</td>
<td>1560</td>
</tr>
<tr>
<td><strong>EcoRI</strong></td>
<td></td>
</tr>
<tr>
<td>GAAATTACGG CGGCGCGCGG GCTGTTGCGG GCTGCGCCCG GCTGCGCCCG</td>
<td>1800</td>
</tr>
<tr>
<td>TTAAAGCCTGA ACCGGGATCG ATCCGGGCGG TGGATTTAGT TCAACACTCT</td>
<td>1860</td>
</tr>
<tr>
<td>ACCGGGATCG ATCCGGGCGG TGGATTTAGT TCAACACTCT</td>
<td>1920</td>
</tr>
<tr>
<td>TGGATTTAGT TCAACACTCT</td>
<td>1980</td>
</tr>
<tr>
<td>ACCGGGATCG ATCCGGGCGG TGGATTTAGT TCAACACTCT</td>
<td>2040</td>
</tr>
<tr>
<td>GAAATTACGG CGGCGCGCGG GCTGTTGCGG GCTGCGCCCG GCTGCGCCCG</td>
<td>2100</td>
</tr>
<tr>
<td><strong>Nde I</strong></td>
<td></td>
</tr>
<tr>
<td>TAATTACCTGG CGTTTGAAGAT CTCTTAAATAT ATGCTGTTTCT CTGTTTAGAA CTGTTATCCG</td>
<td>1500</td>
</tr>
<tr>
<td><strong>Pribnow (trc promoter) -35</strong></td>
<td></td>
</tr>
<tr>
<td>CTACCACATTC CACCACATTAT ACAGACTCTGA TGGATTTAGT TCAACACTCT ACATATTTCG</td>
<td>1560</td>
</tr>
<tr>
<td><strong>EcoRI</strong></td>
<td></td>
</tr>
<tr>
<td>GAAATTACGG CGGCGCGCGG GCTGTTGCGG GCTGCGCCCG GCTGCGCCCG</td>
<td>1800</td>
</tr>
<tr>
<td>TTAAAGCCTGA ACCGGGATCG ATCCGGGCGG TGGATTTAGT TCAACACTCT</td>
<td>1860</td>
</tr>
<tr>
<td>ACCGGGATCG ATCCGGGCGG TGGATTTAGT TCAACACTCT</td>
<td>1920</td>
</tr>
<tr>
<td>TGGATTTAGT TCAACACTCT</td>
<td>1980</td>
</tr>
<tr>
<td>ACCGGGATCG ATCCGGGCGG TGGATTTAGT TCAACACTCT</td>
<td>2040</td>
</tr>
<tr>
<td>GAAATTACGG CGGCGCGCGG GCTGTTGCGG GCTGCGCCCG GCTGCGCCCG</td>
<td>2100</td>
</tr>
</tbody>
</table>
### Appendix A: DNA Sequences

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGTTCCGTTTG</td>
<td>2220</td>
</tr>
<tr>
<td>TCCAGCAAGG</td>
<td>2280</td>
</tr>
<tr>
<td>CTTATACGCG</td>
<td>2340</td>
</tr>
<tr>
<td>CATCCGACAT</td>
<td>2400</td>
</tr>
<tr>
<td>CTGTTAGGAC</td>
<td>2460</td>
</tr>
<tr>
<td>GACCATCGCC</td>
<td>2520</td>
</tr>
<tr>
<td>AAGAAGATCC</td>
<td>2580</td>
</tr>
<tr>
<td>TTCTCTAGG</td>
<td>2640</td>
</tr>
<tr>
<td>TTCTGAGTGC</td>
<td>2700</td>
</tr>
<tr>
<td>GCTTAAATCG</td>
<td>2760</td>
</tr>
<tr>
<td>GACTCCCCGT</td>
<td>2820</td>
</tr>
<tr>
<td>CTGAGCGCGA</td>
<td>2880</td>
</tr>
<tr>
<td>TTACTGCGTG</td>
<td>2940</td>
</tr>
<tr>
<td>GTTTGGCTTC</td>
<td>3000</td>
</tr>
<tr>
<td>CCAAAACCGC</td>
<td>3060</td>
</tr>
<tr>
<td>AGATGATCCG</td>
<td>3120</td>
</tr>
<tr>
<td>ACTCACGCTG</td>
<td>3180</td>
</tr>
<tr>
<td>GGAGCAGGCT</td>
<td>3240</td>
</tr>
<tr>
<td>GTGACCGATG</td>
<td>3300</td>
</tr>
</tbody>
</table>
Appendix A: DNA Sequences

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGGCCGTCAAT ACAGGATAAT ACCGCGCCAC ATAGCAGAAG TTTAAAAGTG TCTATCATTTG</td>
<td>3360</td>
</tr>
<tr>
<td>GCCGAGATT ACGGCTATT TGGCGGCGTG TATCGTCTTG AAATTTTCAC GAGTCAAGCT</td>
<td>3420</td>
</tr>
<tr>
<td>GAAAACGTTC TCCGGGCCGGA AAGACTCTCAA GGAATCTTACC GCTTTGGAAGA TCCAGGTCGA</td>
<td>3480</td>
</tr>
<tr>
<td>TGGAAACCAC TCGTGACCCC AACTGATCTTT CACTGTACCT TACTTTTACC AGCGTTTCTG</td>
<td>3540</td>
</tr>
<tr>
<td>ACATTGGGTG AGCCAGCTGGG TTGACTAGAA GTCTGATGAAA ATGAAAAGTG GTGCAAAGAC</td>
<td>3600</td>
</tr>
<tr>
<td>GCCGCAAGTA AACAGGAAAGG CAAATGCGG CAAAAGAGG AAAAGGCGG ACACGGAATA</td>
<td>3660</td>
</tr>
<tr>
<td>CCACTCGTTT TTGACCTCCC GTTTTTACGC GTTTTTTCCC TTATCCCCG TGTGCTTTTA</td>
<td>3720</td>
</tr>
<tr>
<td>GGTTGAATACT CATACCTCTTC CTTTTTCAAT ATTATTTGAG CATTTTACAG GTTTATTTGC</td>
<td>3753</td>
</tr>
<tr>
<td>CAACTTATGA GTATGAGAAG GAAAAGTTA TAATAACTTC GTAATAGTC CCAATAACAG</td>
<td>3816</td>
</tr>
<tr>
<td>TCATGACGCG ATACATATTT GAATGTATTT AGAAAAATAAA ACAATAGGG GTTCCGCGCA</td>
<td>3880</td>
</tr>
<tr>
<td>AGTTACTCCG TATGCTACCA ATTATCGAAA TCTTTTTATT TGTATCCCC CGAGCGCGGT</td>
<td>3944</td>
</tr>
<tr>
<td>CATTTCCGGG AAGAGTGCCA CCGAGCCTTC AAGAAAACG CATTTATCATG ACATTAACCT</td>
<td>4008</td>
</tr>
<tr>
<td>GTAAAGGGGC TTTTCACGGT GGACTGCAGA TTTTGGTAATA ATATAGTAC TGTAATTGG</td>
<td>4072</td>
</tr>
<tr>
<td>ATGACTCGCC</td>
<td>3788</td>
</tr>
<tr>
<td>Aat II</td>
<td>3788</td>
</tr>
<tr>
<td>Bcl I</td>
<td>1375</td>
</tr>
<tr>
<td>Bgl II</td>
<td>1475</td>
</tr>
<tr>
<td>Cla I</td>
<td>995</td>
</tr>
<tr>
<td>EcoR I</td>
<td>1741</td>
</tr>
<tr>
<td>Nde I</td>
<td>1469</td>
</tr>
<tr>
<td>Nru I</td>
<td>1369</td>
</tr>
<tr>
<td>Pst I</td>
<td>773</td>
</tr>
<tr>
<td>Pvu II</td>
<td>768</td>
</tr>
<tr>
<td>Sca I</td>
<td>3236</td>
</tr>
</tbody>
</table>

This sequence was assembled from information about pBS provided by Stratagene (and annotated by members of the Knowles lab), araD sequence (nt 785 to 1470), and pKK233-2 sequence (nt 1471-1740) (from Genebank).

Some Unique Restriction Sites on pRE1

<table>
<thead>
<tr>
<th>Restriction Site</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aat II</td>
<td>3788</td>
</tr>
<tr>
<td>Bcl I</td>
<td>1375</td>
</tr>
<tr>
<td>Bgl II</td>
<td>1475</td>
</tr>
<tr>
<td>Cla I</td>
<td>995</td>
</tr>
<tr>
<td>EcoR I</td>
<td>1741</td>
</tr>
<tr>
<td>Nde I</td>
<td>1469</td>
</tr>
<tr>
<td>Nru I</td>
<td>1369</td>
</tr>
<tr>
<td>Pst I</td>
<td>773</td>
</tr>
<tr>
<td>Pvu II</td>
<td>768</td>
</tr>
<tr>
<td>Sca I</td>
<td>3236</td>
</tr>
</tbody>
</table>
## Oligonucleotides Used in This Study

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5' to 3')</th>
<th>A</th>
<th>C</th>
<th>G</th>
<th>T</th>
<th>total</th>
<th>mol. wt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AJ1</td>
<td>GCGAGTGCCGTGTTAAACAATGCCGCAATGGA</td>
<td>8</td>
<td>7</td>
<td>10</td>
<td>6</td>
<td>31</td>
<td>9634</td>
</tr>
<tr>
<td>AJ2</td>
<td>CGTGGCAGGAATTCGATAATGCCAAATGCCG</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>29</td>
<td>8970</td>
</tr>
<tr>
<td>AJ3</td>
<td>TGTGCTACAGAATTTCGCCACGACGACCA</td>
<td>7</td>
<td>10</td>
<td>7</td>
<td>5</td>
<td>29</td>
<td>8899</td>
</tr>
<tr>
<td>AJ4</td>
<td>AAAAGCCCTCTTCGAAACACGGCAACTCACCCG</td>
<td>10</td>
<td>13</td>
<td>5</td>
<td>4</td>
<td>32</td>
<td>9740</td>
</tr>
<tr>
<td>AJ5</td>
<td>GTTGGCGTGTGCGAGAGGGCCTTTTCGTAC</td>
<td>4</td>
<td>6</td>
<td>11</td>
<td>11</td>
<td>32</td>
<td>9941</td>
</tr>
<tr>
<td>AJ6</td>
<td>TGGTCGATTCGAAAAGGCCGGCTTTGCTAT</td>
<td>4</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>27</td>
<td>8295</td>
</tr>
<tr>
<td>AJ7</td>
<td>GGGGCGCGGTTCCGAAATGGGACGACACG</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>3</td>
<td>27</td>
<td>8396</td>
</tr>
<tr>
<td>AJ8</td>
<td>GGGGTTCCATTTGCGGATTGCAATGGCCTCGT</td>
<td>8</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>35</td>
<td>10847</td>
</tr>
<tr>
<td>AJ9</td>
<td>CGCCGGACATCTGGAGTTATAGAGTCGACG</td>
<td>7</td>
<td>11</td>
<td>14</td>
<td>7</td>
<td>39</td>
<td>12091</td>
</tr>
<tr>
<td>AJ10</td>
<td>CGAACACGCAACTCACC</td>
<td>6</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>18</td>
<td>5445</td>
</tr>
<tr>
<td>MT04</td>
<td>CCGCAGGGGGATCTTGGCGAGTTAGCTGCGCCG</td>
<td>6</td>
<td>13</td>
<td>12</td>
<td>8</td>
<td>39</td>
<td>12002</td>
</tr>
<tr>
<td>MT05</td>
<td>GGGCCCATGGCCCATAATGTTAGAGATCTCAAACGGGCA</td>
<td>11</td>
<td>12</td>
<td>10</td>
<td>7</td>
<td>40</td>
<td>12315</td>
</tr>
<tr>
<td>MT08</td>
<td>CGACGAGCGCTGACACCCAGCGATGCC</td>
<td>6</td>
<td>9</td>
<td>7</td>
<td>2</td>
<td>24</td>
<td>7389</td>
</tr>
<tr>
<td>MT09</td>
<td>GCAGAGCCGAGGTATGATGGCGGCTG</td>
<td>5</td>
<td>3</td>
<td>12</td>
<td>4</td>
<td>24</td>
<td>7595</td>
</tr>
<tr>
<td>MT12</td>
<td>GGCCTGATTAACACGCGTCATGA</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>22</td>
<td>6833</td>
</tr>
<tr>
<td>MT13</td>
<td>GGGCCAGGGCGGTCAGTGTG</td>
<td>3</td>
<td>5</td>
<td>11</td>
<td>3</td>
<td>22</td>
<td>6915</td>
</tr>
<tr>
<td>MT14</td>
<td>GGGCCTCTGGCTCCATTCCACGGGCGCCT</td>
<td>2</td>
<td>11</td>
<td>8</td>
<td>7</td>
<td>28</td>
<td>8561</td>
</tr>
<tr>
<td>MT15</td>
<td>CGTACCTTCAACCACTTCACGGGT</td>
<td>5</td>
<td>10</td>
<td>3</td>
<td>6</td>
<td>24</td>
<td>7265</td>
</tr>
</tbody>
</table>
Figure B.1. 500 MHz Proton-NMR Spectrum of Commercial D-Xylulose-5-Phosphate (Sodium Salt) in D$_2$O (Solvent Suppressed).
Figure B.2. 500 MHz DQFCOSY of L-Ribulose-5-Phosphate (Free Acid) in D₂O (Solvent Suppressed).
Figure B.3. 500 MHz DQFCOSY of Commercial D-Xylulose-5-Phosphate (Sodium Salt) in D$_2$O (Solvent Suppressed).
Figure B.4. 200 MHz Proton-NMR of Allyl Phosphate (Cyclohexylammonium Salt) in D$_2$O.
Figure B.5. 200 MHz Proton-NMR Spectrum of Glycolaldehyde Phosphate (Ca$^{2+}$ Salt) in Acidic D$_2$O.
Appendix C
Protein Sequence Alignments

Complete Amino Acid Sequence Alignment

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Consensus</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>arad_ecoli</td>
<td>ML EDLKRQVLEANLAPKHN L V</td>
<td>M L L L--</td>
<td>22</td>
</tr>
<tr>
<td>arad_salty</td>
<td>ML EDLKRQVLEANLAPKHN L V</td>
<td>M L L L--</td>
<td>22</td>
</tr>
<tr>
<td>arad_haein</td>
<td>ML AQLKXEVLEANLAPKHN L V</td>
<td>M N R A E S Q X I I D C T C EM T K L G N</td>
<td>23</td>
</tr>
<tr>
<td>fuca_ecoli</td>
<td>M N R A E S Q X I I D C T C EM T K L G N</td>
<td>M N R A E S Q X I I D C T C EM T K L G N</td>
<td>23</td>
</tr>
<tr>
<td>rhad_ecoli</td>
<td>MQNITQ SWFVGM I K A T T D L W K G D E R E N G N L L R L L I D A P Y H D N F H Q</td>
<td>M L L L--</td>
<td>50</td>
</tr>
<tr>
<td>rhad_salty</td>
<td>MQNITQ SWFVGM I K A T T D L W K G D E R E N G N L L R L L I D A P Y H D N F H Q</td>
<td>M L L L--</td>
<td>50</td>
</tr>
<tr>
<td>rhad_haein</td>
<td>MQNITQ SWFVGM I K A T T D L W K G D E R E N G N L L R L L I D A P Y H D N F H Q</td>
<td>M L L L--</td>
<td>50</td>
</tr>
<tr>
<td>fuca_haein</td>
<td>MQNITQ SWFVGM I K A T T D L W K G D E R E N G N L L R L L I D A P Y H D N F H Q</td>
<td>M L L L--</td>
<td>50</td>
</tr>
<tr>
<td>fuca_ecoli</td>
<td>MQNITQ SWFVGM I K A T T D L W K G D E R E N G N L L R L L I D A P Y H D N F H Q</td>
<td>M L L L--</td>
<td>50</td>
</tr>
<tr>
<td>rhad_ecoli</td>
<td>EW EGTKKPSSDTPTHRLLYQAFPSIGG IVHTHSRHATIAQ</td>
<td>I--P--G--Y--V--V--G--G</td>
<td>106</td>
</tr>
<tr>
<td>rhad_salty</td>
<td>EW EGTKKPSSDTPTHRLLYQAFPSIGG IVHTHSRHATIAQ</td>
<td>I--P--G--Y--V--V--G--G</td>
<td>106</td>
</tr>
<tr>
<td>rhad_haein</td>
<td>EW EGTKKPSSDTPTHRLLYQAFPSIGG IVHTHSRHATIAQ</td>
<td>I--P--G--Y--V--V--G--G</td>
<td>106</td>
</tr>
<tr>
<td>fuca_haein</td>
<td>EW EGTKKPSSDTPTHRLLYQAFPSIGG IVHTHSRHATIAQ</td>
<td>I--P--G--Y--V--V--G--G</td>
<td>106</td>
</tr>
<tr>
<td>fuca_ecoli</td>
<td>EW EGTKKPSSDTPTHRLLYQAFPSIGG IVHTHSRHATIAQ</td>
<td>I--P--G--Y--V--V--G--G</td>
<td>106</td>
</tr>
<tr>
<td>rhad_ecoli</td>
<td>NIV EKNKKPSSTPHTHELEYQRFPSSH IG IVTHHSRHATIAQ</td>
<td>I--P--G--Y--V--V--G--G</td>
<td>106</td>
</tr>
<tr>
<td>rhad_salty</td>
<td>NIV EKNKKPSSTPHTHELEYQRFPSSH IG IVTHHSRHATIAQ</td>
<td>I--P--G--Y--V--V--G--G</td>
<td>106</td>
</tr>
<tr>
<td>rhad_haein</td>
<td>NIV EKNKKPSSTPHTHELEYQRFPSSH IG IVTHHSRHATIAQ</td>
<td>I--P--G--Y--V--V--G--G</td>
<td>106</td>
</tr>
<tr>
<td>fuca_haein</td>
<td>NIV EKNKKPSSTPHTHELEYQRFPSSH IG IVTHHSRHATIAQ</td>
<td>I--P--G--Y--V--V--G--G</td>
<td>106</td>
</tr>
<tr>
<td>fuca_ecoli</td>
<td>NIV EKNKKPSSTPHTHELEYQRFPSSH IG IVTHHSRHATIAQ</td>
<td>I--P--G--Y--V--V--G--G</td>
<td>106</td>
</tr>
<tr>
<td>rhad_ecoli</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>rhad_salty</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>rhad_haein</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>fuca_haein</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>fuca_ecoli</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>rhad_ecoli</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>rhad_salty</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>rhad_haein</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>fuca_haein</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>fuca_ecoli</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>rhad_ecoli</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>rhad_salty</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>rhad_haein</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>fuca_haein</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
<tr>
<td>fuca_ecoli</td>
<td>YHILWNLTEAVPFGELFAHFLSHCERIATRIGKDRVIMCHATHMLIALTY</td>
<td>E--K--P--S--H--Y--P--G--H--H--H</td>
<td>150</td>
</tr>
</tbody>
</table>

Sequences were aligned using SEQSEE (Wishart et al., 1994). Key: arad, L-ribulose-5-phosphate 4-epimerase; fuca, L-fuculose-1-phosphate aldolase; rhad, L-rhamnulose-1-phosphate aldolase; ecoli, Escherichia coli; salty, Salmonella thyphimurium; haein, Haemophilus influenzae.
Members of the AraD/FucA Family

Figure C.1. Proteins Sharing a Homologous Domain with E. coli L-Ribulose-5-Phosphate 4-Epimerase. Key: BACSU, Bacillus subtilis; MYCPN, Mycoplasma pneumoniae; HAEIN, Haemophilus influenzae; ECOLI, Escherichia coli; SALTY, Salmonella typhimurium; METJA, Methanococcus jannaschii; ARAD, L-ribulose-5-phosphate 4-epimerase; FUCA, L-fuculose-1-phosphate aldolase; SGAE, probable sugar isomerase; YJFX, probable sugar isomerase SGAE; YIAS, probable sugar isomerase SGBE; YGBL, hypothetical 23.2 kDa protein; YE18, hypothetical protein.
Appendix D

Graphical Representation of Enzyme Kinetics

Figure D.1. Direct Plot of Kinetic Data from Zinc(I)-Substituted L-Ribulose-5-Phosphate 4-Epimerases. A, wild type; B, H95N; C, H97N; D, D76N.
Figure D.2. Direct Plots of Kinetic Data for Zinc(II)-Substituted L-Ribulose-5-Phosphate 4-Epimerases Assayed in the Presence of 0.1 mM Zn^{2+}. A, wild type; B, H95N; C, H97N; D, D76N.
Figure D.3. Direct Plots of Kinetic Data for Cobalt(II)-Substituted L-Ribulose-5-Phosphate 4-Epimerases.
A, wild type; B, H95N; C, H97N; D, D76N.
References


Definably Random Mutagenesis: Improving the Catalytic Potency of an Enzyme.

Hess, D., T.C. Covey, R. Winz, R. W. Brownsey, and R. Aebersold (1993). Analytical and
Micropreparative Peptide Mapping by High Performance Liquid
Chromatography/Electrospray Mass Spectrometry of Proteins Purified by Gel
Electrophoresis. Protein Sci. 2, 1342-1351.

Complete Sequence Analysis of the Genome of the Bacterium Mycoplasma pneumoniae.
Nucleic Acids Res. 24, 4420-4449.

Phosphoenolpyruvate Carboxykinase-Cobalt(III) Complex. Biochemistry 36, 3389-
3403.

Crystalline Thermolysin Suggests A Pentacoordinate Zinc Intermediate in Catalysis.
Biochemistry 20, 6912-6920.


Polymerase II is Homologous to the α-like DNA Polymerases. Mol. Gen. Genet. 226,
24-33.


for Site-Directed Mutagenesis Without PCR Product Purification. BioTechniques 12,
528-534.


