CHARACTERISATION OF STRUCTURAL AND BIOCHEMICAL PROPERTIES OF T. TENAX AND S. CARNOSUS FBP ALDOLASES, INVESTIGATING ALDOL CONDENSATION, ENANTIOPURITY AND THE POTENTIAL FOR THE CATALYZATION OF NOVEL PRODUCTS

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health.

2015

ADAM JAMES FLETCHER
List of Contents

List of Contents .............................................................................................................................................. 2
List of Tables .................................................................................................................................................. 5
List of Figures ................................................................................................................................................ 6
List of Equations and Schemes ..................................................................................................................... 8
List of Abbreviations ..................................................................................................................................... 9
Declaration .................................................................................................................................................... 12

1. INTRODUCTION ......................................................................................................................................... 16
   1.1. Carbon-Carbon Bond Formation .................................................................................................. 16
      1.1.1. The Baylis-Hillman Reaction ................................................................................................ 17
      1.1.2. Enzymatic Catalysis of the Baylis-Hillman Reaction .......................................................... 21
   1.2. The Aldol Condensation Reaction .................................................................................................. 23
      1.2.1. The Aldol Condensation Reaction Mechanism ..................................................................... 23
   1.3. Aldolase Enzymes .......................................................................................................................... 26
      1.3.1. Biotechnological Uses of Aldolases ....................................................................................... 27
      1.3.2. FBP-Aldolases ....................................................................................................................... 27
   1.4. Enzyme Design .................................................................................................................................. 36
      1.4.1. Altering Aldolase Stereochemistry Using Enzyme Design .................................................. 36
      1.4.2. Altering Enzyme Specificity ................................................................................................... 37
   1.5. Designing an Enzyme to Catalyze the MBH Reaction ................................................................. 39
      1.5.1. Comparison between the MBH and the Aldol Reaction ....................................................... 39
      1.5.2. Design of a Baylis-Hillman Enzyme .................................................................................... 40
   1.6. Thesis Aims ......................................................................................................................................... 41

2. MATERIALS & METHODS .......................................................................................................................... 42
   2.1. Materials ............................................................................................................................................ 42
      2.1.1. Enzymes and Chemicals .......................................................................................................... 42
      2.1.2. Escherichia coli Strains ............................................................................................................ 43
      2.1.3. Oligonucleotide Primers .......................................................................................................... 43
      2.1.4. Protein Purification Materials ................................................................................................. 45
      2.1.5. Buffers ........................................................................................................................................ 45
      2.1.6. Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) Components .... 45
   2.2. Determination of DNA Concentration ............................................................................................. 47
   2.3. Polymerase Chain Reaction (PCR) .................................................................................................... 47
      2.3.1. PCR Reaction ............................................................................................................................ 47
      2.3.2. PCR Product Purification ........................................................................................................ 48
   2.4. Agarose Gel Electrophoresis ............................................................................................................. 48
   2.5. DNA Restriction Digest ...................................................................................................................... 49
   2.6. Precipitation of DNA Using Ethanol ............................................................................................... 49
   2.7. Using the In-Fusion Advantage PCR Cloning Method .................................................................... 50
2.8. DNA Sequencing Reactions ................................................................. 50
2.9. E. coli Growth Conditions ................................................................. 50
2.10. Determination of E. coli Optical Density ........................................... 51
2.11. Transformation of Competent E. coli ................................................ 51
2.12. Purification of Plasmid DNA from E. coli ........................................... 51
2.13. Making Glycerol E. coli Stocks for Long Term Storage ....................... 52
2.14. Protein Expression and Purification ................................................... 52
  2.14.1. Protein Expression Trials ............................................................... 52
  2.14.2. Large Scale Protein Expression ..................................................... 52
  2.14.3. Lysis of E. coli by Sonication ....................................................... 53
  2.14.5. Anion Exchange Chromatography ............................................... 54
  2.14.6. Gel Filtration Chromatography .................................................... 54
  2.14.7. Nickel Affinity Chromatography .................................................. 54
  2.14.8. Size Exclusion Chromatography .................................................. 55
2.15. Measuring Protein Concentration ...................................................... 55
2.16. SDS-PAGE Analysis ........................................................................... 55
2.17. Multi-Angle Laser Light Scattering (MALLS) ...................................... 56
2.18. Chemical Synthesis and Analysis ...................................................... 56
  2.18.1. Synthesis of G3P and HBOP Using Dowex-50 Acid Resin .............. 56
2.19. Mass Spectrometry ............................................................................ 57
2.20. Biochemical Analysis .......................................................................... 57
  2.20.1. Steady State Reaction Kinetics ...................................................... 57
  2.20.2. Effect of GADH Titration on Reaction Kinetics ............................... 58
  2.20.3. Effect of Temperature on TtFBPA ................................................ 58
2.21. Nuclear Magnetic Resonance (NMR) Spectroscopy .............................. 59
  2.21.1. Phosphorus NMR Spectroscopy .................................................... 59
  2.21.2. Two-Dimensional Nuclear Magnetic Resonance Spectroscopy (2D NMR) Spectroscopy ................................................................. 59
2.22. X-ray Crystallography ........................................................................ 60
  2.22.1. Screening for Crystallization Conditions Using Commercial Screens ........................................................................................................... 60
  2.22.2. Screening for Crystallization Conditions Using Customized Screens ........................................................................................................... 60
  2.22.3. Cryo-protection of Crystals ......................................................... 61
  2.22.4. Co-crystallization and Ligand Soaking ......................................... 61
  2.22.5. Structure Determination of TtFBPA ............................................. 62
  2.22.6. Structure Determination of ScFruA .............................................. 62
3. EXPRESSION AND PURIFICATION OF TWO FBP ALDOLASES .......... 63
3.1. Background Information ..................................................................... 63
3.2. TtFBPA Heterologous Expression and Purification ............................... 63
3.3. ScFruA Expression and Purification ................................................... 68
  3.3.1. Re-cloning of ScFruA Gene ............................................................ 68
  3.3.2. Expression and Purification of ScFruA in E. coli .............................. 68
3.4. Multi-Angle Laser Light Scattering (MALLS) Analysis of Purified FBP Aldolases... 72
  3.4.1. MALLS Analysis of TtFBPA .......................................................... 72
  3.4.1. MALLS Analysis of ScFruA ......................................................... 76
3.5. Activity Assays of the Recombinantly Overexpressed FBP Aldolases ....... 78
  3.5.1. Reaction Kinetics of TtFBPA Proteins ........................................... 79
3.5.2. Steady State Retro-Aldol Activity of ScFruA .......................................................... 81

3.6. Determination of the Enzymatic Aldol Reaction .......................................................... 82
3.7. Discussion ...................................................................................................................... 83

4. CRYSTAL STRUCTURE OF TWO FBP ALDOLASES ....................................... 85
4.1. Crystallization of WT TtFBPA ..................................................................................... 85
4.1.1. Structure Determination of WT TtFBPA .................................................................. 87
4.1.2. Structure Determination of TtFBPA W144L ............................................................. 91
4.1.3. Co-crystallization of WT TtFBPA with DHAP ......................................................... 92
4.1.4. Co-crystallization of TtFBPA W144L with DHAP ................................................. 96
4.1.5. Co-crystallization of TtFBPA with the MBH Substrate, HBOP ................................ 97
4.2. Crystal Structure Determination of ScFruA .............................................................. 101
4.3. Discussion and Conclusion .......................................................................................... 106

5. NMR-BASED DETECTION OF ALDOLASE ACTIVITY ................................... 108
5.1. Collection of $^{31}$P-NMR of Reference Spectra .............................................................. 109
5.2. $^{31}$P-NMR of the Retro-Aldol Reaction ........................................................................ 111
5.2.1. $^{31}$P-NMR of the Retro-Aldol Reaction Using TtFBPA ........................................... 111
5.3. $^{31}$P-NMR Spectroscopy of the Aldol Reaction Using TtFBPA ..................................... 113
5.3.1. End Point Reaction of Aldol Reaction Using $^{31}$P-NMR Spectroscopy ............... 113
5.3.2. Time Course Reaction of Aldol Reaction Using $^{31}$P-NMR Spectroscopy ............ 116
5.3.3. Identification of TtFBPA Unknown Products ........................................................ 118
5.4. $^{31}$P-NMR Spectroscopy of ScFruA Aldol Activity ...................................................... 119
5.5. $^{31}$P-NMR spectroscopy aimed at Detecting of Baylis-Hillman Activity ..................... 121
5.5.1. Detecting TtFBPA Baylis-Hillman Reaction Activity ............................................. 121
5.6. 2D $^1$H/$^{13}$C-NMR HSQC Spectroscopy ......................................................................... 122
5.6.1. Reference Spectra .................................................................................................. 123
5.6.2. 2D NMR of TtFBPA reaction mixtures ................................................................. 125
5.6.3. 2D NMR of ScFruA Reaction Mixtures ................................................................. 130
5.1. Discussion ..................................................................................................................... 134
5.1.1. TtFBPA Activity .................................................................................................... 135
5.1.2. ScFruA Activity .................................................................................................... 136
5.2. Conclusions .................................................................................................................. 137

6. CONCLUSIONS AND IMPLICATIONS ................................................................... 139
6.1. Detailed Characterization of TtFBA Activity ............................................................... 140
6.2. Characterization of TtFBPA Baylis-Hillman Activity ................................................. 141
6.3. Determination of ScFruA Biochemical and Structural Properties ............................ 142
6.4. Implications and Further Investigations .................................................................... 143

7. BIBLIOGRAPHY ........................................................................................................ 145

8. APPENDIX ................................................................................................................. 153
8.1. Cloning Vectors ......................................................................................................... 153
8.2. Mass Spectrometry References .................................................................................. 155
List of Tables

Table 2.1: Strain information for *E. coli* cells used .............................................................. 43
Table 2.2: PCR primers for ScFruA ................................................................................... 44
Table 2.3: Buffer and reactions component with pH ........................................................... 46
Table 2.4: SDS-PAGE components .................................................................................... 47
Table 2.5: Composition of PCR reactions used to amplify ScFruA ................................. 48
Table 2.6: Temperature and time of PCR cycling steps ....................................................... 48
Table 3.1. Michaelis-Menten kinetics for retro-aldol activity of TtFBPA and ScFruA proteins ....................................................................................................................... 81
Table 4.1 X-ray crystallographic statistics for data collection, processing and refinement of TtFBPA structures in this study ................................................................. 88
Table 4.2: X-ray crystallographic statistics for data collection, processing and refinement of ScFruA structures in this study .................................................................... 103
List of Figures

Figure 1.1: The decameric structure of TtFBPA as published by Lorentzen et al. ............ 33
Figure 3.1: SDS-PAGE analysis of WT TtFBPA anion exchange ...................................... 65
Figure 3.2: SDS-PAGE analysis of W144L, W144Y and K177A TtFBPA anion exchange 66
Figure 3.3: SDS-PAGE analysis of WT, W144L, W144Y and K177A TtFBPA gel filtration. ..................................................................................................................... 67
Figure 3.4: SDS-PAGE analysis of the protein expression trials of E.coli BL21 (DE3) cells containing either pET28a::ScFruA or pET28b::ScFruA plasmid. ......................... 70
Figure 3.5: SDS-PAGE analysis of N- and C-His 6 ScFruA using metal ion affinity chromatography. ........................................................................................................ 71
Figure 3.6: SEC-MALLS chromatogram of WT TtFBPA at 1.8 mg/ml. ......................... 74
Figure 3.7: SEC-MALLS chromatogram of WT TtFBPA at 10 mg/ml. ......................... 74
Figure 3.8: SEC-MALLS chromatogram of W144L TtFBPA at 1.8 mg/ml. .................... 75
Figure 3.9: SEC-MALLS chromatogram of W144Y TtFBPA at 1.8 mg/ml. ................. 74
Figure 3.10: SEC-MALLS chromatogram of N-His 6 ScFruA at 1.8 mg/ml ................... 77
Figure 3.11: SEC-MALLS Chromatogram of C-His 6 ScFruA at 1.8 mg/ml. .................. 77
Figure 3.12: Steady state reaction kinetic data for TtFBPA proteins using FBP ........ 80
Figure 3.13: Steady state reaction kinetic data for ScFruA using FBP ......................... 82
Figure 4.1: Images of X and Y shaped WT TtFBPA crystals. ........................................... 86
Figure 4.2: TtFBPA crystal structure reveals a decameric oligomerisation. ................. 89
Figure 4.3: The active site pocket of WT TtFBPA with free phosphate bound .......... 90
Figure 4.4: Active site of ligand free W144L TtFBPA compared to the WT structure. 92
Figure 4.5: Active site of WT TtFBPA co-crystallised with DHAP. ............................ 94
Figure 4.6: An Excerpt from the Class I Aldolase mechanism showing the formation of stereogenic centre ................................................................................................ 95
Figure 4.7: Co-crystallization of W144L TtFBPA with DHAP compared to the ligand free WT TtBPA crystal structure. ................................................................. 97
Figure 4.8: Co-crystallization of WT TtFBPA with HBOP. 1) ....................................... 99
Figure 4.9: The proposed chemical structures of HBOP bound to Lys-177 via a Schiff base and Amine bond ................................................................. 100
Figure 4.10: Co-crystallization of W144L TtFBPA with HBOP ........................................ 101
Figure 4.11: Images of C-His$_6$ScFruA crystals ................................................................. 102
Figure 4.12: Structure of the ScFruA apo structure compared to the 2IQT homologous structure .................................................................................................................... 104
Figure 4.13: The active site pocket of ligand free ScFruA showing free phosphate bound. 105
Figure 4.14: The active site pocket of ScFruA containing DHAP .................................... 106
Figure 5.1: The inter-conversion of DHAP and Gem DioI DHAP ..................................... 110
Figure 5.2: $^{31}$P-NMR of reference compounds used in the aldol and retro-aldol reactions.. 110
Figure 5.3: $^{31}$P-NMR time course of the retro-aldol reaction using FBP and WT TtFBPA. 113
Figure 5.4: $^{31}$P-NMR spectroscopy of the aldol reaction using TtFBPA ..................... 115
Figure 5.5: $^{31}$P-NMR spectra of a time course reaction of WT TtFBPA aldol condensation. 118
Figure 5.6: $^{31}$P-NMR of the initial and final time course spectra for the WT TtFBPA time course reaction. .......................................................................................................... 119
Figure 5.7: $^{13}$P-NMR spectroscopy of ScFruA aldol activity ........................................... 120
Figure 5.8: $^{31}$P-NMR of WT TtFBPA Baylis-Hillman activity ........................................... 122
Figure 5.9: $^1$H/$^{13}$C HSQC NMR reference spectrum of FBP ................................. 124
Figure 5.10: $^1$H/$^{13}$C HSQC NMR reference spectrum of DHAP ............................... 124
Figure 5.11: $^1$H/$^{13}$C HSQC NMR reference spectrum of G3P ................................. 125
Figure 5.12: $^1$H/$^{13}$C HSQC NMR of the WT TtFBPA aldol reaction ....................... 127
Figure 5.13: $^1$H/$^{13}$C HSQC-TOCSY NMR of the WT TtFBPA aldol reaction ............ 129
Figure 5.14: $^1$H/$^{13}$C HSQC NMR of the ScFruA aldol reaction .................................. 131
Figure 5.15: $^1$H/$^{13}$C HSQC-TOCSY NMR of the ScFruA aldol reaction ................. 134
Figure 6.1: The proposed enzymatic Baylis-Hillman reaction and active site pocket catalyzed by TtFBPA ......................................................................................................... 144
Figure 8.1: Vector map for the pET28a vector ................................................................. 153
Figure 8.2: Vector map for the pET28a vector ................................................................. 154
Figure 8.3: ESI reference spectrum for FBP ................................................................. 155
Figure 8.4: ESI reference spectrum for DHAP ........................................................... 155
Figure 8.5: ESI reference spectrum for G3P ............................................................... 156
List of Equations and Schemes

Scheme 1.1: An overview of the MBH reaction .................................................. 17
Scheme 1.2: The MBH reaction mechanism proposed by Hill and Isaacs in 1990 .......... 19
Scheme 1.3: MBH reaction mechanism proposed by McQuade et al. ..................... 20
Scheme 1.4: The proposed reaction products of the McQuade mechanism ............... 21
Scheme 1.5: The stereogenic centers introduced through the Baylis-Hillman reaction .... 21
Scheme 1.6: The proposed MBH reaction mechanism proposed by Jiang and Yu ........ 22
Scheme 1.7: The aldol condensation reaction in the synthesis of the Wieland-Miescher ketone ......................................................................................................................... 23
Scheme 1.8: Keto-enol tautomerisation under acidic conditions ................................ 24
Scheme 1.9: The aldol condensation reaction under acidic conditions ....................... 24
Scheme 1.10: Keto-enol tautomerisation under basic conditions ............................... 25
Scheme 1.11: The aldol condensation reaction under basic conditions ...................... 25
Scheme 1.12: The different DHAP-dependent aldolases lead to distinct products ...... 28
Scheme 1.13: The proposed aldol condensation mechanism in human FBP aldolase ..... 31
Scheme 1.14: Sialic acid aldolase reaction from Hsu et al. 2005 ............................... 32
Scheme 1.15: Class II aldolase aldol reaction mechanism ........................................ 35
Scheme 1.16: The chemical structure of HBOP ...................................................... 40
Equation 2.1: The Michaelis-Menten equation .................................................... 58
Equation 2.2: ........................................................................................................ 58
List of Abbreviations

Abs  Absorbance
BBSRC Biotechnology and Biological Sciences Research Council
BSA Bovine Serum Albumin
CCD Charge coupled device
C-\text{His6ScFruA} \textit{Staphylococcus carnosus} fructose bisphosphate aldolase containing a C-terminal hexahistidine tag
DABCO 1,4-diazabicyclo[2.2.2]octane
DHA Dihydroxyacetone
DHAP Dihydroxyacetone phosphate
DNA Deoxyribonucleic acid
DNPH 2,4-Dinitrophenol hydrazine
DMSO Dimethyl sulphoxide
DTT Dithiothreitol
EDTA Ethylenediaminetetraacetic acid
ESI Electrospray ionization
\( F_\text{c} \) Calculated structure factor
\( F_\text{o} \) Observed structure factor
FBP Fructose-1,6-bisphosphate
FPLC Fast protein liquid chromatography
FSA Fructose-6-phosphate aldolase
G3P Glyceraldehyde 3-phosphate
GAPDH Glyceraldehyde 3-phosphate dehydrogenase
HBOP 1-hydroxy-3-buten-2-one phosphate
HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His6 Hexahistidine
HSQC Heteronuclear single quantum coherence spectroscopy
IPTG Isopropyl \( \beta \)-D-1 thiogalactopyranoside
IVD \textit{In Vitro} Diagnostic
LB Lysogeny broth
LOD Limit of detection
MALLS Multi-angle laser light scattering
MAD Multi-wavelength anomalous diffraction
MBH Morita-Baylis Hillman
MIR Multiple isomorphous replacement
NADH Nicotinamide adenine dinucleotide (reduced)
NEB New England Biolabs
NEC No enzyme control
N-\text{His6ScFruA} \textit{Staphylococcus carnosus} fructose bisphosphate aldolase containing a N-terminal hexahistidine tag
NMR Nuclear magnetic resonance
OD Optical density
PAD Pixel array detector
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA</td>
<td>Perchloric acid</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein data bank</td>
</tr>
<tr>
<td>PEG</td>
<td>Poly-ethylene glycol</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenyl Methyl Sulphonyl Fluoride</td>
</tr>
<tr>
<td>ppm</td>
<td>Parts per million</td>
</tr>
<tr>
<td>RAMA</td>
<td>Rabbit muscle aldolase</td>
</tr>
<tr>
<td>RLS</td>
<td>Rate Limiting Step</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root mean squared deviation</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>SAD</td>
<td>Single-wavelength diffraction</td>
</tr>
<tr>
<td>ScFruA</td>
<td><em>Staphylococcus carnosus</em> fructose-1,6-bisphosphate aldolase</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SEC</td>
<td>Size exclusion chromatography</td>
</tr>
<tr>
<td>SIR</td>
<td>Single isomorphous replacement</td>
</tr>
<tr>
<td>T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris acetate EDTA</td>
</tr>
<tr>
<td>TagA</td>
<td>Escherichia coli tagatose-1,6-bisphosphate aldolase</td>
</tr>
<tr>
<td>Tal</td>
<td>Transaldolase</td>
</tr>
<tr>
<td>TBP</td>
<td>Tagatose-1,6-bisphosphate</td>
</tr>
<tr>
<td>TBPA</td>
<td>Tagatose-1,6-bisphosphate aldolase</td>
</tr>
<tr>
<td>TE</td>
<td>Tris EDTA</td>
</tr>
<tr>
<td>TIM</td>
<td>Triose phosphate isomerase</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin layer chromatography</td>
</tr>
<tr>
<td>TLP</td>
<td>Thermolysin-like protease</td>
</tr>
<tr>
<td>TOCSY</td>
<td>Total correlation spectroscopy</td>
</tr>
<tr>
<td>TOF</td>
<td>Time of flight</td>
</tr>
<tr>
<td>TtFBPA</td>
<td><em>Thermoproteus tenax</em> fructose-1,6-bisphosphate aldolase</td>
</tr>
<tr>
<td>XDS</td>
<td>X-ray detector software</td>
</tr>
<tr>
<td>ε</td>
<td>Extinction coefficient</td>
</tr>
</tbody>
</table>
Abstract

Characterisation of structural and biochemical properties of *T. tenax* and *S. carnosus* FBP aldolases, investigating aldol condensation, enantiopurity and the potential for catalyzation of novel products.

Adam James Fletcher
The University of Manchester 2015 Doctor of Philosophy

The Morita-Baylis-Hillman (MBH) reaction is a carbon-carbon (C-C) bond forming reaction between an activated alkene and an aldehyde. It is a synthetically useful reaction due to the high atom economy and retention of multiple functional groups. Unfortunately, harsh reaction conditions are required during the MBH reaction and unpredictable product stereospecificity have hampered the widespread application of this reaction.

Catalysis of the MBH reaction by enzymes has the potential to allow the reaction to occur at ambient conditions, while offering scope for improving the stereospecificity. This thesis focussed on the enzyme design of a MBH enzyme using thermostable fructose-1,6-bisphosphate (FBP) aldolases as scaffolds. These enzymes were chosen because there are common features between the aldol and MBH reactions, both making use of an enol intermediate to attack the aldehyde. In addition, aldolases typically accept a wide variety of substrates. Thermostable aldolases were selected for increased temperature tolerance creating a more desirable catalyst for industrial purposes.

*Thermoproteus tenax* FBP aldolase (TtFBPA; WT and W144L, W144Y, K177A variants) and *Staphylococcus carnosus* FBP aldolase (ScFruA) were expressed and purified from *E. coli*. While the retro-aldol reaction catalysed by these enzymes could be easily monitored, the reverse reaction (aldol synthesis) is more difficult to quantify. Multiple methodologies for high throughput spectrophotometric detection of aldol activity were developed as a method of monitoring constructs made during directed evolution of the FBP aldolases. However, none of these proved successful in robustly determining aldol activity.

The dihydroxyacetone phosphate (DHAP) mimic 1-hydroxy-3-buten-2-one phosphate (HBOP) was used to assay for MBH catalysis. While crystallographic studies with TtFBPA suggest that HBOP is bound to W144L TtFBPA in a manner compatible with the MBH reaction. NMR studies could not detect any corresponding activity. This suggests further protein engineering will be required to evolve this FBP aldolase to an MBH catalyst. In addition, our crystallographic and NMR studies with TtFBPA reveals this enzyme is capable of catalysing the formation of both FBP and tagatose-1,6-bisphosphate (TBP).

Additionally, we determined the first structure of ScFruA. Interestingly, NMR experiments suggested ScFruA lacks significant control of the stereospecificity of the aldol condensation reaction and appears to catalyse the formation of FBP, TBP, xyluose-1,6-bisphosphate and psicose-1,6-bisphosphate. We conclude that while FBP aldolases could indeed provide useful scaffolds for the development of an MBH catalyst, the enzymes lack any inherent activity, necessitating the need for future creation of variants. The success of this approach will depend on the ability to screen mutant libraries for MBH product formation.
No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
Acknowledgements

First and foremost, I would like to thank Professor David Leys, Dr. Johanna Avis and Dr. David Berrisford for their invaluable support and guidance throughout this project.

For their contributions to the laboratory work, I would like to thank Dr Mark Dunstan, who helped with performing crystallography trials and his guidance with structural refinement. For their help performing NMR, I would like to thank Andrea Hounslow at the University of Sheffield and Dr Matthew Cliff at the University of Manchester. I would also like to thank Dr Karl Payne for aiding me with enzyme kinetics experimentation.

For their assistance and encouragement with the laboratory work I would like to thank Dr. Karl Fisher and Dr. Natalie Stephenson. For putting up with me throughout my time as a PhD student I would like to thank my family – without their emotional and financial support I would not have gotten this far.

Finally I would like to thank the Biotechnology and Biological Sciences Research Council (BBSRC) for the funding they have provided for this project, without which this research could not have been undertaken.
The Author

The author of this thesis, Adam James Fletcher, received a 2:1 Honours degree in Biology with Industrial Experience at the University of Manchester in 2011. During his degree he spent 12 months working for a small biotechnology company called DxS, where he developed a system of giving greater resolution in allele specific PCR primer design. When he returned to the University of Manchester he completed his undergraduate project in Professor Richard Reece’s laboratory, looking at the potential nuclear migration of regulatory proteins in the \textit{S. cerevisiae} GAL switch. During this time he also worked part-time at DxS/QIAGEN where he developed an \textit{In Vitro} Diagnostic (IVD) screening test for BRAF mutations in cancer to aid in drug treatment. In September 2011 he began the research that makes up this thesis within Dr. Johanna Avis’s laboratory. During this Ph.D. he moved to Dr. David Ley’s laboratory where he completed this work. He is currently working as a Product Development Scientist at Premaitha Health PLC, developing new IVD products for chromosomal trisomy detection using next generation sequencing techniques.
Chapter One

Introduction

1. INTRODUCTION

1.1. Carbon-Carbon Bond Formation

Carbon-carbon (C-C) bonds form the basis of organic chemistry, providing a framework for the production of complex (biological) molecules. The creation of the relatively stable C-C bonds also leads to the addition of a putative stereogenic centre, increasing molecular complexity (Basavaiah, Rao, & Satyanarayana, 2003; Resch, Schrittwieser, Siirola, & Kroutil, 2011). Hence, C-C bond forming reactions play a pivotal role, not only in the synthesis of complex biological molecules such as nucleotides and carbohydrates, but also in the synthesis of useful organic molecules such as pharmaceuticals (e.g. Taxol) (Guillena, Nájera, & Ramón, 2007; Kingston, 2000; Nicolaou et al., 1994). As such, the formation of C-C bonds is highly desirable and the subject of great interest (C.-J. Li, 2005; Miao, Rahimi, Geertsema, & Poelarends, 2015).

In 2010, research into the use of palladium as a catalyst for C-C bond formation was awarded the Nobel Prize in chemistry. This research, conducted by Richard Heck, Ei-ichi Negishi and Akira Suzuki, showed organopalladium successfully catalysed coupling reactions of iodobenzene and styrene resulting in the formation of stilbene via the Heck reaction [Beletskaya & Cheprakov 2000; Tsuji 1969; Zapf & Beller 2002]. In biology, C-C bond forming reactions occur at mild conditions, with high regio- and stereoselectivity (Miao et al., 2015). The key C-C bond-forming reactions focused on in this thesis are discussed in detail below.
1.1.1. The Baylis-Hillman Reaction

The Morita-Baylis-Hillman (MBH) reaction was first discovered in 1968, by Morita et al. (Morita, Suzuki, & Hirose, 1968), when they reacted acrylonitriles and methyl acrylates with aldehydes using an alkyl phosphine catalyst. In 1972, Baylis and Hillman showed a similar reaction with nucleophilic tertiary amine catalysts, such as 1,4-diazabicyclo[2.2.2]octane (DABCO) (A. B. Baylis, 1972; Amarante et al., 2009). This reaction has received considerable interest since the mid-1990s, due to two important features of the reaction: the atom economy and the fact that products are densely functionalized and typically contain an electron withdrawing group (Basavaiah et al., 2003; Basavaiah & Veeraraghavaiah, 2011; Carrasco-Sanchez, Simirgiotis, & Santos, 2009; McDougal & Schaus, 2003). Furthermore, the MBH reaction allows the conversion of cheap starting materials into highly functionalized synthons for use in further reactions (Basavaiah et al., 2003). However, it is not widely used in synthetic chemistry due to its slow reaction rate and inability to control diastereoselectivity (Mansilla & Saá, 2010).

1.1.1.1. The MBH Reaction Mechanism

The MBH reaction is a three-component C-C bond forming reaction that involves the coupling of the $\alpha$-carbon of an activated alkene (such as an acrylate) with a carbon electrophile (such as an aldehyde or imine; Scheme 1.1) (Carrasco-Sanchez et al., 2009).

Scheme 1.1: An overview of the MBH reaction
Hill and Isaacs proposed a three-step MBH reaction mechanism in 1990, which has been widely accepted (Scheme 1.2) (Hill & Isaacs, 1990). The mechanism initially involves the Michael addition of a nucleophilic catalyst (such as the tertiary amine DABCO) with the alkene group of a α,β-unsaturated carbonyl resulting in the formation of a conjugated enolate ion (Carrasco-Sanchez et al., 2009; Mansilla & Saá, 2010). This enolate ion performs a nucleophilic attack on the aldehyde, resulting in the formation of a new C-C bond, this is proposed to be the rate-limiting step (Marko, Giles, & Hindley, 1997). This is then followed by the elimination of the amine catalyst concomitant with deprotonation.

However, this reaction mechanism has received some criticism, due to its inability to account for the presence of unwanted diaxanone byproducts (Iwabuchi, Nakatani, Yokoyama, & Hatakeyama, 1999). Hence, a second reaction mechanism was proposed by McQuade, which may be able to explain the formation of diaxanone and also explains the difficulty in controlling reaction stereochemistry (Scheme 1.3) (Masson, Housseman, & Zhu, 2007; Price, Broadwater, Walker, & McQuade, 2005).

In the McQuade mechanism the initial catalyst binding ($k_1$) and nucleophilic attack of the electrophile by the enolate ($k_2$) is the same as that proposed by Hill and Isaacs. However, in the McQuade mechanism, the product of this step performs a second nucleophilic attack on an electrophile to yield a hemiacetal intermediate. The latter then undergoes deprotonation, which is hypothesized to be the rate-limiting step (RLS) (Carrasco-Sanchez et al., 2009). This deprotonation yields a cyclic intermediate, resulting in catalyst ejection causing the formation of a second hemiacetal intermediate.
Scheme 1.2: The MBH reaction mechanism proposed by Hill and Isaacs in 1990

This second hemiacetal intermediate is able to undergo two different intramolecular transesterification reactions: the first leads to the formation of a MBH product, whilst the second reaction leads to transesterification between the hydroxyl (highlighted: *) and the R group of the alkene, leading to the formation of cyclic dioxanone, seen frequently as a byproduct (Scheme 1.4) (Miao et al., 2015). This mechanism also explains the difficulty in controlling the MBH reaction stereochemistry, due to the presence of multiple stereogenic centers leading to different stereoisomeric products (Illustrated by wavy bonds in Scheme 1.5).

Recent research on the MBH reaction has focused on either improving the diastereoselectivity of the reaction (Masson et al., 2007), or the use of the aza-Baylis-Hillman reaction variant, where an imine group is used in lieu of an aldehyde as the carbon
electrophile (Declerck, Martinez, & Lamaty, 2009; Lindner, Liu, Karaghiosoff, Maryasin, & Zipse, 2013; Martinez, 2009; Masson et al., 2007).

Scheme 1.3: MBH reaction mechanism proposed by McQuade et al.
1.1.2. Enzymatic Catalysis of the Baylis-Hillman Reaction

There have been several recent studies showing catalysis of a Baylis-Hillman type reaction by proteins. Initial catalytic activity was reported by Reetz *et al.* in 2007, using bovine serum albumin (BSA; Fraction V), although the reaction rate was very slow (15% product conversion over 48 hours)(Reetz, Mondière, & Carballeira, 2007). Both Lopez-Iglesias *et al.* and Jiang and Yu reported Baylis-Hillman activity between nitrobenzaldehyde and a methyl vinyl ketone using proteases and biotin esterase, with a 46% conversion rate over 96 hrs (Jiang & Yu, 2014; López-Iglesias & Gotor-Fernández, 2015)
Jiang and Yu reported that *E. coli* BioH esterase showed high levels of catalytic promiscuity for the Baylis-Hillman substrates. The production of Baylis-Hillman products were analyzed and both TLC and $^1$H-NMR (Alcaide, Almendros, Aragoncillo, & Rodríguez-Acebes, 2004). Baylis-Hillman activity was observed only in aldehydes with strong electron withdrawing groups, but while the substrate range of the activated alkene is much wider. Kapoor and Majumder in 2014 demonstrated a lipase-catalyzed Baylis-Hillman reaction (Kapoor, Majumder, & Gupta, 2014). Over the course of the reaction two products were formed; a Baylis-Hillman adduct and an aldol adduct (Kapoor et al., 2014).

**Scheme 1.6: The proposed MBH reaction mechanism proposed by Jiang and Yu**

![Scheme 1.6: The proposed MBH reaction mechanism proposed by Jiang and Yu](image)
1.2. The Aldol Condensation Reaction

The aldol condensation reaction is a reversible C-C bond forming reaction involving an enol or enolate ion reacting with a carbonyl molecule to form a β-hydroxy carbonyl (such as an aldehyde) called an aldol (Machajewski & Wong, 2000; Slough, 1989). The aldol condensation reactions are both biologically and industrially useful as a relatively simple method of forming C-C bonds to yield a more complex molecule. A common industrial use for the aldol reaction is the synthesis of the Wieland-Miesher ketone (Scheme 1.7) (List, Lerner, & Barbas III, 2000; List, Shabat, Barbas III, & Lerner, 1998; López-Iglesias & Gotor-Fernández, 2015). This is a highly versatile synthon utilised in numerous pharmaceutical products including one synthesis method of the anti-cancer drug Paclitaxel (Taxol) by the Danishefsky method (Danishefsky et al., 1996; Kingston, 2000; Su et al., 1997)

Scheme 1.7: The aldol condensation reaction in the synthesis of the Wieland-Miesher ketone

1.2.1. The Aldol Condensation Reaction Mechanism

The aldol reaction can be performed under both basic and acidic conditions, as both can lead to the inter-conversion between the ketone and enol/enolate tautomers (Scheme 1.8 and Scheme 1.10) (Guillena et al., 2007).
Scheme 1.8: Keto-enol tautomerisation under acidic conditions

Under acidic conditions, the aldehyde becomes protonated rendering the molecule highly electrophilic, allowing a nucleophilic attack from the \(\alpha\)-carbon of the enol. This results in the formation of new carbon-carbon bond, via aldol addition. This molecule then undergoes condensation through the removal of water, yielding a \(\beta\)-hydroxyl carbonyl (Scheme 1.9)(Beletskaya & Cheprakov, 2000).

Scheme 1.9: The aldol condensation reaction under acidic conditions
Scheme 1.10: Keto-enol tautomerisation under basic conditions

To complete the aldol addition reaction the enolate ion nucleophilically attacks the electrophilic carbonyl carbon of the aldehyde molecule, which causes electron migration to form an alkoxide salt with a metal ion (M⁺). The reaction is completed through the condensation of water with the alkoxide group (Scheme 1.11)(Falcicchio, Wolterink-Van Loo, Franssen, & van der Oost, 2014).

Scheme 1.11: The aldol condensation reaction under basic conditions
The aldol reaction results in the formation of products that can contain multiple stereogenic centres, producing a racemate. However, it has been shown that the inclusion of a proline catalyst can allow enantioselectivity for the aldol reaction further increasing its potential industrial applications (Guillena et al., 2007; Liu et al., 2008).

### 1.3. Aldolase Enzymes

Aldolase enzymes are capable of catalysing aldol reactions in nature and function in particular during the catabolism of highly oxygenated metabolites (retro-aldol reaction) or during the synthesis of carbohydrates and amino acids (aldol reaction) (Gijsen, Qiao, Fitz, & Wong, 1996; Samland & Sprenger, 2006). Many aldolase enzymes are capable of controlling the reaction stereoselectivity (Machajewski & Wong, 2000).

There are two classes of aldolase enzymes found in nature: Class I aldolases catalyse the reversible formation/cleavage of C-C bonds through a catalytic lysine residue (Bednarski et al., 1989), while Class II aldolases catalyse the reaction using a divalent metal ion (Blom, Tètreault, Coulombe, & Sygusch, 1996; Marsh & Lebherz, 1992).

Aldolase enzymes have been shown to typically display high degrees of specificity for the nucleophilic donor molecule, with a more relaxed specificity for the electrophilic acceptor molecule (Gijsen et al., 1996). As a result, aldolase enzymes are classified according to their primary donor molecule: dihydroxyacetone phosphate (DHAP) dependent aldolases, dihydroxyacetone (DHA) aldolases, pyruvate aldolases, acetaldehyde aldolases and glycine aldolases (Falcicchio et al., 2014).
The most exploited group of aldolases in synthetic chemistry are the DHAP-dependent enzymes, the most readily used of which are the fructose-1,6-bisphosphate (FBP) aldolases (Gijsen et al., 1996; Machajewski & Wong, 2000; Samland & Sprenger, 2006). There are four DHAP-dependent aldolases, which utilize either D-glyceraldehyde-3-phosphate or D-Glyceraldehyde as secondary substrates, which yield differential stereochemistry from one another around the C3 and C4 carbons of the final product (Scheme 1.12) (Suau, Álvaro, Benaiges, & López-Santín, 2006).

1.3.1. Biotechnological Uses of Aldolases

Given the readily available starting materials, DHAP-dependent aldolases have been used extensively in the asymmetric synthesis of multiple products such as deoxy sugars, cyclitols and more complex polycyclic biological molecules such as pancratistatin and (-)-syringolides (Samland & Sprenger, 2006).

Unfortunately, the major disadvantage of these DHAP-dependent aldolases is the cost and stability of the starting material DHAP, hampering the large scale utilisation of these enzymes (Schümperli, Pellaux, & Panke, 2007). Therefore, research has focused on improving the substrate scope of the aldolase and overcoming the dependency on DHAP.

1.3.1. FBP-Aldolases

FBP aldolases function in the metabolic glycolysis/gluconeogenesis pathway to catalyse the reversible breakdown of FBP into two triose phosphates, DHAP and glyceraldehyde-3-phosphate (G3P). In glycolysis, FBP aldolase is key in the conversion of glucose into pyruvate, converting FBP into two 3-carbon molecules, which are further processed by the
glycolysis pathway (X. Li, Gu, & Zhou, 2015). In this pathway, glucose is initially phosphorylated, then isomerised into fructose-6-phosphate, and a further irreversible phosphorylation leads to FBP. FBP aldolase then converts FBP into DHAP and G3P, triose phosphate isomerase (TIM) then catalyses the DHAP:G3P interconversion (Wierenga, Kapetaniou, & Venkatesan, 2010).

Scheme 1.12: The different DHAP-dependent aldolases lead to distinct products

D-fructose-1,6-bisphosphate (FBP Aldolase)

D-tagatose-1,6-bisphosphate (TBP Aldolase)

G3P

Glyceraldehyde

L-fuculose-1-phosphate (Fuc-1P aldolase)

L-rhamnulose-1-phosphate (Rha-1P aldolase)
The enzyme G3P dehydrogenase then produces 1,2-bisphosphoglycerate, which is converted into pyruvate before it enters the Krebs cycle (X. Li et al., 2015; Richards & Rutter, 1961). Through the consumption of G3P, the glycolysis reaction sequence keeps the bidirectional FBP aldolase in the C-C bond breaking direction (the retro-aldol reaction). Conversely during gluconeogenesis, FBP aldolases catalyse the formation of a new C-C bond producing FBP (Say & Fuchs, 2010). The fructose-1,6-bisphosphatase enzyme irreversibly removes the C1 Phosphate from FBP, preventing an equilibrium being formed and ensuring FBP aldolase functions to form C-C bonds (X. Li et al., 2015). Two distinct types of FBP aldolases are found in nature, these are categorised into class I and II FBP aldolases.

1.3.1.1. Class I FBP Aldolases

Class I FBP aldolases are the most common variety of aldolase, found in the majority of cells, spanning both prokaryotic and eukaryotic organisms (Tittmann, 2014). Class I aldolases all contain a catalytic lysine residue within the active site, which covalently bonds to the carbonyl group of DHAP, forming a Schiff base intermediate. The catalytic mechanism of the class I FBP aldolases was determined using human, rabbit and fruit fly aldolases (Alefounder, Baldwin, Perham, & Short, 1989; Brenner-Holzach O, 1967; Meyeroff & Lohmann, 1934).

In human FBP aldolases the side chain ε-nitrogen (NH$_3^+$) of the catalytic lysine (Lys-229) points into the active site cavity (Morris & Tolan, 1994). The ε-nitrogen of Lys229 performs a nucleophilic attack on the carbonyl carbon of DHAP [Littlechild, J.A. & Watson 1993]. The Lys-229 ε-nitrogen is then deprotonated by Lys-146 allowing Asp-33 to donate a proton to the hydroxyl group of the intermediate releasing oxygen from the
molecule via dehydration. The resulting formation is a Schiff base between Lys-229 and DHAP (Scheme 1.13). Following the formation of the Schiff base/iminium intermediate, proton abstraction leads to a nucleophilic attack on G3P, which leads to the formation of the new C-C bond (H.P. Brockamp, 1990; Zannetti, Walter, Knorst, & Fessner, 1999).

The *Staphylococcus carnosus* FBP aldolase (ScFruA) has been observed to be unusually pH and temperature stable as well as accept a broad range of substrates [Brockamp & Kula 1990; Witke & Gotz 1993]. While ScFruA is likely to share significant active site homology with other aldolases, no structure has been reported (Zannetti et al., 1999). Additionally, ScFruA has been combined with DHAP synthesizing enzymes to make a ‘one-pot four enzyme system’ capable of producing sugar molecules directly [Li et al. 2012].

Another Class I aldolase family are Sialic acid aldolases, these enzymes naturally catalyse the reversible reaction between N-acetylmannosamine and pyruvic acid to Sialic Acid. Sialic acid aldolases typical of most aldolase enzyme comprising of eight repeating α/β barrels (a TIM barrel)(Hsu, Hong, Wada, Franke, & Wong, 2005). Enzyme engineering studies have been conducted on Sialic acid aldolases with the aim of broadening the substrate specificity of the enzyme and creating pharmaceutically useful products (Woodhall, Williams, Berry, & Nelson, 2005).

Various studies have looked at the specificity of Sialic acid aldolases, one such study showed that stereospecificity of the commercially bought enzyme can be altered thermodynamically without mutagenesis (Monosaccharides, 1992). Mutagenesis have however been conducted, one such study, conducted by Hsu *et al.* aimed at designing an enzyme capable of catalysing the formation of L-3-deoxy-manno-2-octulosonic acid (L-KDO). L-KDO is a useful intermediary in the development of bacterial cell-surface
selectors for the study of protein-carbohydrate interactions (Woodhall et al., 2005). After multiple mutagenesis-screening rounds were performed a total of eight amino acid substitutions resulted in an over 1000 fold improvement in the L-KDO specificity constant (kcat/km) relative to Sialic acid.

**Scheme 1.13: The proposed aldol condensation mechanism in human FBP aldolase**

![Diagram of the proposed aldol condensation mechanism in human FBP aldolase](image)
Another study by Woodhall et al. aimed at engineering Sialic acid mimetics which competitively inhibit sialidases. The drug Zanamivir inhibits influenza A and B sialidases although it is chemically synthesised. Woodhall et al. proposed a secondary synthesis mechanism using an enzyme-catalysed aldol reaction between an aldehyde and pyruvate followed by additional functional group manipulation steps.

**Scheme 1.14: Sialic acid aldolase reaction from Hsu et al. 2005**

![Scheme 1.14: Sialic acid aldolase reaction from Hsu et al. 2005](image)

*N*-Acetyl-D-mannosamine  \( \xrightarrow{\text{D-sialic acid aldolase}} \) D-Sialic acid

1.3.1.2. **Class IA FBP Aldolases**

The most common glycolysis pathways observed in biology are the Embden-Meyerhof and Entner-Doudoroff pathways, though many Archaea use modified versions (Sato & Atomi, 2011). Thermophilic archaea have been shown to contain a bifunctional FBP aldolase, for use in both glycolysis and gluconeogenesis, which shares little homology with bacterial or eukaryotic aldolases (Siebers et al., 2001). However, analysis of FBP aldolases from the hydrogen-sulphur autotrophic and hyperthermophilic archaea *Thermoproteus tenax* and *Pyrococcus furiosus* revealed that aldolase activity was not only metal-independent but also had a mechanism similar to Class I aldolases (E. Lorentzen et al., 2003; Esben Lorentzen, Siebers, Hensel, & Pohl, 2005). Hence, they were categorized as Class IA Aldolases (Siebers et al., 2001).
The *T. tenax* FBP aldolase (TtFBPA) is a hyper-thermostable protein that forms a 280kDa protein complex composed of two pentameric rings. Each monomer forms an \((\alpha\beta)_8\) barrel, or TIM barrel, structure which forms a dipentameric ring structure with a diameter of 100 Å and a height of 75 Å (Figure 1.1) (E. Lorentzen et al., 2003). Pentameric ring formation occurs due to the hydrophobic interactions between \(\alpha\)-helices and \(\beta\)-sheets, while five residues are thought to be involved in the formation of the decamer (E. Lorentzen et al., 2003).

Evolutionary comparison between archaeal Class IA aldolases and eukaryotic Class I aldolases showed there were six residues within the substrate-binding pocket of TtFBPA, which were found to have similar spatial distributions relative to type I aldolases (E. Lorentzen et al., 2003).

**Figure 1.1: The decameric structure of TtFBPA as published by Lorentzen et al.** The top and side view of the TtFBPA W144E, Y146F double mutant decamer structure and monomer structure
Furthermore, the phosphate moiety of DHAP is bound by the nitrogen backbone within both $\beta 7\alpha 7$ and $\beta 8\alpha 8$ of the TIM barrel, in both classes (E. Lorentzen et al., 2003). However, there is still some controversy surrounding which amino acids serve as a general acid in mediating Schiff base formation and which serve as a general base in mediating C3-C4 bond cleavage in Class I FBP Aldolases (Esben Lorentzen et al., 2005). Crystallographic data obtained by Maurady et al. suggested that Glu-187 could act as the acid (Maurady, Zdanov, de Moissac, Beaudry, & Sygusch, 2002), whilst Morris and Tolan proposed Asp-33 as the base mediating C3-C4 bond cleavage (Morris, Davenport, & Tolan, 1996).

An aspartate residue is conserved between eukaryotic and archaeal aldolases: Asp-24 in TtFBPA and Asp-33 in Human Aldolase I (Esben Lorentzen et al., 2005). These aspartate residues may serve the same catalytic role. The crystal structure showing a carbinolamine intermediate bound to TtFBPA, reveals the carboxyl group of Asp-24 to be just 2.9Å from the C4 hydroxyl of the carbinolamine intermediate, suggesting it is in the perfect position to assist in C3-C4 cleavage (E. Lorentzen et al., 2003).

1.3.1.3. Class II FBP Aldolases

Class II FBP aldolases are typically seen in fungi and some bacteria, such as Mycobacterium tuberculosis, Pseudomonas aeruginosa and Bacillus anthracis (Alefounder et al., 1989; Labbé et al., 2012). Unlike class I and IA aldolases, which catalyze the formation of FBP via a Schiff base, class II aldolases are metalloproteins which use a divalent ion as a cofactor (typically Zn$^{2+}$, but also Co$^{2+}$ and Fe$^{2+}$ (Marsh & Lebherz, 1992). This divalent metal ion is coordinated in the active site pocket by multiple histidines. Catalysis occurs through deprotonation of DHAP via a basic residue resulting in the formation of an enolate ion, with the divalent metal ion acting as the Lewis acid (Plater et
al., 1999). A subsequent nucleophilic attack from DHAP, to electrophilic G3P, results in the formation of FBP (Scheme 1.15).

As highlighted, many of the organisms containing metal-dependent class II aldolases cause infectious diseases in both humans and plants. As such, class II aldolases have been viewed as potential druggable target against bacteria and fungi, with research focusing on the use of metal chelating agents as Class II aldolase inhibitors [Labbé et al. 2012]. The study showed the successful binding of an inhibitor exhibiting sulfonate and phosphonate groups in addition to zinc chelating activity.

**Scheme 1.15: Class II aldolase aldol reaction mechanism**
1.4. Enzyme Design

1.4.1. Altering Aldolase Stereochemistry Using Enzyme Design

Utilizing enzymes capable of synthesizing enantiomerically-pure products is vital in pharmaceutical production as stereoisomers of small molecule therapeutics can have dramatically different properties. An example of side effects associated with a racemic drug was seen with Thalidomide (Ito, Ando, & Handa, 2011). This drug was used in the 1950s and 60s to combat morning sickness, to which R-Thalidomide is effective. However, the S-enantiomer is teratogenic, resulting in severe developmental defects of embryos exposed to the drug (Ito et al., 2011).

Improving the stereochemical outcome in aldolase reactions has been the subject of study, due to the two stereogenic centers produced during the aldol condensation reaction (Bolt, Berry, & Nelson, 2008). Various methodologies have been employed in an attempt to enhance enzyme stereochemical control including directed evolution (randomized mutagenesis or DNA shuffling and rational design) (Baker & Seah, 2012; Royer et al., 2010; Toscano, Müller, & Hilvert, 2007; Williams, Woodhall, Farnsworth, Nelson, & Berry, 2006).

Directed evolution has been successfully utilized to alter the stereospecificity of enzymes, such as tagatose-1,6-bisphosphate aldolase (TBPA) (Williams, Domann, Nelson, & Berry, 2003), which was successfully converted into a FBP aldolase, altering the stereochemistry of the aldol condensation reaction from (S)-C3 and (S)-C4 enantiomers to (S)-C3 and (R)-C4 enantiomers (Williams et al., 2003). In this study, three rounds of DNA shuffling were performed on TBPA, which identified four point mutations (H26Y, D104G, V121A and P256L) that had a 4:1 preference for FBP over TBP (Stemmer, 1994).
Rational design was conducted on class II aldolase BphI, which catalyzes the conversion of 4-hydroxy-2-oxopentanoate into acetaldehyde and pyruvate. The study successfully switched the stereochemistry of the aldolase to catalyze 4-hydroxy-2-oxohexanoate into propionaldehyde and pyruvate (Baker, Carere, & Seah, 2011; Baker & Seah, 2012). An initial enzyme design study compared BphI to the structurally similar enzyme DmpG, and used site directed mutagenesis based on key differences to modify BphI (Manjasetty, Powlowski, & Vrielink, 2003). This succeeded in increasing the affinity of BphI to propionaldehyde compared to its native substrate acetaldehyde (Baker et al., 2011). WT BphI displays highly selective nucleophilic attack of the pyruvate enolate si face to the aldehyde carbonyl producing 4(S)-hydroxyl-2-oxopentanoate. Mutagenesis of Tyr-290 resulted in the removal of stereochemical control producing both 4(S) and 4(R) products. Baker and Seah deduced that the hydroxyl group on Tyr-290 was able to sterically hinder the binding of the 4(R) enantiomer, therefore removing this resulted in a loss of stereochemical control (Baker & Seah, 2012).

1.4.2. Altering Enzyme Specificity

Enzymes typically show high levels of activity for a narrow range of substrates. However, aldolases show relatively high levels of substrate promiscuity. There have been multiple studies conducted that have attempted to alter or broaden the range of substrates in both the aldehyde acceptor and ketone donor species (Schümperli et al., 2007). Unfortunately, the enzyme has a much higher degree of specificity for the ketone donor species than the aldehyde acceptor (Falcicchio et al., 2014). This is problematic as DHAP is expensive, relatively unstable and contains a phosphate group, which is unlikely to be utilized in further synthesis steps (Z. Li et al., 2012). There have been multiple attempts to redesign the phosphate binding region of the active site pocket in DHAP-dependent aldolases to
allow these enzymes to accept a non-phosphorylated donor molecule (Garrabou et al., 2011; Schurmann, 2001). Garrabou et al. successfully showed the mutagenesis of rhamnulose-1-phosphate aldolase (a DHAP dependent aldolase) to accept DHA as the donor substrate (Garrabou et al., 2011).

Due to the limitations of donor specificity within DHAP-dependent aldolases, research interests have moved towards modifying the catalytic activity of fructose-6-phosphate adolases (FSA)(Schurmann, 2001; Sugiyama et al., 2007). FSAs display heightened donor substrate promiscuity accepting DHA, hydroxyacetone and hydroxybutanone (Schurmann, 2001; Soler et al., 2015). A number of studies have used rational design to alter the substrate specificity of FSAs to create biologically useful molecules, such as the modification of G3P by replacing the phosphate moiety with a Cbz protecting group, a useful group in industrial chemistry (Gutierrez, Parella, Joglar, Bujons, & Clapés, 2011; Sugiyama et al., 2007).

Mutagenesis studies of transaldolase (Tal) successfully removed the enzyme’s dependence on phosphorylated donor molecules, allowing it to accept DHA rather than DHAP (S. Schneider, Sandalova, Schneider, Sprenger, & Samland, 2008; Sarah Schneider et al., 2010). Mutagenesis of the Tyr-178 residue showed increased activity for non-phosphorylated donors by blocking phosphorylated molecules from binding (S. Schneider et al., 2008). Furthermore, additional mutagenesis rounds were performed which successfully removed the phosphate dependence of the acceptor aldehyde (Sarah Schneider et al., 2010).
1.5. Designing an Enzyme to Catalyze the MBH Reaction

1.5.1. Comparison between the MBH and the Aldol Reaction

The Mortia-Baylis-Hillman reaction and the aldol reaction have a large degree of similarity. The catalyst binding step in the MBH reaction involves the nucleophilic attack by a lone pair of electrons from the tertiary amine DABCO on the C=C group [Basavaiah & Veeraraghavaiah 2012]. The net result is the formation of an enolate ion. This is mirrored in the aldolase reaction where catalyst binding involves the nucleophilic attack of the catalytic lysine on DHAP, yielding an enolate ion (Miao et al., 2015). However, Schiff base formation in the aldolase reaction is mediated through an additional dehydration assisted by amino acids (Esben Lorentzen et al., 2005; Nes, 2005). The nucleophilic attack of the catalyst-bound species on the electrophilic acceptor results in the formation of new C-C bond, while bound to the catalyst in both the MBH reaction and aldolase reaction. The MBH reaction mechanism initially suggested by Hill and Isaacs is similar to that of the aldolase reaction; however, catalyst release is mediated by the formation of a C=C bond in the MBH reaction while catalyst release is mediated via a hydrolysis reaction resulting in the formation of a carbonyl group in the aldol reaction (Hill & Isaacs, 1990). Due to the number of similarities between the aldolase and MBH reactions, an FBP aldolase was determined as a good target for the production of a Baylis-Hillman catalysing enzyme. Hence, an \( \alpha,\beta \)-unsaturated ester structurally similar to DHAP was designed and synthesized by the Berrisford Group as a MBH substrate, this was named 1-hydroxy-3-buten-2-one phosphate (HBOP, Scheme 1.16) (Swiatyj, 2011).
1.5.2. Design of a Baylis-Hillman Enzyme

Two thermostable FBPAs (TtFBPA and ScFruA) were chosen as starting points for the design of an MBH enzyme. Thermostable enzymes are advantageous for their use in industrial chemistry, as enzymatic reactions could then be coupled to other reactions requiring higher temperatures. Additionally, ScFruA is monomeric, offering the possibility of being easily linked to a number of other enzymes, allowing whole reaction pipelines producing useful products from simple starting materials in an enzyme pathway (Witke & Gotz, 1993; Zannetti et al., 1999).

Initial mutagenesis conducted by the Avis lab produced a W144L and W144Y mutant of TtFBPA. Lorentzen et al. showed that a W144E/Y146F double mutant removed all aldol catalytic activity while still displaying substrate binding. Hence, it was likely that mutation of the tryptophan at position 144 to leucine, and, to a lesser extent, tyrosine would serve to enlarge to active site pocket of TtFBPA without affecting substrate binding properties. To determine if this is the case, structural determination of these aldolases, is necessary. Further rational enzyme design mutagenesis can then be conducted with the goal of enhancing the enzymes catalytic activity towards the MBH reaction.

Scheme 1.16: The chemical structure of HBOP
1.6. Thesis Aims

As mentioned above, aldolase enzymes could offer a starting point for designing a novel “Baylis-Hillmanase” enzyme. In short, aldolases accept a wide number of acceptor molecules, mutagenesis studies have successfully altered the donor substrate specificity, and the mechanisms of the Baylis-Hillman and aldolase reactions are similar. Furthermore, there are a number of aldolases that are known to be thermostable such as; the thermophilic archaeal FBP aldolase from \textit{T. tenax}, (TtFBPA; stable up to 100 °C) and the monomeric FBP aldolase from \textit{S. carnosus} (ScFruA; up to 60 °C) (Buchanan, Connaris, Danson, Reeve, & Hough, 1999; Z. Li et al., 2012). Therefore, the primary aim of this study is to further characterize these aldolases, as a starting point for the design of Baylis-Hillman activity. In order to characterize these enzymes a number of techniques will be utilized, including MALLS, NMR and X-ray crystallography. These will be aimed at determining the structure with the HBOP substrate, as well as test for activity with HBOP of both WT and mutated enzymes.

Furthermore, in order to assess the enzyme activity of these two aldolases a screening method is necessary that can be applied to high throughput screening. The latter is required to allow for the characterisation of large numbers of mutations produced through directed evolution.
Chapter Two

Materials and Methods

2. MATERIALS & METHODS

2.1. Materials

2.1.1. Enzymes and Chemicals

2.1.1.1. Enzymes

All restriction enzymes, ligases and polymerases purchased were obtained from New England Biolabs (NEB) with the exception of In-fusion polymerase (from Clontech), while rabbit muscle aldolase (RAMA), glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and triose phosphate isomerase (TIM) were purchased from Sigma-Aldrich.

2.1.1.2. Chemicals

General chemicals were purchased from Sigma-Aldrich, Fisher Scientific or Formedium with the exception of those listed below:

- Complete EDTA-free protease inhibitor tablets (Roche Diagnostics)
- Dialysis tubing (Medicell International Ltd.)
- Dithiothreitol (DTT, Fluka Analytical)
- Spin mini prep kit (QIAGEN)
- Phenyl Methyl Sulphonyl Fluoride (PMSF, ICN Biomedicals)
- Complete mini EDTA-free Protease Inhibitor Tablets (Roche Diagnostics)
2.1.2. *Escherichia coli* Strains

During the studies, multiple *E.coli* strains were used for plasmid propagation, protein expression or cloning. Table 2.1 details the *E.coli* strains used.

<table>
<thead>
<tr>
<th>Table 2.1: Strain information for <em>E.coli</em> cells used</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strain</strong></td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>DH5α</td>
</tr>
<tr>
<td>BL21(DE3)</td>
</tr>
<tr>
<td>Stellar competent</td>
</tr>
</tbody>
</table>

2.1.3. Oligonucleotide Primers

All oligonucleotide primers used in recloning experiments were purchased from Eurofins mwg|operon. The sequences of these primers are shown in Table 2.2.
<table>
<thead>
<tr>
<th>Vector</th>
<th>Description</th>
<th>Sequence</th>
<th>Tm (°C)</th>
<th>Restriction Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET28a</td>
<td>Forward Infusion Primer (Recloning)</td>
<td>5’-CGCGGCAGCCATATGAATCAAGAGCAGTTCGACAAATC-3’</td>
<td>83.44</td>
<td>Nde1</td>
</tr>
<tr>
<td>pET28a</td>
<td>Reverse Infusion Primer (Recloning)</td>
<td>5’-GGTGGTGCTCGAGTTAGGCTTTATTAACGCTGGCATC-3’</td>
<td>79.41</td>
<td>Xho1</td>
</tr>
<tr>
<td>pET28b</td>
<td>Forward Infusion Primer (Recloning)</td>
<td>5’-AGGAGATATACCATGAATCAAGAGCAGTTGCAC-3’</td>
<td>68.25</td>
<td>Nco1</td>
</tr>
<tr>
<td>pET28b</td>
<td>Reverse Infusion Primer (Recloning)</td>
<td>5’-GGTGCTCGAGGCTTTATTAACGCTGG-3’</td>
<td>79.47</td>
<td>Xho1</td>
</tr>
</tbody>
</table>

Underlined nucleotides denote restriction sites
*Additional G from Nco1 restriction site was removed during Infusion cloning
2.1.4. Protein Purification Materials

The materials used in protein purification steps are listed below:

- Q-Sepharose anion exchange resin (GE Healthcare)
- P6DG gel filtration resin (Bio-Rad)
- Ni-NTA resin (Qiagen)
- HisTrap column (Crude, GE Healthcare)
- Superdex-200 gel filtration column (GE Healthcare)

2.1.5. Buffers

The components of buffers used in experiments are listed in Table 2.3 below. pH measurements were taken at room temperature (approximately 20 °C).

2.1.6. Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Components.

The components of buffers used in SDS-PAGE are listed in Table 2.4 below.
<table>
<thead>
<tr>
<th>Buffer</th>
<th>Components</th>
<th>pH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>50x TAE Buffer</strong></td>
<td>242 g/L Tris-base, 57.1 mL Glacial Acetic acid, 0.5 mM EDTA</td>
<td>8.2-8.4</td>
</tr>
<tr>
<td><strong>6x DNA Loading Dye</strong></td>
<td>0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>E.coli Lysis Buffer</strong></td>
<td>50 mM Tris, 150 mM NaCl, 1 mM EDTA, 5 mM DTT, 0.1 mM PMSF</td>
<td>8.0</td>
</tr>
<tr>
<td><strong>Anion Exchange Buffer</strong></td>
<td>50 mM Tris, 1 mM EDTA, 5 mM DTT, 0.1 mM PMSF</td>
<td>8.0</td>
</tr>
<tr>
<td><strong>Gel Filtration Buffer</strong></td>
<td>25 mM HEPES, 150 mM NaCl, 1 mM EDTA, 5 mM DTT, 0.1 mM PMSF</td>
<td>7.0</td>
</tr>
<tr>
<td><strong>His Trap Buffer</strong></td>
<td>10 mM HEPES, 0.3M NaCl, 5 mM Imidazole</td>
<td>7.9</td>
</tr>
<tr>
<td><strong>SEC Buffer</strong></td>
<td>25mM HEPES, 150 mM NaCl</td>
<td>7.0</td>
</tr>
<tr>
<td><strong>Brady’s Reagent</strong></td>
<td>1.2 g 2,4-Dinitrophenol Hydrazine (DNPH), 8 mL H2O, 20 mL Ethanol, 6 mL Concentrated Sulphuric Acid</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>31P NMR Buffer</strong></td>
<td>25mM HEPES</td>
<td>8.05</td>
</tr>
<tr>
<td><strong>13C NMR Buffer</strong></td>
<td>25 mM Sodium Phosphate Buffer</td>
<td>7.98</td>
</tr>
</tbody>
</table>
### Table 2.4: SDS-PAGE components

<table>
<thead>
<tr>
<th>Name</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2x SDS Loading Dye</strong></td>
<td>62.5 mM Tris-HCl (pH 6.8), 2% Sodium Dodecyl Sulphate (SDS), 25% Glycerol, 2.5% β-mercaptoethanol, 0.01% Bromophenol Blue</td>
</tr>
<tr>
<td><strong>SDS Running Buffer:</strong></td>
<td>25 mM Tris-HCl, 250 mM Glycine, 0.1% SDS</td>
</tr>
<tr>
<td><strong>TGX Strain Free Gels</strong></td>
<td>4-20% Polyacrylamide Gels, premade from BioRad</td>
</tr>
<tr>
<td><strong>Instant Blue Stain:</strong></td>
<td>Premade from Expedeon</td>
</tr>
</tbody>
</table>

2.2. Determination of DNA Concentration

DNA concentration was determined using a Nanodrop Instrument (ND2000) following the manufacturer’s instructions for DNA quantification using $e_{260}/e_{280}$ ratio.

2.3. Polymerase Chain Reaction (PCR)

2.3.1. *PCR Reaction*

Nucleotide primers used in PCR are detailed in Table 2.2. PCR reactions were setup as detailed in Table 2.5, placed in a thermal cycler and run as detailed in Table 2.6. Following PCR, samples were digested with DpnI to remove methylated template DNA and analyzed on a 1% agarose gel.
Table 2.5: Composition of PCR reactions used to amplify ScFruA

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template DNA</td>
<td>100 ng</td>
</tr>
<tr>
<td>dNTP mix</td>
<td>200 μM</td>
</tr>
<tr>
<td>Pfu DNA Polymerase(^a)</td>
<td>2 Units</td>
</tr>
<tr>
<td>Forward Primer</td>
<td>200 nM</td>
</tr>
<tr>
<td>Reverse Primer</td>
<td>200 nM</td>
</tr>
<tr>
<td>10x PCR Reaction Buffer</td>
<td>2.5 μL</td>
</tr>
<tr>
<td>Sterile filtered water</td>
<td>Up to a final volume of 25 μL</td>
</tr>
</tbody>
</table>

\(^a\) Pfu UltraII fusion HS DNA polymerase enzyme (Stratagene)

Table 2.6: Temperature and time of PCR cycling steps

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
<th>Number of Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>95 °C</td>
<td>2 minutes</td>
<td>1</td>
</tr>
<tr>
<td>95 °C</td>
<td>20 seconds</td>
<td></td>
</tr>
<tr>
<td>T(_m)</td>
<td>30 seconds</td>
<td>30</td>
</tr>
<tr>
<td>72 °C</td>
<td>1 minute per kb</td>
<td></td>
</tr>
<tr>
<td>72 °C</td>
<td>3 minutes</td>
<td>1</td>
</tr>
</tbody>
</table>

------------------------  HOLD at 4°C  ------------------------

2.3.2. PCR Product Purification

PCR product purification was performed using the Molecular Biology Kit (NBS Biologicals Ltd.) following the manufacturer’s instructions.

2.4. Agarose Gel Electrophoresis

A 1 % agarose gel was prepared by heating 1 g agarose (Melford) in 100 mL 1x TAE buffer (from 50x TAE buffer) until molten. The molten agarose was allowed to cool to
approximately 50 °C and SimplyBlue SafeStain (Life technologies) was added to a final concentration of 0.5 µg/mL. The gel was then cast and cooled to room temperature. To prepare samples for loading, 1 µL of 6x DNA loading dye was added to 5 µL of DNA sample. Samples were loaded onto the gel alongside the GeneRuler™ 1 kb DNA ladder (Fermentas) as a DNA marker.

Agarose gel electrophoresis was run at 50 V in 1x TAE buffer (from 50x TAE buffer) and visualized using a UV transilluminator (Geneflash syngene bioimaging).

2.5. DNA Restriction Digest

Restriction digests were set up using approximately 1 µg of vector DNA and PCR products. The DNA was digested at the enzyme(s) optimum temperature for 1-2 hours in 2 µL of 10x reaction buffer as determined by New England Biolabs double digest guidelines. A total of 5-10 units of restriction enzyme(s) were used and the reaction made up to a final volume of 20 µL.

2.6. Precipitation of DNA Using Ethanol

0.1x reaction volume of 3 M sodium acetate (pH 5.2) was added to the DNA sample followed by 2-2.5x reaction volume of 100 % ethanol and incubated at -80 °C for 20 minutes. The sample was then centrifuged at 16,000 xg for 15 minutes at room temperature. The supernatant was removed and the pellet washed with 100 µL of 70% ethanol. The pellet was further centrifuged at 16,000 xg for 5 minutes at room temperature; the supernatant was then removed. The pellet was air dried to remove residual ethanol and resuspended in 10 µL dH2O.
2.7. Using the In-Fusion Advantage PCR Cloning Method

The Clontech online In-Fusion design tool was utilised to determine correct cloning primer sequence compositions and the correct primer-template ratios within the reaction based upon DNA length of the insert and template, while maintaining a ratio of a 2:1 insert to vector [http://bioinfo.clontech.com/infusion/molarRatio.do]. Reactions were setup with 3 μL 5x In-fusion buffer, 1.5 μL In-fusion enzyme (Clontech), the 100 ng of insert and 50 ng vector with the final reaction volume made up to 15 μL. The reaction was incubated at 37 °C for 20 minutes, followed by 50 °C for 15 minutes. 15 μL 15 mM TE buffer was added and the sample was mixed by inversion. 5 μL of this reaction was used to transform Stellar competent cells following manufacturer’s instructions.

2.8. DNA Sequencing Reactions

DNA samples were sent to Eurofins mwg|operon at a concentration of approximately 100 ng in 15 μL for Sanger sequencing targeted around the T7 polymerase. Sequence analysis and alignment was calculated using Serial Cloner (Serial Basics) to determine whether the insert was correctly inserted in the right direction into the plasmid [http://serialbasics.free.fr/Serial_Cloner.html] (Appendix Figure 8.1 and Appendix Figure 8.2).

2.9. *E. coli* Growth Conditions

*E. coli* was cultured in Lysogeny broth (LB) (1 % bacto-tryptone, 0.5 % yeast extract and 1 % NaCl), liquid *E. coli* cultures were incubated in a shaking incubator at 37 °C or 30 °C at 250 rpm overnight. *E. coli* was also cultured on LB-Agar plates which is
incubated at 37 °C overnight and stored at 4 °C for up to six weeks. The additions of antibiotics were used as a selection marker to a final concentration of 30 μg/mL kanamycin or 100μg/mL ampicillin.

2.10. Determination of *E. coli* Optical Density

*E. coli* cell concentration was determined by optical density measurements at 600 nm (OD$_{600}$) performed using a Cary 50Bio UV-Vis Spectrophotometer (Varian).

2.11. Transformation of Competent *E. coli*

Vials containing 30 μL aliquots of competent cells (Novagen) were thawed on ice, after which 1-2 μL of 50-100 ng/μl plasmid DNA was added. The cells were incubated on ice for at least one hour. The cells were then heat shocked at 42 °C for 45 seconds and transferred to ice for several minutes. 500 μL of SOC media was added and the cells were incubated at 37 °C for 40 minutes. Samples were plated on LB-Agar plate containing the appropriate antibiotics for selection and then incubated overnight at 37 °C.

2.12. Purification of Plasmid DNA from *E. coli*

A Spin MiniPrep Kit (QIAGEN) was used following the manufacturer’s instructions with the following alterations to produce a second, identical un-used plasmid stock for long-term storage. A 6 mL culture was grown overnight with the appropriate antibiotics and pelleted by centrifugation at 9,000 xg for 2 minutes. All buffer solution
volumes were doubled and then split into two equal reactions on the MiniPrep spin column. The purified plasmids were eluted in buffer EB and stored at -20 °C.

2.13. Making Glycerol E. coli Stocks for Long Term Storage

Cells containing the relevant plasmid were grown overnight in a shaking incubator at 37 °C at 200 rpm. A 600 μL volume of overnight culture was added to a sterile cryogenic vial with a rubber gasket in addition to 600 μL of sterile 40 % glycerol. The vials were mixed by pipetting and flash frozen using liquid nitrogen and stored at -80 °C.

2.14. Protein Expression and Purification

2.14.1. Protein Expression Trials

For protein expression trials 200 mL cultures were grown at both 30 °C and 37 °C in LB and auto-induction media (containing 30 μg/mL kanamycin). When an OD$_{600}$ of 0.6 was reached, IPTG was added to a final concentration of 0.5 mM to induce protein expression. Cells were grown for a further 20 hours at 30 °C or 6 hours at 37 °C and harvested by centrifugation in a Beckman Coulter Avanti J-26XP centrifuge using a JLA8.1 rotor. Cells were spun at 6900 xg for 10 minutes and the supernatant discarded. Cell pellets were then stored at -20 °C until required. Cell pellets were lysed and analysed by SDS-PAGE.

2.14.2. Large Scale Protein Expression

For large scale expression of protein, 50 mL of LB containing the appropriate antibiotics (either 100 μg/mL ampicillin or 30 μg/mL kanamycin) were inoculated
with a single colony of *E. coli* BL21 (DE3) containing the relevant plasmid and incubated overnight in a shaking incubator at 37 °C and 200 rpm.

A 10 mL volume of an overnight starter culture was inoculated into 1L of LB containing either 100 μg/mL ampicillin (for the TtFBPA) or 30 μg/mL kanamycin (in case of ScFruA) and incubated at 37 °C until an OD$_{600}$ of 0.6 was reached. IPTG was added to a final concentration of 0.5 mM to induce protein expression. Cells were then grown at either 37 °C for 6 hours (for the TtFBPA) or at 30 °C for 20 hours (in case of the ScFruA). Cells were harvested by centrifugation in a Beckman Coulter Avanti J-26XP centrifuge using a JLA8.1 rotor. Cells were spun at 6900 xg for 10 minutes and the supernatant was discarded. Cell pellets were then stored at -20 °C until required.

### 2.14.3. Lysis of *E. coli* by Sonication

Cell pellets were thawed and re-suspended in 50 mL lysis buffer in addition to one complete EDTA free protease inhibitor tablet (Roche) and 10 μg/μl DNase. Cells were mixed to homogeneity at 4 °C for 10 minutes then lysed by sonication using six thirty second bursts with a thirty second rest between each burst. Cell debris was removed by centrifugation using a Beckman Coulter Optima C-100XP Ultra centrifuge with a 45 Ti rotor at 125,000 xg for 1 hour and 10 minutes.

### 2.14.4. Heat Treatment of Cell Lysate

After lysis, samples were incubated at 90 °C for 40 minutes. Precipitated protein was removed by centrifugation using a Beckman Coulter Optima C-100XP Ultra centrifuge with a 45 Ti rotor at 125,000 xg for 1 hour and 10 minutes.
2.14.5. Anion Exchange Chromatography

Heat treated protein was firstly dialyzed into Anion Exchange buffer containing no salt (no salt buffer; Table 2.3) overnight. The sample was then applied to self-packed column containing 5 mL of Q-Sepharose resin, equilibrated with 5 column volumes of no salt buffer; the flow through used for SDS-PAGE analysis.

The column was washed with 10 mL fractions of no salt buffer, followed by a stepwise increase in concentration of NaCl in anion exchange buffer. The NaCl stepwise gradient was: 100 mM, 200 mM, 500 mM and 1 M NaCl. All fractions eluted were subjected to SDS-PAGE analysis.

2.14.6. Gel Filtration Chromatography

Protein purified using anion exchange was concentrated using a Vivaspin 20 centrifugal concentrator (MWCO 10,000 Da, Generon) and buffer exchanged by dialysis into gel filtration buffer (Table 2.3). The concentrated protein was applied to a self-packed column containing P60G gel filtration resin and 1 mL fractions of gel filtration buffer were applied and collected from the resin until 15 mL of buffer had been applied to the column.

2.14.7. Nickel Affinity Chromatography

Soluble *E. coli* lysate prepared (section 2.14.3) was applied to a 50 mL self-packed Ni²⁺-NTA resin (QIAGEN) that had been equilibrated with 5 column volumes of HisTrap Buffer, the flow-through was saved for SDS-PAGE analysis. The column was then washed alternating 10 mL fractions of either HisTrap Buffer or a stepwise increase in concentration of imidazole in HisTrap Buffer. The imidazole stepwise
_gradients used were: 50 mM, 100 mM, 200 mM and 500 mM imidazole. All washes and elusions were saved for SDS-PAGE analysis.

2.14.8. Size Exclusion Chromatography

Chromatography was conducted using a Superdex 10/200 GL size exclusion column (GE Healthcare) on an AKTA FPLC system. The column was equilibrated with 5 column volumes of filtered and degassed size exclusion chromatography (SEC) Buffer. A 200 μL volume of sample at approximately 10 mg/ml was injected onto the column; the column was then eluted at 0.3 mL/min and collected in 1 mL fractions. Fractions containing the peak protein elution were pooled and protein concentration quantified using a NanoDrop. Fractions were analysed by SDS-PAGE to determine protein size and purity. Protein concentration was adjusted to 10 mg/mL and flash frozen using liquid nitrogen for long term storage.

2.15. Measuring Protein Concentration

The concentrations of pure TtFBPA and ScFruA aldolases were calculated using a NanoDrop using the proteins 260/280 nm ratio (extinction coefficient’s 38390 and 10430 M⁻¹cm⁻¹ respectively).

2.16. SDS-PAGE Analysis

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was used to visualize protein samples to assess purity and molecular weight. A 10 μL volume of sample was added to 10 μL of 2x SDS Loading Buffer and boiled at 95 °C for 5
minutes. A 10 μL volume of the boiled sample was loaded onto a Biorad SDS-PAGE precast gel in addition to Biorad Precision Plus Stain-free Ladder. The gel was run at 200 V in SDS Running Buffer for approximately 20 minutes. Gels were then stained by submerging in InstantBlue (Expedeon).

2.17. Multi-Angle Laser Light Scattering (MALLS)

Protein sample was diluted to 1 mg/mL in filtered and de-gassed 25 mM HEPES pH 7.0 with 150 mM NaCl and run on a Superdex 200 10/300 GL size exclusion column (GE Healthcare). Samples were eluted at 0.25 mL/min through a flow cell in a DAWN-EOS MALLS spectrometer instrument (Wyatt Technology) with a quasi-elastic light scattering detector. From the light scattering data, Zimm fitting using the software provided with the instrument could derive molecular mass.

2.18. Chemical Synthesis and Analysis

2.18.1. Synthesis of G3P and HBOP Using Dowex-50 Acid Resin

The following protocol was adapted from the datasheet for DL-G3P diethyl acetal with the following alterations:

A 1.5 g quantity of Dowex-50 Acid resin was suspended in 3 mL dH₂O, to which 100 mg of either DL-G3P diethyl acetal monobarium salt or HBOP monobarium salt was added. The sample was then boiled at 95 °C for 3 minutes shaking every 15 seconds, and then chilled on ice for 5 minutes. The resin was then centrifuged at 4,000 xg for 30 minutes at 4 °C to separate resin from supernatant. The supernatant was collected and the resin was re-suspended in 1 mL H₂O and centrifuged again at 4,000 xg for 30
minutes at 4 °C with supernatant collected again. Both supernatants were combined and equilibrated to pH 7.0, flash frozen in liquid nitrogen and lyophilized. This process generated approximately 200 mmoles of DL-G3P or HBOP.

2.19. Mass Spectrometry

The Mass Spectrometry Facility in the University of Manchester conducted TOF/TOF spectrometry. For each sample, 20 μL of 1 mg/ml sample was injected into a Bruker Ultraflex II TOF/TOF spectrometer.

2.20. Biochemical Analysis

2.20.1. Steady State Reaction Kinetics

Steady state reaction kinetics were performed measuring the oxidation of nicotinamide adenine dinucleotide (NADH) in a coupled enzyme assay, by monitoring a decrease in absorbance at 340 nm (NADH $\varepsilon_{340} = 6.21 \text{ mM}^{-1} \text{ cm}^{-1}$) over two minutes at 50 °C.

Each steady state reaction contained 50 mM HEPES pH 7.0, 50 mM NADH, 5 μL GADH/TIM (75-200 U/mg) and between 1 μM and 500 mM FBP. The spectrophotometer was blanked using a reaction without NADH; a control without protein was also used.

All assays were initiated through the addition of 100 μg FBP aldolase, after which the cuvette was mixed by inversion and placed in the spectrophotometer; reactions were performed in triplicate to calculate an initial rate of NADH oxidation.
Initial rates for reactions were plotted against ligand concentration and data was fitted to a Michaelis-Menten function (Equation 2.1) using GraphPad Prism.

**Equation 2.1: The Michaelis-Menten equation**

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

\( V_{\text{max}} \) is the enzymes theoretical maximal rate at substrate saturation, \( [S] \) is the substrate concentration at rate \( v \), \( K_M \) is the substrate concentration where the rate is half the \( V_{\text{max}} \). From this information, the turnover number (\( k_{\text{cat}} \)) can be obtained (Equation 2.2).

**Equation 2.2:**

\[ k_{\text{cat}} = \frac{V_{\text{max}}}{[E_T]} \]

2.20.2. *Effect of GADH Titration on Reaction Kinetics*

In the GADH titration study, reactions were set up as described in section 2.20.1 with the following alterations; FBP concentration was maintained at 100 mM while the concentration of GADH and TIM were altered to: 4U/20U GADH/TIM, 8U/40U GADH/TIM and 18U/80U GADH/TIM.

2.20.3. *Effect of Temperature on TtFBPA*

In the temperature study, reactions were set up as shown in section 2.20.1 with the following alterations: FBP concentration was maintained at 100 mM, reactions were measured over 10 minutes and the temperature was altered to: 20 °C, 30 °C, 40 °C and 50 °C.
2.21. Nuclear Magnetic Resonance (NMR) Spectroscopy

2.21.1. Phosphorus NMR Spectroscopy

A total of 50 mM DL-G3P, 25 mM DHAP and 25 mM FBP each was reconstituted in $^{31}$P NMR Buffer, reference standard spectra were obtained with 500 μL of each substrate/product. A mixture containing DL-G3P (25 mM) and DHAP (12.5 mM; total volume 500 μL) was incubated at 50 °C for 30 minutes to determine degradation levels under these conditions. Enzymatic reactions were carried out using either 45 mM FBP or 23 mM DL-G3P and 11.5 mM DHAP and 100 μg of aldolase.

$^{31}$P-NMR experiments were acquired on a Bruker Avance 500 MHz NMR spectrometer equipped with a broadband direct detection cryoprobe with a varying number of scans, work was conducted by staff in the NMR facility within the University of Sheffield.

2.21.2. Two-Dimensional Nuclear Magnetic Resonance Spectroscopy (2D NMR) Spectroscopy

Total correlation spectroscopy (TOCSY) was performed to determine the chemical structure of products of the aldolase reactions by determining which nuclei are coupled together and secondly which nuclei are coupled by a chain of coupling.

A total of 50 mM DL-G3P, 25 mM DHAP and 25 mM FBP each were reconstituted in $^{13}$C-NMR Buffer. Reference compound spectra were obtained with 500 μL 50 mM DL-G3P, then 25 mM of DL-G3P and 12.5 mM DHAP made up to 500 μL which was incubated at 50 °C for 30 minutes. Enzymatic reactions were conducted with 23 mM DL-G3P and 11.5 mM DHAP and 100 μg of Aldolase.
Both $^1$H/$^{13}$C HSQC and $^1$H/$^{13}$C HSQC TOCSY NMR experiments were acquired on a Bruker Avance 800 MHz NMR spectrometer conducted by staff in the NMR facility within the University of Manchester.

2.22. X-ray Crystallography

2.22.1. Screening for Crystallization Conditions Using Commercial Screens

In order to determine potential crystallization conditions, the aldolases were tested against five commercially available crystallization screens: JCSG+, PACT premier™, Morpheus® and Clear Strategy™ Screens I and II crystallisation trial plates (Molecular Dimensions). These were prepared using a Mosquito liquid handling robot (Molecular Dimensions) on sitting drop 96-well trays (MRC Wheldan). The Mosquito liquid handling robot was used to pipette 200 nL of protein solution of varying concentrations, into the sample well, followed by addition of 200 nL of mother liquor. Trays were immediately examined using a stereoscopic light microscope (Nikkon) and stored at 4 °C for 24 hours before re-examination.

2.22.2. Screening for Crystallization Conditions Using Customized Screens

Customized crystallography screens were designed using the Hampton Research crystal tray design tool [http://hamptonresearch.com/make_tray.aspx]. The crystallisation conditions for TtFBPA were split into two different screening conditions. The first set of conditions used an increasing concentration of PEG 4000 from 6-10% in 1% increments. The second set of conditions consisted of a 2D matrix of increasing PEG 4000 concentration from 6-10%, in 1% increments, against increasing sucrose and ethylene glycol concentrations between 1-8%, again in 1%
increments. Both screening conditions contained a constant concentration of 0.1 M sodium acetate buffer (pH 5.0) and 12% glycerol throughout the screen.

The crystallisation conditions for ScFruA were adapted from the Clear Strategy 1 screen (position A2) which gave the best crystal morphology. This screen contained a constant concentration of 0.1 M sodium acetate buffer (pH 5.5) with a 2D matrix of increasing PEG 2000 MME concentration from 1-11% in 2% increments against an increasing lithium sulphate concentration of 150-300 mM in 25 mM increments.

### 2.22.3. Cryo-protection of Crystals

Once crystal hits were identified, samples were cryo-protected and mounted in a CryoLoop™ (Hampton Research) and flash-cooled in liquid nitrogen. In order to cryo-protect the crystals, 12% glycerol was added to the mother liquor.

### 2.22.4. Co-crystallization and Ligand Soaking

Co-crystallization and ligand soaking experiments are performed to introduce relevant molecules into the crystal lattice to observe protein-ligand interactions. Co-crystallization methods employ the addition of ligand while the protein is in solution and crystal trays are setup as normal. Soaking experiments attempt to introduce ligand into a protein crystal through water channels within the crystal.

For co-crystallization experiments, a suitable protein-ligand mixture was achieved through two different methods: the first simple addition of between 2-10 mM ligand to the concentrated protein solution. The second consisted of addition of 2 mM ligand to a 1mg/mL protein concentration, this sample was then concentrated to 10 mg/mL. In
ligand soaking experiments, solidified ligand was added to the drop containing protein crystals.

2.22.5. Structure Determination of TtFBPA

The *T. tenax* crystal structure bound to inorganic phosphate was determined previously (PDB ID: 1OJX) and used as initial model. This model was manually refined in an iterative process using the *Refmac5* and Coot in the CCP4 suite [Winn et al. 2011]. The resulting higher resolution model was then used as the initial template for all other TtFBPA models. Bespoke ligands were designed using the PRODRG server [Schüttelkopf & Van Aalten 2004].

2.22.6. Structure Determination of ScFruA

An initial model for the ScFruA was determined using the molecular replacement method (PhaserMR, CCP4 suite) with the *P. gingivalis* FBP Aldolase (PDB ID: 2IQT) as search model which has 55% homology. Following molecular replacement and initial model was constructed using Buccaneer (CCP4 Suite) [Winn et al. 2011] followed by ARP/wARP [Morris et al. 2003]. The ARP/wARP model further refined though an iterative process of refinement using Refmac5 followed by real space refinement using Coot.
3. EXPRESSION AND PURIFICATION OF TWO FBP ALDOLASES

3.1. Background Information

A Baylis-Hillman reaction catalyzing enzyme would be industrially useful as a novel method of producing C-C bonds with additionally highly functionalized groups. As this reaction is similar to that catalysed by aldolases, TtFBPA was chosen as a putative “Baylis-Hillmanase” due to its very high thermostability which is linked to the decameric state of the enzyme (Siebers, Klenk, & Hensel, 1998). In addition, a monomeric aldolase from a thermophilic organism was tested, ScFruA (Witke & Gotz, 1993).

Expression and purification experiments were conducted on TtFBPA and ScFruA with the aim of producing large levels of high purity protein for use in X-ray crystallography and NMR. MALLS analysis was conducted with the aim of determining whether TtFBPA is a decameric protein as described in the literature and whether ScFruA is a monomeric protein. In addition, the activity of these enzymes and selected variants was measured.

3.2. TtFBPA Heterologous Expression and Purification

Initial cloning experimentation performed by the Avis lab aimed to optimize expression and purification efficiencies of TtFBPA through the addition of either C or N terminal
polyhistidine tags. The introduction of these tags however led to the inactivation of the protein. Rather than perform refolding experimentation, which may not yield meaningful results, cloning of the *TtFBPA* gene would be conducted without a tag and the purification strategy would be based upon biochemical properties of the enzyme.

The *TtFBPA* genes (WT, W144L, W144Y and K177A variants) were cloned into the pET15b vector with the polyhistidine tag removed and used to transform BL21 (DE3) cells. The expression conditions published by Siebers *et al.* does not mention the incubation temperature and time for the *E. coli* cells. However, initial experimentation conducted by the Avis lab showed that incubation of 37 °C for 6 hours post induction, resulted in a high yield of soluble protein (Siebers *et al.*, 2001). Levels of insoluble and soluble protein obtained from large scale (24 L) expression using these conditions are shown in Figure 3.1 and Figure 3.2, lanes 2 and 3 respectively.

The purification conditions published by Seibers *et al.* show that an initial heat treatment step followed by Q-sepharose anion exchange purification resulted in pure protein product. Therefore the four *TtFBPA* proteins (WT and variants) were purified using these two steps. Prior to heat treatment, samples of raw cell lysate were removed for SDS-PAGE gel analysis. Following heat treatment at 95 °C, a further sample was removed for SDS-PAGE gel analysis (Figure 3.1 to Figure 3.2 A-C). The majority of the protein is expressed in soluble form and, following heat treatment, the majority of high molecular weight proteins and a number of lower molecular weight contaminants have been removed.

Following heat treatment, anion exchange chromatography was conducted, and samples from each anion exchange fraction were taken for SDS-PAGE gel analysis (Figure 3.1 and Figure 3.2 A-C, lanes 5 to 12). Pure protein was obtained in fractions eluting between 100-200 mM NaCl.
Following anion exchange, SDS-PAGE analysis showed small levels of contaminating proteins remained. Therefore, for crystallographic purposes, further purification of the protein was achieved using gel filtration chromatography (Figure 3.3; lane 2-8 (WT) and lane 9-15 (W144L); Figure 3.3; lane 2-8 (W144Y) and lane 9-15 (K177A)). This resulted in highly pure TtFBPA WT and variants suitable for crystallography. Proteins were concentrated to 12 mg/ml before being flash frozen and stored at -20 °C.
Figure 3.2: SDS-PAGE analysis of W144L, W144Y and K177A TtFBPA anion exchange. Coomassie stained SDS-PAGE analysis for W144L, W144Y and K177A shows the E. coli whole-cell insoluble fraction, whole-cell soluble lysate, heat treated soluble lysate, anion exchange flow through, wash step with native buffer, native buffer with 100 mM NaCl, second wash step with native buffer, 200 mM NaCl flow through, third wash step, and 1 M NaCl. Molecular weights are given for the Biorad Precision Plus Stain-free Ladder 15-150 kDa are shown.
Figure 3.3: SDS-PAGE analysis of WT, W144L, W144Y and K177A TtFBPA gel filtration. Coomassie stained SDS-PAGE analysis shows the gel filtration purification of WT, W144L, W144Y and K177A. SDS-PAGE analysis of these gel filtration experiments shows the heat-treated soluble protein fraction, flow through and the first five 1 mL fractions eluting from the gel filtration column for each enzyme. Molecular weights are given for the Biorad Precision Plus Stain-free Ladder (lane 1) 15-150 kDa are shown.
3.3. ScFruA Expression and Purification

3.3.1. Re-cloning of ScFruA Gene

A ScFruA construct was obtained from the Turner lab. However, the construct contained an elaborate His\(_6\)-tag, making it unsuitable for crystallographic purposes. The \textit{ScFruA} gene was re-cloned into two vectors, pET28a and pET28b, allowing the introduction of a short His\(_6\)-tag to either the N- (pET28a) or C-terminus (pET28b; Appendix Figure 8.1). In this case, His\(_6\) tagging was deemed appropriate as the Turner group had demonstrated ScFruA activity using His\(_6\) tagged constructs. To determine whether re-cloning of \textit{ScFruA} into the pET28a/b vectors was successful, sequencing was performed using vector specific T7 DNA polymerase primers.

3.3.2. Expression and Purification of ScFruA in \textit{E. coli}

The aim of the expression and purification trials for ScFruA was to identify a construct which allowed for large scale expression and high efficiency purification of the recombinant protein without the requirement for bulky tags which may affect crystallogenesis. Both pET28a and pET28b constructs were transformed into \textit{E. coli} BL21 (DE3) cells as an inducible expression system through the introduction of IPTG. Levels of protein expression were tested at a variety of different incubation temperatures with various media compositions using LB media induced with 0.5 mM IPTG or auto-induction media (Figure 3.4). The expression trials of \textit{ScFruA} containing an N-terminal His\(_6\) tag (N-His\(_6\)ScFruA; pET28a vector; Figure 3.4) show lower protein yields following incubation at 37 \(^\circ\)C compared to those observed following incubation at 30 \(^\circ\)C. The amount of protein present in the soluble and insoluble fractions was approximately equal following expression at 37 \(^\circ\)C; whereas at 30 \(^\circ\)C a higher level of protein was observed in the soluble
fractions compared to the soluble fractions. No difference was observed in expression and protein yields when comparing LB media with auto-induction media at 37 °C.

The expression trials of ScFruA containing a C-terminal His₆ tag (C-His₆ScFruA; pET28b vector; Figure 3.4) show protein yields were much lower at 37 °C compared to those obtained following incubation at 30 °C. As observed for N-His₆ScFruA, there was no difference in the protein yields using different media at 37 °C, while much higher levels of insoluble protein were observed for auto-induction media. Overall, both N- and C-His₆-ScFruA constructs resulted in high levels of soluble protein when using LB media induced with 0.5 mM IPTG at an OD_{600} of 0.6 followed by incubation for 20 hours at 30 °C. Large-scale protein expression was performed using these conditions. Cells were lysed by sonication with a probe sonicator. Following the addition of the soluble lysate to a Ni-NTA column, both N-His₆ScFruA and C-His₆ScFruA are eluted from the column between 50 mM – 200mM imidazole. Elusions at 100-200 mM imidazole show high levels of ScFruA with a minimum of contaminating proteins (Figure 3.5).
Figure 3.4: SDS-PAGE analysis of the protein expression trials of *E.coli* BL21 (DE3) cells containing either pET28a::ScFruA or pET28b::ScFruA plasmid. Coomassie stained SDS-PAGE analysis shows the total cellular soluble and insoluble fractions for *E.coli* BL21 (DE3) cells containing either pET28a::ScFruA or pET28b::ScFruA plasmid induced under different conditions, such as; using both LB and auto induction media, at temperatures of 30°C for 24 hours after induction or 37°C for 6 hours after induction. Molecular weights are given for the Biorad Precision Plus Stain-free Ladder 10-250 kDa are shown.
Figure 3.5: SDS-PAGE analysis of N- and C-HisScFruA using metal ion affinity chromatography. Coomassie stained SDS-PAGE analysis shows the expression profile for E.coli BL21 (DE3) cells containing both pET28a::ScFruA (N-terminal tag) and pET28b::ScFruA (C-terminal tag) plasmid. The gel shows total cellular soluble lysate before and after induction with IPTG, protein flow through and wash steps with native buffer. Protein was eluted with running buffer containing 50 mM, 100 mM, 200 mM and 500 mM imidazole. Each elution was followed by a wash step, rinsing the column with native buffer. Biorad Precision Plus Stain-free Ladder 20-250 kDa was used.
3.4. Multi-Angle Laser Light Scattering (MALLS) Analysis of Purified FBP Aldolases

3.4.1. MALLS Analysis of TtFBPA

MALLS was conducted to determine uniformity and the oligomerisation state of TtFBPA. The uniformity of the protein solution is useful to determine as a mono-disperse protein has a higher likelihood of crystallizing than poly-dispersed proteins. Furthermore, it is important to determine whether TtFBPA is a dodecamer in solution, as suggested by the previously published crystal structure.

Following purification, MALLS analysis of the WT TtFBPA was conducted at two concentrations (1.8 and 10 mg/ml). 1.8 mg/ml is a typical input concentration for MALLS while 10 mg/ml was chosen as high input concentration which is more closely mimicking protein crystallization conditions. A high input concentration was chosen to also determine whether protein concentration had an impact on oligomerisation, this concentration was only conducted WT TtFBPA.

The SEC-MALLS chromatogram for WT TtFBPA at 1.8 mg/ml (Figure 3.6) shows two peaks eluting around 8 mL and 15 mL. The initial minor peak begins to elute at 7.4 mL as part of the void volume, therefore the molecular weight of this fraction is above 28 MDa and is presumably an aggregation product of the protein. The major peak begins to elute at 14.4 mL reaching a peak at 15 ml. This elution peak contains a protein with an average molecular weight of approximately 280 kDa, which can be attributed to the molecular weight of the TtFBPA dodecamer.

In order to determine whether concentration had an impact on oligomerisation, MALLS analysis was also performed at 10 mg/ml to recreate concentrations similar to those used during crystal screening. The SEC-MALLS chromatogram for WT TtFBPA at 10 mg/ml
(Figure 3.8) shows a major peak, which elutes from the column between 11 and 14 mL, with a peak at 12.5 mL. This corresponds to a protein of approximately 280 kDa, which can be attributed to the dodecameric TtFBPA. The absorbance of this peak sharply decreases after the initial peak at approximately 280 kDa. The difference in elution volumes between the MALLS chromatogram at 1.8 mg/ml and 10 mg/ml can be accounted for by increased flow rate used with the 10 mg/ml sample. A minor peak eluting from the column between 15 and 17 mL shows a protein complex with a molecular weight of approximately 200 kDa. This complex could either be attributed to contaminants that are present at higher concentration in this sample.

In addition to the MALLS analysis of the WT enzyme, MALLS was also performed using both W144L and W144Y mutant aldolases at 1.8 mg/ml (Figure 3.8 and Figure 3.). The SEC-MALLS chromatograms for W144L and W144Y aldolases (1.8 mg/mL) show similar profiles to WT TtFBPA (1.8 mg/mL).
Figure 3.6: SEC-MALLS chromatogram of WT TtFBPA at 1.8 mg/ml. The MALLS chromatogram for WT TtFBPA shows both the differential refractive (black) and molecular mass (red) against elution volume. The average molecular weight for the first peak eluting from the column at 8.2 mL is 28 MDa, representing a protein aggregation product. The second peak eluting from the column at 15 mL has an average molecular weight of 280 kDa, the weight of the di-pentamer complex. Run using a 5 mL Superdex 200 column.

Figure 3.7: SEC-MALLS chromatogram of WT TtFBPA at 10 mg/ml. The MALLS chromatogram for WT TtFBPA shows both the differential refractive (black) and molecular mass (red) against elution volume. The first peak eluting from the column at 12.5 mL has an average molecular weight of 280 kDa, the weight of the di-pentamer complex. The second peak eluting from the column at 16 mL has an average molecular weight 200 kDa. Run using a 5 mL Superdex 200 column.
Figure 3.8: SEC-MALLS chromatogram of W144L TtFBPA at 1.8 mg/ml. The MALLS chromatogram for W144L TtFBPA shows both the differential refractive (black) and molecular mass (red) against elution volume. The average molecular weight for the first peak eluting from the column at 8.2 mL is 28 MDa, representing a protein aggregation product. The second peak eluting from the column at 15 mL has an average molecular weight of approximately 280 kDa, indicating a di-pentamer. Run using a 5 mL Superdex 200 column.

Figure 3.9: SEC-MALLS chromatogram of W144Y TtFBPA at 1.8 mg/ml. The MALLS chromatogram for W144L TtFBPA shows both the differential refractive (black) and molecular mass (red) against elution volume. The average molecular weight for the first peak eluting from the column at 8.2 mL is 28 MDa, representing a protein aggregation product. The second peak eluting from the column at 15 mL has an average molecular weight of approximately 280 kDa, indicating a di-pentamer. Run using a 5 mL Superdex 200 column.
3.4.1. MALLS Analysis of ScFruA

MALLS analysis of ScFruA was performed to determine whether the protein sample was monodisperse and verify the oligomeric state of ScFruA. As with the TtFBPA experiments MALLS was performed on solutions at concentrations of 1.8 mg/ml, additionally, both the N- and C-His$_6$ScFruA proteins were analysed to determine whether the different tags affected folding and dimerization, the results are shown in Figure 3.9 and Figure 3.10.

MALLS chromatograms for N- and C-His$_6$ScFruA at 1.8 mg/ml show a single peak eluting from the column between 14.7 and 18.2 mL for both tagged versions. These correspond to proteins eluting with an average molecular weight of approximately 30 kDa, which can be correlated to the molecular weight of the ScFruA monomer. A minor discrepancy in molecular mass between the N- and C-terminally His$_6$ tagged proteins is present, this could either be due to the inherent inaccuracy in MALLS, although the difference could be caused by the linker regions between the protein construct and His$_6$-tag (N-His$_6$ScFruA: linker SSGLVPRGS, 859 Da; C-His$_6$ScFruA: linker LE, 260 Da). The absence of multiple peaks indicates that these samples are mono-dispersed and monomeric as the molecular weight of the enzyme corresponds to the predicted protein size.
Figure 3.9: SEC-MALLS chromatogram of N-His₆ScFruA at 1.8 mg/ml.
MALLS chromatogram for N-His₆ScFruA of differential refractive index (black) and molecular mass (red) against elution volume. The average molecular weight for the peak eluting from the column at 16.5 mL is 34 kDa, which is the predicted molecular mass of the monomer unit. Run using a 5 mL Superdex 200 column.

Figure 3.10: SEC-MALLS Chromatogram of C-His₆ScFruA at 1.8 mg/ml.
MALLS chromatogram for C-His₆ScFruA of differential refractive index (black) and molecular mass (red) against elution volume. The average molecular weight for the peak eluting from the column at 16.2 mL is 34 kDa, which is the predicted molecular mass of the monomer unit. Run using a 5 mL Superdex 200 column.
3.5. Activity Assays of the Recombinantly Overexpressed FBP Aldolases

Using a well-established enzyme coupled assay for determining retro-aldol activity (Fondy, Levin, Sollohub, & Ross, 1968), the retro-aldol activity of the TtFBPA and the ScFruA enzymes was assessed. Care was taken to ensure GADH and TIM were not rate limiting during these assays.

The enzyme coupled retro-aldol monitors the catalysis of FBP into DHAP and G3P through the oxidation of NADH to NAD$^+$ through the activity of a coupled enzyme. The enzyme GADH controls the NADH dependent catalysis of DHAP into Glycerol-3-phosphate, and additional enzyme TIM catalyses the interconversion of G3P into DHAP. These coupled enzymes serve to drive the equilibrium of the enzyme in the retro-aldol direction through the removal of DHAP.

Kinetic characterisation of WT TtFBPA had already been established by Seibers et al. although no characterisation of W144L and W144Y have not been characterised (Siebers et al., 2001). Although the reported $K_m$ was 9 $\mu$M and the $V_{max}$ was 0.23 U/mg, there are no reported errors for these values making interpretation of these results difficult.

WT TtFBPA was therefore re-characterised to determine true kinetics with error values, in addition W144L and W144Y were also characterised in order to determine whether destabilising and enlarging the active site pocket has an effect on reaction kinetics.

No kinetic characterisation of ScFruA has been published and therefore will be conducted to determine the enzymes activity.
3.5.1. Reaction Kinetics of TtFBPA Proteins

In order to determine the catalytic activity of the aldolase enzymes, their kinetic properties were ascertained. Under specific conditions, enzymes obey the Michaelis-Menten model, which was used to determine the various kinetic parameters for each enzyme. Michaelis-Menten models are used to determine the maximum rate ($V_{\text{max}}$, calculated in U/mg) and the Michaelis-Menten constant ($K_m$, calculated in $\mu$M), the substrate concentration needed to give a rate that equals half $V_{\text{max}}$.

Initial experimentation to determine a suitable TtFBPA input concentration concluded that 50 $\mu$g of enzyme yielded suitable retro-aldol activity at an easily quantifiable rate. Additionally, a coupling-enzyme titration study was performed to determine input concentrations of GADH and TIM to ensure their presence in excess of the reaction. This was to ensure the presence of the coupling enzymes would not be rate limiting for the reaction.

The retro-aldol steady-state reaction kinetics for WT TtFBPA and mutants are described in Table 3.1. The catalytically inactive mutant (K177A), as expected, displays no enzymatic activity.

Comparison of the reaction rates for the different TtFBPA enzymes highlights that WT displays the highest maximal reaction rate at 0.23 U/mg compared to 0.15 and 0.11 U/mg for W144Y and W144L respectively. However, WT TtFBPA also exhibits the highest $K_m$ of all the mutants at 6.0 $\mu$M, a two-fold increase over the next mutant W144Y at 3.0 $\mu$M. These results indicate that the mutations introduced are altering substrate-binding affinities at the expense of activity (Table 3.1).
Figure 3.11: Steady state reaction kinetic data for TtFBPA proteins using FBP. The graph shows the initial rate of reaction for 50 μg TtFBPA proteins against FBP concentration. The reactions were conducted at 50 °C with 0.375 U GADH and 1 U TIM and 50 mM NADH. Michaelis-Menten kinetic analysis determined the $K_m$ and $V_{max}$ of these constructs, data shown in Table 3.1. Graph created in GraphPad Prism 7.0.
Table 3.1. Michaelis-Menten kinetics for retro-aldol activity of TtFBPA and ScFruA proteins

<table>
<thead>
<tr>
<th></th>
<th>$K_m$ (μM)$^a$</th>
<th>$V_{max}$ (U/mg)$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TtFBPA WT</td>
<td>6.0 ± 0.7</td>
<td>0.23 ± 0.01</td>
</tr>
<tr>
<td>TtFBPA W144L</td>
<td>1.9 ± 0.4</td>
<td>0.11 ± 0.01</td>
</tr>
<tr>
<td>TtFBPA W144Y</td>
<td>3.0 ± 0.6</td>
<td>0.15 ± 0.01</td>
</tr>
<tr>
<td>TtFBPA K177A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ScFruA WT</td>
<td>146 ± 14</td>
<td>106 ± 2</td>
</tr>
</tbody>
</table>

3.5.2. Steady State Retro-Aldol Activity of ScFruA

Initial experiments determining a suitable concentration of ScFruA were conducted and showed that 5 μg of enzyme provided a rate of activity that could be easily monitored with the available setup. Michaelis-Menten kinetic values (Table 3.1) showed that ScFruA has a much higher rate of reaction ($V_{max} = 106$ U/mg) compared to WT TtFBPA ($V_{max} = 0.23$ U/mg) while the binding affinity for WT TtFBPA was much higher ($K_m$: ScFruA = 146 μM, TtFBPA WT = 6 μM).
3.6. Determination of the Enzymatic Aldol Reaction

Although the enzyme-coupled reaction utilized above is an extremely useful tool for determining aldolase kinetic properties, it is limited by the fact it can only be used to assess the properties of the retro-aldol reaction (the degradation of FBP). In order to determine the kinetic properties of the aldol reaction (synthesis of FBP) another methodology needed to be developed. Ideally, this reaction should be simple to perform and have the capability to be used as part of a high throughput screening methodology. With these requirements, it was determined that spectrophotometric approaches would be the most suitable.

Figure 3.13: Steady state reaction kinetic data for ScFruA using FBP. The graph shows the initial rate of reaction for 5 µg ScFruA against FBP concentration. The reactions were conducted at 50 °C with 0.375 U GADH and 1 U TIM and 50 mM NADH. Michaelis-Menten kinetic analysis is shown in Table 3.1, graph created in GraphPad Prism 6.0.
Several methodologies were developed using different chemical components, which react with aldehydes and ketones such as the compounds 2,4-dinitrophenyl hydrazine and PurpAld™ (Results not shown). Unfortunately, although these compounds reacted with aldehydes, the results were not reproducible and therefore could not be used to quantify kinetic trends. The conclusion for these studies was that a spectrophotometric method of determining aldol activity with sufficient reproducibility could not be developed and that other methodologies should be explored.

3.7. Discussion

Both wild type \textit{T. tenax} and \textit{S. carnosus} FBP aldolases, in addition to three mutant \textit{T. tenax} constructs, were successfully cloned into expression vectors and transformed into \textit{E. coli} cells. The expression of these recombinant proteins was optimized and a purification strategy was developed for each protein. In addition, MALLS analysis was performed on the active constructs, which confirmed the TtFBPA is dodecameric, while the ScFruA protein is monomeric. The activity of the purified proteins was then assessed using a well-characterized enzyme coupled retro-aldol assay.

The steady state reaction kinetics for the TtFBPA constructs were performed to determine whether the purified enzymes were active. Previous work published for the WT TtFBPA reported a $K_m$ of 9 $\mu$M and a $V_{\text{max}}$ of 0.23 U/mg (Siebers et al., 2001). However, no error rates about these results were reported making a full comparison difficult.

The calculated $K_m$ of WT TtFBPA in this study was 6.0 ± 0.7 $\mu$M, which is similar, but ultimately lower, than that reported by Siebers \textit{et al}. The differences between reported $K_m$ values could arise if the standard errors of the two $K_m$ values reported overlapped. However, this hypothesis is difficult to confirm without the standard error from the
published data. Another potential option could be the differential reaction conditions used between the two studies, although the temperatures and coupling enzyme concentrations were identical, the buffering conditions were different (100 mM Tris, pH 7.0 in Siebers and 50 mM HEPES, pH 7.0 in this study). This could potentially lead to different pH values during the reaction due to the differing buffering capacities between Tris and HEPES at high temperatures. Interestingly however, the calculated $V_{\text{max}}$ values of TtFBPA obtained in this study and published by Seibers were both 0.23 U/mg.

The kinetic parameters reported for Rabbit Muscle Aldolase (RAMA), the most well studied FBP aldolase, displayed a $V_{\text{max}}$ of $20.8 \pm 0.5$ U/mg at 37 °C, a value 100-fold higher than that observed for TtFBPA and a $K_m$ of $14.3 \pm 0.7 \mu$M, a two-fold increase over TtFBPA (Morris & Tolan, 1994).
Chapter Four

Results

4. CRYSTAL STRUCTURE OF TWO FBP ALDOLASES

The aim of the work described here is to use X-ray crystallography of TtFBPA as a tool to study putative reactions with the Baylis-Hillman substrate HBOP. Seibers et al. originally solved the crystal structure of the ligand free TtFBPA, and Lorentzen et al. performed ligand trapping experiments to show both Schiff base formation with DHAP and carbinolamide bond formation with FBP. The aim was to use soaking and or co-crystallisation with HBOP to study HBOP complex formation, and if successful, the Baylis-Hillmann C-C bond formation reaction itself. A distinct, thermostable aldolase from S. carnosus was also studied, ScFruA, for which no crystal structure is available yet. While our results reveal TtFBPA does not bind HBOP in a manner compatible with the Baylis-Hillmann reaction, we reveal an unusual sugar is trapped in the active site of the WT enzyme when incubated with DHAP. This suggests TtFBPA does not exert strict stereochemical control over the C-C bond formation. We also report on the structure of ScFruA.

4.1. Crystallization of WT TtFBPA

Previously, Lorentzen et al. reported the TtFBPA crystal structure. The crystallization conditions reported were used as a starting point for TtFBPA crystallization screens using sitting drop diffusion (Lorentzen et al., 2003). Efficient nucleation and crystal growth for the ligand free enzyme was obtained at 7 % and 9 % poly-ethylene glycol (PEG) 4000, 0.1
M sodium acetate buffer (pH 5.0) and 12 % glycerol. The crystals formed at 7 % PEG 4000 tended to be larger and more elongated as compared to the 9 % PEG 4000 crystals (Figure 4.1) However, when the enzyme was co-crystallized with DHAP, successful crystal nucleation required additives (2-5% sucrose and ethylene glycol) to aid crystallogenesis. Most crystals obtained were small, hampering data collection to sufficiently high resolutions. Hence, a screen was designed with incremental increase in the concentration of various compounds (6% and 10 % PEG 4000, 0.1 M sodium acetate buffer (pH 5.0) and 12 % glycerol). Multiple conditions from this screen led to high quality crystals suitable for X-ray diffraction. Crystals deemed to be of high quality (i.e. large in size, uniform in shape) were mounted in nylon loops and flash-cooled in liquid nitrogen prior to data collection; additional cryoprotection was not necessary as 12% glycerol is present in the mother liquor.

Figure 4.1: Images of X and Y shaped WT TtFBPA crystals. The X shaped crystals (A) were grown in 7% PEG 4000 while Y shaped crystals (B) were grown at 9% PEG 4000.
4.1.1. Structure Determination of WT TtFBPA

Flash-cooled crystals were sent to the Diamond Light Source (Oxford, UK), the X-ray collection data from all data sets was reduced and scaled using X-ray Detector Software (XDS) (Performed by Dr Colin Levy). The best data set collected for the WT TtFBPA apoprotein yielded information to 1.64 Å, 0.24 Å lower than the published TtFBPA (1OJX). The structure was solved by molecular replacement using the 1OJX structure. The initial model was completed through iterative cycles of manual model building and real-space refinement (Coot version 0.8.1; CCP4) followed by crystallographic refinement (Refmac5; CCP4) (Winn et al., 2011). The final model for the monoclinic crystals of WT TtFBPA (Figure 4.3) achieved a final R<sub>work</sub>/R<sub>free</sub> of 0.143/0.180, and consists of ten proteins (253 residues per molecule) and 2520 water molecules per asymmetric unit. Data processing and final refinement statistics are presented in Table 4.1.

The TtFBPA crystal structure shows a putative decamer within the asymmetric unit. MALLS experimentation has shown that the protein also exists in solution as a decamer with an approximate molecular weight of 280 kDa (Figure 3.6). The decamer is constructed from two stacked pentamers, and each monomer forms a TIM barrel structure, typically seen with type 1 aldolases (Figure 4.2).

The active site WT TtFBPA structure shows little difference with the published structure of 1OJX (Figure 4.3) with a root mean squared deviation (RMSD) of 0.37 Å, indicating minimal structural differences (Pymol Molecular Graphics System, Version 1.7.4.4, Schrödinger). However, two phosphate groups can clearly be observed in the active site electron density and were included in the model (Figure 4.3) these were not modelled in the 1OJX structure.
Table 4.1 X-ray crystallographic statistics for data collection, processing and refinement of TtFBPA structures in this study

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>WT Monoclinic</th>
<th>W144L Monoclinic</th>
<th>WT DHAP Co-crystal</th>
<th>W144L DHAP Co-crystal</th>
<th>WT HBOP Co-crystal</th>
<th>W144L HBOP Co-crystal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ligand</td>
<td>-</td>
<td>-</td>
<td>DHAP</td>
<td>DHAP</td>
<td>HBOP</td>
<td>HBOP</td>
</tr>
<tr>
<td>Space group</td>
<td>P 1 2 1</td>
<td>P 1 2 1</td>
<td>P 1 2 1</td>
<td>P 1 2 1</td>
<td>P 1 2 1</td>
<td>P 1 2 1</td>
</tr>
<tr>
<td>Cell Dimensions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>82.85, 158.73, 01.78</td>
<td>83.00, 158.93, 101.94</td>
<td>82.91, 158.48, 101.94</td>
<td>82.86, 158.60, 101.77</td>
<td>83.17, 159.00, 101.88</td>
<td>82.86, 158.60, 101.77</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.00, 108.0, 90.00</td>
<td>90.00, 108.0, 90.00</td>
<td>90.00, 108.0, 90.00</td>
<td>90.00, 108.0, 90.00</td>
<td>90.00, 108.13, 90.00</td>
<td>90.00, 108.01, 90.00</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>41.36 (1.64-1.68)</td>
<td>41.42 (1.71-1.75)</td>
<td>42.85 (1.67-1.73)</td>
<td>55.90 (1.61-1.652)</td>
<td>56.03 (2.03-2.083)</td>
<td>55.90 (1.61-1.65)</td>
</tr>
<tr>
<td>R_{meas}</td>
<td>0.051 (0.564)</td>
<td>0.074 (0.621)</td>
<td>0.062 (0.765)</td>
<td>0.053 (0.498)</td>
<td>0.116 (0.72)</td>
<td>0.053 (0.498)</td>
</tr>
<tr>
<td>I/σI</td>
<td>14.2 (2)</td>
<td>11.4 (1.8)</td>
<td>13.2 (1.9)</td>
<td>12.1 (1.8)</td>
<td>12 (1.8)</td>
<td>12.1 (1.8)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.96</td>
<td>99.71</td>
<td>99.26</td>
<td>99.48</td>
<td>98.37</td>
<td>99.72</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>1.64</td>
<td>1.71</td>
<td>1.67</td>
<td>1.61</td>
<td>2.03</td>
<td>1.61</td>
</tr>
<tr>
<td>No. reflections</td>
<td>28666</td>
<td>255861</td>
<td>271974</td>
<td>304500</td>
<td>151149</td>
<td>304500</td>
</tr>
<tr>
<td>R_{work}</td>
<td>0.140</td>
<td>0.210</td>
<td>0.175</td>
<td>0.146</td>
<td>0.150</td>
<td>0.156</td>
</tr>
<tr>
<td>R_{Free}</td>
<td>0.180</td>
<td>0.243</td>
<td>0.205</td>
<td>0.180</td>
<td>0.202</td>
<td>0.185</td>
</tr>
<tr>
<td>RMS deviations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond Lengths (Å)</td>
<td>0.022</td>
<td>0.019</td>
<td>0.022</td>
<td>0.024</td>
<td>0.018</td>
<td>0.024</td>
</tr>
<tr>
<td>Bond Angles (°)</td>
<td>2.098</td>
<td>1.975</td>
<td>2.176</td>
<td>2.318</td>
<td>1.858</td>
<td>2.275</td>
</tr>
<tr>
<td>Ligand</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean B value (Å²)</td>
<td>-</td>
<td>-</td>
<td>20.002</td>
<td>28.955</td>
<td>30.296</td>
<td>28.266</td>
</tr>
</tbody>
</table>
Figure 4.2: TtFBPA crystal structure reveals a decameric oligomerisation. Ribbon and surface diagram of the WT TtFBPA crystal structure reveals two stacked ring structures (A) each formed from five monomers (B). Each monomer is coloured separately in the ribbon diagram, solvent accessible surface of the structure is shown as transparent grey. The view in B is rotated forward 90° from A. Images made in PyMol version 1.7.4.4
As we sought to determine crystal structures of TtFBPA with various ligands or during turnover, several approaches were used in an attempt to remove the bound phosphates. These included multiple rounds of dialysis and/or treatment with magnesium chloride of the protein solution prior to crystallogenesis. Unfortunately, none of these were successful. To explore whether a WT TtFBPA without bound phosphate could be obtained, we attempted to use distinctly shaped crystals that might have different crystal packing. However, these crystals did not diffract to a suitable level.

Figure 4.3: The active site pocket of WT TtFBPA with free phosphate bound. A close up of the active site WT TtFBP monomer with the corresponding 2Fo-Fc density for free phosphate contoured at 1σ (black). Images made in Pymol version 1.7.4.4.
4.1.2. Structure Determination of TtFBPA W144L

Crystals for W144L TtFBPA were obtained in similar conditions as the WT enzyme, and a similar resolution was achieved (1.71 Å). The initial structure was solved using the monoclinic WT TtFBPA structure; the model was completed through iterative manual model building and real-space refinement (Figure 4.4). The W144L active site pocket is slightly larger compared to WT TtFBPA due to the W to L mutation. This is likely to alter the binding affinity and correlates with the altered $K_m$ and $V_{max}$ values for W144L when compared to WT (Figure 3.11). The increase in size and flexibility could lead to the W144L protein accommodating larger or chemically different substrates. In contrast to the WT structure, only a single free phosphate ion is present in the active site of W144L TtFBPA, supporting the notion that the removal of W177 affects the substrate binding affinity.
4.1.3. Co-crystallization of WT TtFBPA with DHAP

During co-crystallization experiments, over 50 mM DHAP was added, resulting in the formation of crystals typically at lower precipitant concentrations compared to the ligand free protein. Crystals were mounted in nylon loops and flash-cooled in liquid nitrogen prior to data collection, no additional cryoprotectant was added as the mother liquor contains 12% glycerol. Flash-cooled crystals were sent to the Diamond Light Source (Oxford, UK). Data were reduced and scaled using XDS (Performed by Dr Colin Levy). The best data set obtained included reflections up to 1.67 Å resolution.

The high resolution WT TtFBPA structure was used as an initial model, followed by iterative cycles of model building and refinement. Inspection of the initial electron density maps reveals additional electron density within the active site of the enzyme (Figure 4.5).
It is immediately clear this corresponds to a sugar-like molecule containing two phosphate
groups. It would appear that an FBP or FBP-like product is formed in the presence of
DHAP. This is likely due to the fact DHAP, over time, isomerizes to G3P in solution,
providing both substrates required for FBP formation. When FBP is modelled into the
electron density, both the phosphate groups and the sugar ring correspond well to the
density. However, the density around the FBP C4 does not match to the position of the
FBP C4-OH group (Figure 4.5b). In order for the C4-OH to correspond to the electron
density, the C4 stereocentre requires inversion, leading to the diastereomer of FBP,
tagatose bisphosphate (TBP) (Figure 4.5a). TBP being bound in the active site accounts for
the observed electron density (Figure 4.5).
Figure 4.5: Active site of WT TtFBPA co-crystallised with DHAP. The structure of WT TtFBPA (blue) showing the binding of tagatose-1,6-bisphoshate (green), with FBP overlayed (grey) with corresponding Fo-Fc density of the ligand contoured to 3σ (black) (1), in addition, a close up of the C4-OH of TBP (green) with corresponding FBP C4-OH (2). The electron density for the ligand in all 10 subunits (3, A-J). Images created using PyMol version 1.7.4.4.
The TBP ligand modelled into the active site of WT TtFBPA shows the C4-OH group matches the electron density well (Figure 4.5b). This suggests that either TBP is formed over time from DHAP and bound by the enzyme. Additionally, the Aldolase enzyme itself may be responsible for catalyzing the formation of TBP, although further experimentation is required to support this hypothesis. Figure 4.6 shows an excerpt from the class I aldolase mechanism which can result in the formation of either FBP or TBP dependent upon the positioning of the DHAP hydroxyl. Comparison with Figure 4.5B shows that hydrogen bonding of the TBP C4 hydroxyl is stabilised to W144, while FBP C4 hydroxyl would not be stabilised by any residues. This indicates that WT may be able to bi-functionally form FBP or TBP, while W144 mutants may not be able to form TBP.

Figure 4.6: An Excerpt from the Class I Aldolase mechanism showing the formation of stereogenic centre

TBP has not previously been reported as a TtFBPA product and therefore further experimentation will be conducted with the aim of determining whether TBP is indeed the product formed. The crystal structure reveals the TBP molecule is bound to the Lys-177 via a carbinolamine bond, suggesting it corresponds to the aldol reaction product trapped in the active site pocket prior to product release. This could reflect that, at lower
temperatures, the rate-limiting step of this thermophilic enzyme is product release, with TBP release significantly slower than the corresponding FBP product, leading to TBP accumulation in the crystalline sample over time, although further experimentation will be conducted.

4.1.4. Co-crystallization of TtFBPA W144L with DHAP

Co-crystallization experiments using DHAP (over 50mM) were repeated with TtFBPA W144L using similar conditions and crystal handling methods as those for W144L ligand free crystals. Data was collected at the Diamond Light Source (Oxford, UK).

Comparison between the ligand free W144L structure (i.e. containing phosphate) and the co-crystallized W144L:DHAP complex reveals additional electron density that connects the bound phosphate and the Lys-177 amine group (Figure 4.7). Modelling DHAP into the electron density suggests a Schiff base is formed between Lys-177 and DHAP. However, the electron density does not perfectly match DHAP, suggesting at least some level of disorder, multiple species in the active site and/or lower occupancy of the DHAP ligand.

The marked difference observed in the bound active site species for W144L as opposed to the WT TtFBPA structure (i.e. DHAP versus TBP) suggests that the W144L TtFBPA is unable to form sufficient amount of aldol condensation products during the co-crystallization experiment. A result which corresponds to the enzymatic kinetics shown in Figure 3.11. Furthermore, as hypothesised the absence of W144 may have prevent the formation of TBP.
In order to determine if the Baylis-Hillman substrate (HBOP) can bind to the TtFPBA active site, HBOP soaking experiments were carried out using both WT and W144L TtFPBA crystals. Unfortunately, no HBOP was observed in electron densities obtained from soaked crystals. Co-crystallization experiments with HBOP were carried out, and, due to precipitation of the enzyme upon HBOP addition, required the enzyme to be diluted in buffer containing 500 μM HBOP. The TtFPBA:HBOP solution was slowly concentrated to 10 mg/ml TtFPBA. This allowed crystals to be obtained using similar conditions as those for the TtFPBA ligand free crystals. Data sets to 2.03 and 1.61 Å were collected at Diamond light source for the WT and W144L TtFPBA HBOP co-crystals respectively. An
initial model based on the WT TtFBPA DHAP co-crystal structure was used as a starting model for the refinement of WT TtFBPA:HBOP co-crystals (Figure 4.8).

Figure 4.8 shows the in complex enzyme substrate complex with HBOP, which appears bound to the Lys-177 via a Schiff base linkage with the C of the HBOP carbonyl group. Additionally, the F_o-F_c map (Figure 4.8, Section 1) shows the electron density of the typical subunit within the TtFBPA decamer. A structure alignment of each subunit is displayed in Figure 4.8 section 2 which shows the variations in each subunit in the decamer, subunit J is highlighted in green due to a large structural change relative to subunits A-I. The electron density for each subunit is displayed in Figure 4.8 section 3.

Although in general the electron density of HBOP is highly ordered, the alkene moiety of HBOP appears more disordered. This may be further highlighted in subunit J, although a full HBOP molecule is modelled into the electron density, another possible outcome could be that the HBOP alkene moiety is not ordered and therefore a definable electron density pattern cannot be determined. Furthermore, the electron density connecting the HBOP carbonyl with Lys-177 is weakened in multiple subunits too, also suggesting incomplete occupancy within the crystal.

Although HBOP can successfully be resolved in the active site of the WT enzyme, the binding of HBOP in the active site via a Schiff base suggests that a Baylis-Hillman reaction with HBOP is unlikely to be catalyzed by the WT enzyme due to fundamental differences in catalyst binding between the aldol and Baylis-Hillman chemical mechanisms.
Figure 4.8: Co-crystallization of WT TtFBPA with HBOP. 1) The structure of WT TtFBPA (blue) showing the binding of HBOP (magenta) with corresponding Fo-Fc density of the ligand contoured to 3σ (black). 2) An alignment model of all 10 subunits to show structural variations subunits A-I (blue) and subunit J (Green). 3) The structure of each subunit binding HBOP, subunits A-I (blue) and subunit J (Green), with corresponding Fo-Fc density of the ligand contoured to 3σ (black). Images created using PyMol version 1.7.4.4.
A different picture is observed when using W144L TtFBPA for HBOP co-crystallization. The TtFBPA W144L:HBOP crystal structure reveals that all subunits appear to contain HBOP bound in the active site (Figure 4.10). Section 1 shows the F_o-F_c map for a typical subunit bound HBOP while Section displays each individual subunit. Alignment of the subunits was now shown as no major structural deviations were observed.

In the W144L active site, HBOP appears to be covalently bound to Lys-177 in all subunits (Figure 4.10, Section 2). However, electron density suggests the covalent bond between HBOP and Lys-177 is not a Schiff base between the carbonyl carbon of HBOP but rather an amine bond formed between Lys-177 and the HBOP alkene group. This would suggest amine addition to the alkene group concomitant with formation of an enol intermediate. This results in the same intermediate as observed with DABCO binding in the chemical Baylis-Hillman reaction (Scheme 1.2). The electron density between the C1 and C2 of HBOP is reduced compared to the rest of the density; this is likely to result from a large level of fluidity around this bond.

Figure 4.9: The proposed chemical structures of HBOP bound to Lys-177 via a Schiff base and Amine bond
4.2. Crystal Structure Determination of ScFruA

As a starting point for the production of ScFruA crystals, several commercial crystallization screens were used. Both C- and N-His$_6$ScFruA were screened for crystallization, at both 6 and 10 mg/ml concentration. Showers of crystals were observed in the C-His$_6$ScFruA at 100 mM sodium acetate buffer (pH 5.5), between 1-11 % PEG 2K MME and between 150-300 mM lithium sulphate. No crystals formed for N-His$_6$ScFruA.

Figure 4.10: Co-crystallization of W144L TtFBPA with HBOP. 1) The structure of W144L TtFBPA (green) showing the binding of HBOP (magenta) with corresponding Fo-Fc density of the ligand contoured to 3σ (black). 2) The structure of each subunit binding HBOP, subunits A-J (Green), with corresponding Fo-Fc density of the ligand contoured to 3σ (black). Images created using PyMol version 1.7.4.4
The small C-His$_6$ScFruA crystals were used as a seed stock for further rounds of microseeded crystal trials using the same screen. The majority of crystals obtained were fragile and plate-like. Some more robust crystals were identified during the screening (Figure 4.11).

![Figure 4.11: Images of C-His$_6$ScFruA crystals. Crystals were grown between 3-6\% PEG 2000 MME.](image)

Crystals were flash frozen in liquid nitrogen following addition of 12\% PEG 200. A complete diffraction data set was obtained from a single crystal at the Diamond Lightsource facility (Oxford, UK). Data were reduced and scaled using XDS (Performed by Dr Colin Levy). The crystal belongs to space group P2$_1$ and diffracted to 2.17 Å (Table 4.2). The structure was solved using molecular replacement using PhaserMR (McCoy et al., 2007), and the *P. gingivalis* FBP aldolase crystal structure as a search model (PDB code 2IQT, 55\% sequence identity). Following molecular replacement, an initial automated model-building round was performed using Buccaneer (Cowtan, 2006). The auto-building process was completed using ARP/wARP (Morris, Perrakis, & Lamzin, 2003). This model was further refined through iterative cycles of manual model building and refinement using Refmac5 (Willis & Pryor, 2003). Final data collection and refinement statistics are given in Table 4.2. The final ScFruA structure compared to 2IQT is shown in Figure 4.12. The ScFruA asymmetric unit shows a protein tetramer, similar to other FBP
aldolases such as RAMA. However, MALLS (Section 3.4) data showed that ScFruA is a monomeric 34 kDa protein.

Table 4.2: X-ray crystallographic statistics for data collection, processing and refinement of ScFruA structures in this study

<table>
<thead>
<tr>
<th></th>
<th>ScFruA</th>
<th>ScFruA:DHAP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ligand</strong></td>
<td>-</td>
<td>DHAP</td>
</tr>
<tr>
<td><strong>Space group</strong></td>
<td>P 1 2, 1</td>
<td>P 1 2, 1</td>
</tr>
<tr>
<td><strong>Cell Dimensions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>76.12, 96.72, 84.31</td>
<td>76.43, 96.95, 84.24</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.00, 98.18, 90.00</td>
<td>90.00, 98.00, 90.00</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>19.85-2.17</td>
<td>32.79-1.95</td>
</tr>
<tr>
<td></td>
<td>(2.22-2.17)</td>
<td>(2.00-1.95)</td>
</tr>
<tr>
<td>R_pim</td>
<td>0.112 (0.637)</td>
<td>0.084 (0.579)</td>
</tr>
<tr>
<td>I/σl</td>
<td>8.4 (1.8)</td>
<td>11.4 (2.1)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.46</td>
<td>99.60</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>2.17</td>
<td>1.95</td>
</tr>
<tr>
<td>No. reflections</td>
<td>59912</td>
<td>83916</td>
</tr>
<tr>
<td>Mean B value (Å²)</td>
<td>28.48</td>
<td>25.903</td>
</tr>
<tr>
<td>R_work</td>
<td>0.200</td>
<td>0.209</td>
</tr>
<tr>
<td>R_Free</td>
<td>0.271</td>
<td>0.249</td>
</tr>
<tr>
<td><strong>RMS deviations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond Lengths (Å)</td>
<td>0.014</td>
<td>0.018</td>
</tr>
<tr>
<td>Bond Angles (°)</td>
<td>1.612</td>
<td>1.901</td>
</tr>
<tr>
<td><strong>Ligand</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean B value (Å²)</td>
<td>-</td>
<td>35.481</td>
</tr>
</tbody>
</table>

Comparison of the carbon backbone between ScFruA and *P. gingivalis* FBP aldolase (2IQT) reveals that the majority of the α-helix and β-sheet structures are conserved. Variation between both structures emerges within loop regions connecting these secondary structures, such as an elongated loop on ScFruA (highlighted with an * Figure 4.12b). Comparison between the active site of ScFruA and 2IQT reveals very similar active sites,
with only slight positional changes in the active site residues. The active site of ligand free ScFruA is shown in Figure 4.13.

Figure 4.12: Structure of the ScFruA apo structure compared to the 2IQT homologous structure. The entire ScFruA asymmetric unit (A) and the comparison of the ScFruA peptide backbone (yellow) against the 2IQT homologous structure (grey). PyMol version 1.7.4.4.
Following crystallization of ScFruA, both ligand soaking and co-crystallization experiments were attempted with the DHAP substrate. Unfortunately co-crystallization experiments did not yield any suitable crystals, but ligand-soaking experiments with DHAP were successful. Protein crystals were soaked with DHAP (at concentrations above 50 mM) for 30 seconds. The electron density (Figure 4.14) for the ScFruA:DHAP active site reveals additional electron density now extending from the bound phosphate group towards the catalytic Lys-212 residue. This can be modelled as DHAP bound to Lys-212, although density for the C3-OH group is poor. This either suggests multiple conformation and/or species in the active site, or a lower occupancy of this hydroxyl group of DHAP. A similar observation was made for the W144L TtFBPA:DHAP complex.
4.3. Discussion and Conclusion

HBOP co-crystallization experiments with both WT TtFBPA and the W144L variant reveal some binding of HBOP within the active site. In the case of WT TtFBPA:HBOP crystals, a subset of subunits contain what appears to be a Schiff base adduct to Lys177, a conformation incompatible with Baylis-Hillman activity. By contrast, the HBOP:W144L complex contains distinct electron density that appears to show a covalent linkage formed between the HBOP C=C group and Lys-177. This can be interpreted as an amine linkage to an enol tautomer of HBOP. However, the electron density does indicate a significant degree of disorder or lower levels of occupancy and this interpretation must be viewed cautiously, as the ligand B factors are 30.296 in WT and 28.266 in W144L compared to overall B factors of 25.184 and 20.018 respectively. Conversely however, although a

Figure 4.14: The active site pocket of ScFruA containing DHAP. A close up of the ScFruA active site monomer (yellow) containing DHAP bound to Lys-212 (magenta) with the corresponding Fo-Fc density for DHAP contoured at 3σ (black). Images made in Pymol version 1.7.4.4.
positive enzyme-HBOP complex is observed with binding via K177 and the alkene moiety of HBOP a subsequent N-alkylation reaction may have occurred, protonating the highly reactive lone pair on lysines R-chain. Further mass spectrometry experimentation needs to be conducted to determine whether this Nitrogen is protonated.

WT TtFBPA co-crystallization and ligand soaking experiments both yielded slightly different results. Co-crystallisation experiments, which involved the slow exposure of the enzyme to DHAP during crystallogenesis showed the product TBP in the active site bound covalently to Lys-177. Modelling of FBP into this electron density gives a negative density around the C4 hydroxyl group which highlights that this group does not occupy that space. When the C4 hydroxyl group of TBP is modelled in however, it fits the density.

ScFruA crystallization experiments yielded an end result of obtaining an apo-crystal structure and the structure bound DHAP. ScFruA is markedly different to TtFBPA, with the exception of the catalytic lysine, there is no structural homology between the two enzymes, this is because ScFruA is a class I aldolase and TtFBPA is a class IA archaeal aldolase. ScFruA ligand soaking experiments were conducted as a potential method to identify any potential products formed, unidentified, novel products were observed in NMR experiment. Unfortunately, no such products were observed in DHAP soaked crystals.

In conclusion, although ligand binding has been observed in TtFBPA variants and in ScFruA, enzymatic activity cannot be determined. In order to determine whether true Aldol or Baylis-Hillman products are formed, multiple NMR techniques will be used to determine product formation.
Chapter Five

Results

5. NMR-BASED DETECTION OF ALDOLASE ACTIVITY

Although suitable spectrophotometric methodologies for determining aldol activity were tested, they did not yield any successful results. Therefore the development of a high throughput method of determining aldol activity was abandoned. Instead, methodologies were investigated that could offer a robust method of determining aldol and Baylis-Hillman enzyme activity. NMR is a technique for determining the composition of solutions. Various types of NMR have been utilised in determining enzyme activity [Szwergold et al. 1995; List et al. 1998], though importantly Berry et al. recently showed $^{31}$P-NMR to be a suitable method for showing aldol activity in both FBP and TBP aldolases [Williams et al. 2003]. This technique was not only used to detect products at the end point of the enzymatic reaction but also monitored product formation and substrate depletion as part of a time-course reaction. Therefore, $^{31}$P-NMR experiments will be performed to assess aldol activity, in addition to determining whether any potential Baylis-Hillman activity is observed in TtFBPA.

Further NMR studies utilising $^1$H/$^{13}$C 2D HSQC NMR and HSQC TOCSY NMR will also be performed. This will aim to provide greater structural detail about the products formed during the reactions, and characterise the stereospecificity of TtFBPA and ScFruA.
5.1. Collection of $^{31}$P-NMR of Reference Spectra

Mass spectrometry reference spectra were obtained for DHAP, G3P and FBP (in D$_2$O) to determine their purity (Appendix Section 8.2). The results showed that contaminants were present in both DHAP and G3P and as such $^{31}$P-NMR spectrum were also collected for these compounds to serve as future reference. Furthermore, reference spectra are important when performing $^{31}$P-NMR as the spectrum is heavily dependent on exact buffer conditions used. Therefore, the best method for product identification is via comparison to known reference spectra collected under near identical conditions. The $^{31}$P-NMR spectrum of FBP (Figure 5.2) shows four distinct peaks – the two major peaks at 3.5 and 3.12 ppm correspond to $\beta$-FBP, the major epimer of FBP, while the two minor peaks at 4.29 and 2.8 ppm correspond to $\alpha$-FBP, the minor epimer in solution, this is known due to the peak assignment by Williams et al. no free phosphate could be detected\cite{Williams2003}. The DHAP $^{31}$P-NMR spectrum shows three main peaks – one at 1.63 ppm that can be attributed to free phosphate and two additional peaks at 3.44 and 4.26 ppm (Figure 5.2, DHAP spectrum). The latter are assigned as the DHAP phosphorus and the gem-diol DHAP form phosphorus respectively (Chemical structures in Figure 5.1). As all peaks observed for the DHAP $^{31}$P-NMR spectra can be accounted for, contaminants detected previously by mass spectrometry are unlikely to be phosphorylated or occur at minute amounts.
Figure 5.1: The inter-conversion of DHAP and Gem Diol DHAP

Figure 5.2: $^{31}$P-NMR of reference compounds used in the aldol and retro-aldol reactions. $^{31}$P NMR spectra for FBP (25 mM; purple), DHAP (25 mM; green), G3P (25 mM; red) and a combination of DHAP and G3P (12.5 mM each, blue).
The $^{31}$P-NMR spectrum obtained for G3P contains two main peaks (Figure 5.2): the major peak observed at 3.85 ppm is allocated to the phosphate group attached to G3P, with the peak observed at 1.65 ppm attributed to free phosphate. No significant signals corresponding to a contamination with FBP or other phosphorylated sugars can be observed.

In addition to collection of reference spectra of DHAP and G3P, a spectrum of sample containing both DHAP and G3P was obtained, to ensure the individual peaks corresponding to both species could be resolved. The $^{31}$P-NMR spectrum for an equimolar G3P and DHAP mixture shows the combination of these two compounds causes a chemical shift in the profiles observed, likely as a result of pH change as decomposition of compounds occurs when compared to the individual compound solution (pH 7.01 in D$_2$O) (Figure 5.2). In this case, the DHAP peaks have changed from 4.26 and 3.44 ppm to 4.06 and 3.29 ppm upon the addition of G3P, while the G3P associated peak has shifted from 3.85 ppm to 3.39 ppm.

5.2. $^{31}$P-NMR of the Retro-Aldol Reaction

5.2.1. $^{31}$P-NMR of the Retro-Aldol Reaction Using TtFBPA

With the individual reference spectra obtained, experiments were performed to monitor the ability of TtFBPA to convert FBP into DHAP and G3P, the retro-aldol reaction, via $^{31}$P-NMR. Multiple spectra were collected for a retro-aldol reaction (WT TtFBPA mixed with FBP) on the same sample over a 1.5 hour period, to determine whether $^{31}$P-NMR is a suitable method of identifying any differences in reaction products and substrates following catalysis with aldolase enzyme (Figure 5.3).
An increased number of scans were collected for the time course reaction sample, resulting in improved peak intensities when compared with the reference compounds. The spectrum for the 30 minute time point of the WT retro-aldol reaction clearly contains four FBP peaks ($\beta$-FBP at 3.38 and 3.14 and $\alpha$-FBP at 4.37 and 2.95). Three additional peaks at 4.24, 3.77 and 3.52 ppm are present. The position of these peaks does not, on initial inspection, correspond to the main peaks observed for the reference spectrum of DHAP and G3P, which occur at 4.06, 3.39, and 3.29 ppm. However, comparison between the free phosphate peaks between the DHAP and G3P control spectrum and the 30 min spectrum show a chemical shift of 0.29 ppm, likely due to pH change. Taking this into consideration, peaks observed at 4.24 ppm and 3.52 ppm can be allocated to DHAP. However, a corresponding G3P associated peak would be expected at approximately 3.15 ppm, which is not present. An Additional peak at 3.77 ppm is also observed which is likely to be G3P. However, the concentration of these products is similar to the minor epimer $\alpha$-FBP species (peaks at 4.37 and 2.95 ppm) suggesting only limited retro-aldol activity. The one hour time point spectrum reveals an additional peak unknown peak at 3.79 ppm. While peaks corresponding to the DHAP and G3P products increase slightly over the time course of this experiment, the most dramatic increase occurs in the levels of free phosphate, suggesting significant product hydrolysis occurs. These results demonstrate that, in principle, retro-aldol activity can be observed using $^{31}$P-NMR. Unfortunately, due to the changing in the peaks within the enzymatic reactions it highlights that the buffer concentration was not strong enough to prevent changes in pH.
5.3. End Point Reaction of Aldol Reaction Using $^{31}$P-NMR Spectroscopy

As $^{31}$P-NMR spectroscopy was able to detect some retro-aldol activity for WT TtFBPA, this technique was used to study whether the aldol condensation of DHAP and G3P into FBP by TtFBPA (WT and variants) can be monitored.

Given the inherent reactivity of DHAP and G3P, a reaction mixture containing just DHAP and G3P was incubated at 50 ºC for 30 minutes, and the $^{31}$P-NMR measured to determine the effect incubation at elevated temperatures on a mixture of these two compounds (Figure 5.4). A large increase in the amount of free phosphate is detected (peak at 1.47 ppm).
ppm) when comparing to the spectrum for DHAP and G3P prior to incubation (peak at 1.23 ppm). Furthermore, the level of G3P has dropped dramatically following incubation (peak at 3.67 ppm, compared to 3.39 ppm).

Therefore, it appears incubation at elevated temperature leads to G3P degradation and an increase in free phosphate levels. This in turn appears to have altered the pH of the reaction, resulting in the observed shift of 0.24 ppm.

When monitoring a DHAP and G3P mixture containing WT TtFBPA (500 μg) in addition, four major peaks are observed at 4.29, 3.71, 3.47 and 1.55 ppm (Figure 5.4) corresponding to DHAP (4.29 and 3.47 ppm, red labels), G3P (3.71 ppm, green labels) and free phosphate species (1.55 ppm, blue labels). In addition to the substrate peaks, the spectrum of the WT TtFBPA reaction mixture contains three additional peaks at 3.36, 3.16 and 3.00 ppm. The peaks at 3.36 and 3.00 ppm could correlate to β-FBP (Black Labels) suggesting catalysis by the WT TtFBPA of the aldol condensation reaction can be monitored. The additional peak at 3.16 ppm is unaccounted for.
The spectrum of a TtFBPA W144L (500 μg) containing reaction mixture shows the same substrate derived peaks as for the WT enzyme (peaks at 4.32, 3.74 and 3.49 ppm), as well as the free phosphate peak at 1.59 ppm. An additional minor peak is observed at 3.05 ppm, possibly indicating formation of β-FBP albeit at low levels. This correlates to kinetics data shown in Section 3.3, showing that W144L has an enlarged K_m relative to WT TtFBPA.

The spectrum observed for a reaction containing TtFBPA W144Y (500 μg) also reveals strong substrate peaks (4.27, 3.67 and 3.46 ppm) and a large amount of free phosphate (1.51 ppm) with only two small peaks at 3.34 and 2.97 ppm corresponding to β-FBP.

Figure 5.4: 31P-NMR spectroscopy of the aldol reaction using TtFBPA. 31P-NMR control spectra for FBP (25 mM; pink), equimolar DHAP:G3P (12.5 mM each, red) and the same equimolar DHAP:G3P sample after a 30 minute incubation at 50 °C (12.5 mM; yellow). TtFBPA enzymes; WT (green), W144L (light blue) and W144Y (dark blue) were incubated with DHAP and G3P for 30 mins at 50 °C before analysis. Peak assignment has been labelled with DHAP (red), G3P (Green), FBP (Black) and Phosphate (Blue).
results show that $^{31}$P-NMR spectroscopy can be used to monitor the WT TtFBPA aldolase activity, with diminished catalytic activity observed for the TtFBPA variants.

5.3.2. Time Course Reaction of Aldol Reaction Using $^{31}$P-NMR Spectroscopy

Given the low levels of aldol condensation products for the TtFBPA detected after 30 minutes incubation at 50 °C, a similar reaction was monitored for a 15 hour time course, monitoring the reaction every hour (Figure 5.5). Given the extended time frame of this experiment, the incubation temperature was dropped to 25 °C (instead of 50 °C) to minimize thermal substrate degradation.

The first spectrum collected following 1 hour of incubation shows peaks at 4.24, 3.43 and 3.62 ppm, corresponding to DHAP and G3P respectively. A number of other peaks are observed which increase over the course of the experiment. The peaks at 2.95 and 3.35 ppm likely correspond to the major β-FBP epimer. The other peaks observed are currently unknown (labeled [1] to [6]). The peaks at 3.65 and 3.3 ppm (labeled [2] and [4]) show a clear increase over the course of the experiment, indicating that there is a second molecule with a similar chemical shift to FBP that accumulates over a much longer time period. This unidentified species could correspond to another phosphorylated sugar, such as tagatose-1,6-bisphosphate (TBP; an epimer of FBP which differs in the chirality at the C4-OH group). Interestingly, these peaks are barely observed within the first 2 hours and appear in higher quantities after this time. These results mimic the crystallographic experiments whereby overnight soaking with DHAP led to the presence of TBP in the active site of TtFBPA. The concentration of both FBP and TBP does not increase from 6 hours onwards, suggesting the reaction may have reached equilibrium, while the DHAP and G3P continue to be depleted with the free phosphate concentration increasing.
5.3.3. Identification of TtFBPA Unknown Products

The time course experiment performed for WT TtFBPA reveals formation of as yet unidentified products, suggesting that the enzyme is catalyzing the formation of products other than FBP. There are several FBP aldolases which have been shown to catalyse the formation of tagtose-1,6-bisphosphate (TBP) and it is possible that TtFBPA is capable of catalysing the formation of TBP [LowKam et al. 2010; Zgiby et al. 2000; Manjasetty et al. 2003]. The activity of *E. coli* tagatose-1,6-bisphosphate aldolase TagA (a gift from the Turner group) was measured using $^{31}$P-NMR and compared to the aldol activity of WT TtFBPA (Figure 5.6).

When the position of the free phosphate peaks is compared with other spectra, a large shift is observed for the TagA reaction mixture spectrum. This suggests that 25 mM HEPES pH 7.0 was ineffective in buffering the reaction. Comparison of the chemical shifts between WT TtFBPA and TagA free phosphate peaks gives a chemical shift of 0.8 ppm while the shift in DHAP peaks between TagA (4.62 and 3.74 ppm) and WT TtFBPA (4.24 and 3.44) is around 0.4 ppm. As such, it is difficult to correlate the reaction products of both mixtures with any certainty. Ideally, a reference spectrum of TBP is required. Unfortunately, this product is not commercially available.
The NMR experiments with TtFBPA demonstrate both aldol and retro-aldol activity can be monitored in principle. Additionally, further hints of a broader product profile than was originally expected were obtained for this enzyme. In light of this, similar experiments were performed with the ScFruA to investigate whether this enzyme is also capable of producing novel products. The $^{31}$P-NMR spectrum of an aldol reaction mixture for ScFruA, is shown in Figure 5.7.

**Figure 5.6: $^{31}$P-NMR of the initial and final time course spectra for the WT TtFBPA time course reaction.** Eqimolar DHAP:G3P (red) and FBP (purple), additionally an aldol condensation of \textit{E.coli} TagA spectrum is shown (blue) which denotes the formation of TBP (4.18 and 4.06 ppm). The WT aldol spectrum after one hour (light green) shows 11 identified peaks and the spectrum after 15 hours the WT aldol spectrum (dark green) shows 12 identified peaks.

**5.4. $^{31}$P-NMR Spectroscopy of ScFruA Aldol Activity**

The NMR experiments with TtFBPA demonstrate both aldol and retro-aldol activity can be monitored in principle. Additionally, further hints of a broader product profile than was originally expected were obtained for this enzyme. In light of this, similar experiments were performed with the ScFruA to investigate whether this enzyme is also capable of producing novel products. The $^{31}$P-NMR spectrum of an aldol reaction mixture for ScFruA, is shown in Figure 5.7.
The ScFruA spectrum contains a series of partially overlapping peaks, three of which can be allocated to DHAP (4.46 and 3.71 ppm) and G3P (3.80 ppm). The height of these peaks (as compared to the reference spectrum) suggests that their concentration has dropped. The peaks at 3.61 and 3.40 ppm likely correspond to β-FBP and the peaks at 4.29 and 3.26 ppm can be allocated to α-FBP. Inorganic phosphate is present at high levels (1.85 ppm). This leaves two peaks unaccounted for (3.89 and 3.5 and ppm). The two peaks at 3.89 and 3.8 ppm could potentially correspond to TBP or another bisphosphorylated sugar. The identity of the additional peak at 3.5 ppm cannot be determined by $^{31}$P-NMR. The presence and stacking of so many peaks makes product identification difficult, therefore, alternative approaches will be used.

Figure 5.7: $^{31}$P-NMR spectroscopy of ScFruA aldol activity. $^{31}$P-NMR reference spectra for equimolar DHAP:G3P (red; 12.5 mM each) and FBP (blue, 25 mM). The ScFruA aldol reaction (green) shows 11 identified peaks.
5.5. $^{31}$P-NMR spectroscopy aimed at Detecting of Baylis-Hillman Activity

5.5.1. Detecting TtFBPA Baylis-Hillman Reaction Activity

The Baylis-Hillman reaction consists of a condensation between an alkene, in this case HBOP, and an aldehyde, such as G3P (Scheme 1.2 and Scheme 1.3). This has some similarity with the aldol condensation reaction, which utilizes the ketone DHAP, and G3P (Figure 5.6). To determine whether TtFBPA can catalyse a Baylis-Hillman-type reaction, $^{31}$P-NMR was performed using HBOP and G3P as the substrates (Figure 5.8). The reaction conditions used in these experiments were identical to those used for the aldol condensation reaction studied (i.e. a 1 hour incubation at 50 °C).

A reference spectrum was obtained for HBOP and shows 3 major singlet peaks at 3.31, 2.96 and 1.34 ppm and a doublet peak at 3.20 and 3.16 ppm. The peak observed at 1.34 ppm corresponds to free phosphate. The presence of 5 total peaks greatly suggests the presence of HBOP isomers (such as germinal diol HBOP), or reaction by-products as a result of reaction with the C=C function group of HBOP and possible contaminants. Unfortunately, characterization of the phosphorus peaks is difficult as $^{31}$P-NMR does not determine any structural information. Therefore, determination of HBOP from any potential contamination or isomers is difficult. Comparison of the spectrum obtained for the TtFBPA WT with HBOP and G3P mixture with reference spectra reveals clearly defined peaks corresponding to HBOP (3.40, 3.28, 3.26 and 3.10 ppm). Furthermore, a small peak appears at 3.71 ppm, which likely corresponds to G3P (Black Label). The height of this peak suggests a high rate of G3P hydrolysis, linked to the increase observed in the free phosphate peak (1.48 ppm). Further peaks are observed at 4.02 and 2.49 ppm, though these are also observed in the HBOP reference spectrum ([1] and [2], red label). A shoulder is observed in the peak at 3.40 ppm in addition to a peak at 3.24 ppm, which is not observed in either the HBOP or G3P reference spectra, and could be a consequence of
the increased resolution with a higher number of scans performed for the enzymatic reaction sample ([3], red label). Hence, it can be concluded that there is no Baylis-Hillman activity under the conditions used.

**Figure 5.8: $^{31}$P-NMR of WT TtFBPA Baylis-Hillman activity.** Reference spectra for FBP (purple; 25 mM), G3P (red, 25mM) and the Baylis-Hillman substrate HBOP (turquoise, 25 mM) were used as positive control reactions. The WT TtFBPA Baylis-Hillman reaction shows 13 identified peaks with 5 major peaks. G3P peak is labeled on both the reference and HBOP & G3P spectra (Black). Additional peaks are labelled [1-3] (Red).

### 5.6. 2D $^1$H/$^{13}$C-NMR HSQC Spectroscopy

The $^{31}$P-NMR spectra for both TtFBPA and ScFruA contained peaks that could not be readily identified. These unassigned peaks could correspond to alternative products; however, $^{31}$P-NMR provides insufficient information to allow for identification of these products. For this reason 2D $^1$H/$^{13}$C-NMR was used to provide insight into the chemical structure of the TtFBPA and ScFruA reaction products.
5.6.1. Reference Spectra

Although there are reference spectra for DHAP and FBP available in both the Human Metabolome Database and the Yeast Metabolome Database, these were repeated to account for possible changes in pH, temperature, buffering conditions and concentration. New reference spectra are shown for FBP (Figure 5.9), DHAP (Figure 5.10) and G3P (Figure 5.11).

The annotated heteronuclear single quantum coherence spectroscopy (HSQC) reference spectrum for FBP contains the FBP structure (Figure 5.9) and reveals both the α and β form of FBP (labeled α1-α6 and β1-β6). The $^{31}$P-NMR spectrum of DHAP (Figure 5.2) contains two distinct peaks indicating the presence of both ketone and gem-diol forms. The four peaks observed on the $^1$H/$^{13}$C-NMR spectrum exhibit very tightly coupled doublet peaks, corresponding to the two protons bound to the C1 and C3 carbons. It is likely that the more downfield cross peaks are the gem-diol DHAP hydrate. The $^1$H/$^{13}$C 2D NMR spectrum for G3P shows three cross peaks labelled C1 to C3, the more downfield peak is labelled C3 due to the carbonyl group which is very electron withdrawing, C1 is labelled due to the doublet peak from C2.
Figure 5.9: $^1$H/$^{13}$C HSQC NMR reference spectrum of FBP. The $^1$H/$^{13}$C NMR spectrum contains 25 mM FBP in D$_2$O. The structure of FBP has been annotated with carbon identity [1-6]. These carbons have been annotated onto the NMR cross peaks. C1-6 cross peaks are identified where possible. The identity of $\alpha$-FBP C3-5 couldn’t be fully identified. C2 is not visible due to the absence of protons.

Figure 5.10: $^1$H/$^{13}$C HSQC NMR reference spectrum of DHAP. The $^1$H/$^{13}$C NMR spectrum contains 25 mM DHAP in D$_2$O. The structure of DHAP has been annotated with carbon identity [1-3]. These carbons have been annotated onto the NMR cross peaks as belonging to DHAP or the gem-diol form of DHAP. C2 is not visible due to the absence of protons.
5.6.2. 2D NMR of TtFBPA reaction mixtures

With the HSQC reference spectra for FBP, DHAP and G3P established, aldol condensation reactions were conducted using TtFBPA. The reaction mixture was analysed by HSQC $^1$H/$^{13}$C 2D NMR to determine carbon cross peaks (Figure 5.11) and HSQC TOCSY $^1$H/$^{13}$C 2D NMR to determine coupled carbon cross peaks (Figure 5.13).

Utilising the reference HSQC spectra for the FBP, DHAP and G3P, cross-peaks which correspond to these molecules are labelled according to their carbon positions. Comparison of the cross peaks of the reference spectra to those observed in the spectrum for the TtFBPA reaction shows the presence of both substrates (Figure 5.11). In addition, there are also cross peaks which can be identified as FBP. However, there are also several unidentified cross peaks in the spectra labelled $x1$-$5$.

![Figure 5.11: $^1$H/$^{13}$C HSQC NMR reference spectrum of G3P.](image)

The $^1$H/$^{13}$C NMR spectrum contains 25 mM G3P in D$_2$O. The structure of FBP has been annotated with carbon identity [1-3]. These carbons have been annotated onto the NMR cross peaks.
Cross peaks labelled $x_{1-3}$ are single peaks indicating that they are CH or CH$_3$ groups, whilst the tight clustering of these peaks indicates that they may form a cyclic sugar group, as seen with FBP. These cross peaks $x_{1-3}$ could be derived from C3, C4 and C5 carbons of TBP or another sugar. As is the case with FBP, the C2 carbon of TBP has no coupled protons and would not be detected. The nature of cross peaks $x_{4-5}$ remains uncertain. In addition to the $x_{1-3}$ singlet peaks, there is also another CH$_2$ group ($x_6$) present that is likely to be another CH$_2$-PO$_4$ group with subtle electrical differences as compared to FBP C1 and C6 and G3P C1. This could be attributed to the slight change in overall molecular structure between FBP and TBP. Additionally, it could also indicate the formation of a monophosphorylated sugar such as F1P or F6P.

To provide a more detailed structural insight into how these protons are coupled a HSQC-TOCSY experiment was setup using the WT TtFBPA reaction mixture (Figure 5.13). Comparison of the cross peaks identified in Figure 5.12 to those observed in Figure 5.13 shows a number of additional coupled cross peaks present, corresponding to the interactions occurring between coupled protons. Cross peaks from Figure 5.12 have been annotated on to Figure 5.13 to differentiate the cross peaks from the coupled cross peaks.
Figure 5.13 shows two distinct product being formed. FBP is the expected product and can be identified due to the reference spectra, however, a second product is also being formed. Distinct coupling of cross peak FBP C6 and $x_1$ linked by the coupling cross peak at ppm f1:71.5, f2:3.87. Due to the tight clustering of the cross peaks labelled $x_1-3$, any potential coupled cross peaks would be hard to identify due to likely overlap. As mentioned previously, these three peaks likely correspond to CH groups bonded in the cyclic ring of TBP. The carbon atoms bonded to phosphate groups are likely to overlap with either the FBP C1, FBP C6 or G3P C1. The $^{31}$P-NMR spectrum of the same reaction (Figure 5.6) in addition to the crystallography date (Figure 4.5) suggests that this phosphorylated sugar can be identified as TBP. As TBP is a diastereomer of FBP, it is expected that the C1 and C6 carbons of these two molecules would have a similar chemical shift, although this could potentially not be the case. As the C4 carbon of TBP is shifted into the cis position the cross peaks observed for this carbon, along with the C3 and C5 carbons are likely to shift due to small changes in the electronegativity around these carbons. As with FBP, the C2 carbon of TBP is not coupled to any protons and therefore has no visible crosspeak on the 2D spectra. However, no evidence for a set coupled cross peaks linking two CH$_2$-PO$_4^{2-}$ groups exists. This indicates that infact a mono-phosphorylated sugar such as F1P or F6P may be formed. The cross peak labelled $x_5$ is coupled to $x_4$ by the coupling cross peak at ppm f1:58.5 f2:3.78. This compound is unlikely to be another phosphorylated sugar, due to a lack of coupling cross peaks to a CH$_2$ group. However, these two peaks appear unconnected to any other peaks.
5.6.3. 2D NMR of ScFruA Reaction Mixtures

In order to help determine the identity of reaction products observed in the $^{31}$P-NMR for ScFruA reactions, $^1$H/$^{13}$C 2D NMR HSQC was performed (Figure 5.14).

Cross peaks identified within the reference HSQC spectra for the FBP, DHAP and G3P were labelled (Figure 5.14). From this comparison to the reference spectra both substrates can be identified in the reaction. In addition, cross peaks corresponding to FBP are also observed within this reaction spectrum, as with the TtFBPA experiments there are a number of unidentified cross peaks labelled $x1$-$14$.

The initial observation from Figure 5.14 is that multiple products are formed which are distinct from FBP. Cross peaks labelled $x1$-$3$ are a series of tightly clustered singlet cross peaks that were identified in Figure 5.13, although a definitive identification was not concluded. The cross peaks labelled $x4$-$6$ and $x15$ are CH$_2$ doublet peaks, and most likely correspond to CH$_2$-PO$_4^{2-}$. Peaks $x9$-$14$ corresponds to unidentified singlet peaks, and may account for the additional features seen during $^{31}$P-NMR on the same reaction (Figure 5.10). In order to determine which of these unidentified cross peaks are connected HSQC-TOCSY was performed. The annotated 2D HSQC-TOCSY spectrum for the aldol reaction catalysed by ScFruA is shown in Figure 5.15.

The HSQC-TOCSY spectra identified 3 distinct products separate from FBP. There are a number of different species of coupled cross peaks, which have been annotated with coloured arrows (Green, Black and Blue). A distinct coupling cross peak is observed between the FBP C6 and $x1$ cross peaks at ppm f1:71.5, f2:3.87, as observed in the HSQC-TOCSY of the TtFBPA aldol reaction (blue arrows in Figure 5.15).
In addition, the x4 and x5 cross peaks appear to be able to couple to either a1 (ppm f1:62.5, f2:4.06) or a2 (f1:62.5, f2:4.27) however, coupling is difficult to interpret. The coupling cross peak a1 links x4 and x5 to cross peak x9 (black arrows in Figure 5.15). In turn, x9 may potentially couple to both x11 and x13, suggesting that these peaks belong to one species, although this result could be artificial. Although a coupling cross peak between x11 and x13 is not seen, the three peaks (x9, x11 and x13) could correspond to carbons within the cyclic portion of a phosphorylated sugar. As this product does not show the coupling to two groups likely to be CH$_2$-PO$_4^{2-}$ groups, this product is likely to be a mono-phosphorylated sugar such as F1P or F6P. The creation of multiple phosphorylated sugars with five member rings is supported by the presence of multiple coupled carbon cross peaks which in turn couple to phosphate bound carbons (species x4, x5 and x6). This indicates the formation of multiple phosphorylated, six carbon sugars, and indicates the aldolase does not strictly control the stereochemistry of the reaction.

The coupling cross peak a2 links x4 and x5 to x12, suggesting a second unknown species being formed within this reaction (green arrows within Figure 5.15). The carbon peak at x12 also couples to x14 which shows further coupling to carbon peak x6. In addition to x6 there could be coupling to x7 and x8; however, any potential coupling is difficult to observe due to the close proximity of the cross peaks. There also appears to be additional coupling of x12 to the x10 peak through a coupling peak at ppm f1: 77.5 f2:4.05. This species is harder to identify. However, as there is direct coupling from one phosphorylated
carbon (peaks x4-5) through to another phosphorylated carbon (x6) this molecule is not likely to be a cyclic sugar. This molecule is difficult identify without reference standards. However, it is possible to postulate that an alternate phosphorylated six carbon sugar is being produced.

Using HSQC-TOCSY the further resolution of the aldol reactions for both TtFBPA and ScFruA has been possible. Although full product identification was unsuccessful, the ability of these aldolases to catalyse the formation of multiple products has been demonstrated.

5.1. Discussion

In contrast to the retro-aldol reaction catalysed by aldolases, which can be easily monitored using enzyme coupled assays, the aldol reaction is more difficult to study. Furthermore, the Baylis-Hillman reaction has similarity to the aldol condensation reaction, such as its difficulty to monitor. In this study, NMR was used to directly assess product formation, both for retro-aldol, aldol and Baylis-Hillman reactions. Experiments were performed to monitor aldol activity of TtFBPA and ScFruA where 31P-NMR revealed TtFBPA formed a secondary product, postulated to be TBP or a singularly phosphorylated fructose sugar. In the case of ScFruA, the spectrum contained multiple peaks that could not confidently be assigned. Unfortunately, one fundamental problem with the phosphorus NMR conducted during this study was the inadequate buffering conditions used during the experimentation.
The buffering conditions used had not been sufficient to minimise pH changes during this study, this does not invalidate the results however but it makes analysis more difficult. If the study were to be repeated a significantly higher buffer concentration of 100 mM HEPES pH 7.0 would be used.

The use of $^1\text{H}/^{13}\text{C}$-NMR confirmed that TtFBPA catalysis formation of two distinct molecules while ScFruA was shown to form at least three molecules, although products were not identified. Some of these products are thought to be cyclic due to the similarity of cross peaks between the cycle portion of FBP and the unidentified products.

5.1.1. TtFBPA Activity

An initial proof of concept reaction using TtFBPA revealed that $^{31}\text{P}$-NMR was a suitable methodology for determining the retro-aldol activity of aldolases. As $^{31}\text{P}$-NMR only shows phosphorus atoms, Williams et al. used $^{31}\text{P}$-NMR as a tool for determining the presence of known products within aldol reaction [Williams et al. 2003].

The ability for TtFBPA to catalyze an aldol reaction was assessed using $^{31}\text{P}$ and $^1\text{H}/^{13}\text{C}$-NMR. $^{31}\text{P}$-NMR identified the formation of FBP based upon reference compounds. In addition, the formation of TBP was alluded to. $^1\text{H}/^{13}\text{C}$-NMR showed the formation of FBP, additionally, Figure 5.12 also showed three additional peaks (x1-3), which were tightly clustered group of singlet peaks. The WT TOCSY reaction reveals coupling cross peaks linking the FBP C6 cross peaks with the x1/cross peak. This molecule could potentially to be another bisphoshorylated, six carbon molecule with a five membered ring, and potentially an epimer of FBP. Additionally, the products of this reaction may also be fructose-1-phosphate or fructose-6-phosphate due to incomplete cross peak coupling. These sugars would display similar cross peaks to FBP, which is the observed result.
Unfortunately, the evidence is insufficient to determine the identity of the additional products formed just the presence of additional products.

Baylis-Hillman activity was also investigated using $^{31}$P-NMR, this was assessed as developing enzymatic reactions capable of developing new carbon-carbon bonds is useful in synthetic biology. Some changes were observed in the spectra compared to the HBOP and G3P reference spectra alone – specifically two peaks appearing as shoulders of the HBOP peaks. This most likely from higher resolution data obtained for the reaction spectrum compared to the reference spectrum. No further experimentation was performed as it was observed that no Baylis-Hillman activity was observed with these TtFBPA enzymes. Future experimentation to be conducted would be further NMR experimentation on a second round of TtFBPA mutants which have been designed to have enhanced HBOP binding.

5.1.2. ScFruA Activity

NMR was also used to determine the aldol activity of ScFruA, with the aim of identifying the products formed. $^{31}$P-NMR revealed a number of overlapping peaks for the aldol reaction. DHAP, G3P and $\alpha$ and $\beta$ FBP were identified, showing aldol activity. Additional peaks were observed which were thought to correspond to TBP. The formation of FBP and TBP was confirmed by HSQC-TOCSY.

The $^{31}$P-NMR spectrum contained too many peaks to accurately determine product identity. Based on the HSQC-TOCSY NMR spectrum for this reaction, it was determined that these peaks were likely to correspond to at least one phosphorylated cyclic sugar. HSQC NMR of ScFruA shows additional peaks relative to the TtFBPA. However, peaks x1-3 are identical for the peaks observed in the TtFBPA aldol reaction. Two additional
species were identified within the HSQC-TOCSY. Determining the full coupling cross peak profile is difficult and as a result not all peaks will have coupling cross peaks. However, peaks $x_9$, $x_{11}$ and $x_{13}$ all couple together. This molecule is likely to be another phosphorylated sugar, whether the molecule is cyclic is difficult to determine. However, the fact that cross peak $x_9$ links to both $x_{11}$ and $x_{13}$ alludes to this. In addition, species $a_2$ shows the coupling of either $x_4$ or $x_5$ to $x_6$. As with species $a_1$, species $a_2$ contains a cross peak ($x_{12}$) which couples directly to two other cross peaks. This is likely to be a second phosphorylated sugar similar to species $a_1$.

In total, the ScFruA reaction has potentially produced a total of four phosphorylated cyclic sugars. There are four epimers/diastereomers of FBP, which differ in their chirality of the hydroxyl groups at positions C3 and C4. The presence of four phosphorylated carbon molecules, which show similar coupling patterns across different cross peaks indicates that structurally similar products are being made. Minor differences in their electronic environment produces different cross peak profiles. These unidentified species could potentially be psicose-1,6-bisphosphate and xylulose-1,6-bisphosphate, the final two epimers/diastereomers of FBP. Alternatively, some of these products may be mono-phosphorylated fructose sugars such as F1P or F6P. The NMR evidence presented in this study is insufficient to accurately identify the additional products.

5.2. Conclusions

$^{31}$P-NMR has been known to be a successful method of demonstrating the enzymatic activity of aldolases in both the aldol and retro-aldol directions. These experiments revealed that TtFBPA and ScFruA not only synthesize FBP, but both also product secondary unknown products, which was confirmed using 2D HSQC-TOCSY NMR. The
identity of all of the products produced for aldol reactions using TtFBPA and ScFruA were not fully determined. However, both $^{31}$P-NMR and HSQC-TOCSY showed that both TtFBPA and especially ScFruA were catalysing the formation of novel products. Experiments were also performed to determine whether TtFBPA can function as a “Bayliss-Hillmanase”. Unfortunately, no activity could be observed under the conditions used.

In conclusion, both TtFBPA and ScFruA catalysed the formation of secondary products they have previously not been show to. However, the NMR evidence presented within this study is not sufficient to fully identify the products, just determine the presence of them. Combination of this data with additional mass spectrometry experimentation could combine coupled cross peak information with the total mass of the products which will help identify the unknown products.
Chapter Six

Conclusions and Implications

6. CONCLUSIONS AND IMPLICATIONS

This study sought to determine whether aldolases could be used as a template for the design or evolution of a new Baylis-Hillmanase enzyme. The rationale behind this approach relied on the similarity in mechanism of the aldol condensation and the Baylis-Hillman reaction. Two DHAP dependent aldolases were chosen for this study, both active at higher temperatures. One of these is the decameric TtFBPA while the other is the monomeric ScFruA. No crystal structure was available for the latter. However, it offers a number of advantages with regards to the industrial use of this enzyme: the generation of new enzyme fusions is much easier using monomeric enzymes as compared to multimeric species. Numerous studies have shown that aldolases are very promiscuous with regard to their acceptor molecule specificity, whilst maintaining a much higher level of specificity for the donor molecule [Blom et al. 1996; Williams et al. 2003; Turner 2009; Li et al. 2012; Fushinobu et al. 2011; Baker & Seah 2012]. This promiscuity makes the aldolases good targets for rational enzyme design to catalyze a new reaction. Hence, the Baylis-Hillman substrate HBOP was synthesized to mimic the native donor DHAP.
6.1. Detailed Characterization of TtFBA Activity

Kinetic characterization of the retro-aldol reaction for TtFBPA using an enzyme coupled activity assay revealed that TtFBPA and selected variants were able to convert FBP into DHAP and G3P. Crystallographic approaches to show ligand binding of DHAP were conducted. Although W144L showed the binding of DHAP to K177 via a Schiff base, WT TtFBPA showed the presence of a six carbon sugar molecule, which was determined to be TBP. WT was thought to be able to form TBP due to the hydrogen bonding and therefore stabilization of W144 with TBP C4-OH.

Although the retro-aldol reaction can be monitored easily, the aldol reaction is much more difficult to study. Williams et al. demonstrated, $^{31}$P-NMR can be used to study the aldolase reaction, and to differentiate between FBP and TBP products [Williams et al. 2003]. Hence, similar reactions were monitored using $^{31}$P-NMR with TtFBPA. The NMR experiments performed in this study showed that TtFBPA was capable of catalyzing the formation of multiple sugar species, in this case FBP and potentially TBP.

Although the $^{31}$P NMR experimentation could not definitively show the formation of TBP, in combination with crystallography experiments it shows that under at ambient temperatures and high substrate availability, TtFBPA is able to catalyze the formation of TBP. Our data demonstrates that TtFBPA does not fully control stereo specificity of the aldol reaction, a result which has not been observed before. 2D NMR was conducted with the aim of providing additional information to identify any products formed during the aldol reaction. The results revealed that although there is evidence of TBP production, it is not definitive in NMR, additionally, there is also evidence for the production of F1P or F6P rather than di-phosphorylated sugars. The conclusions from the NMR experimentation
is that a secondary product is being formed, however, its identity is yet unknown, but in combination with the crystallographic data is likely to be TBP.

**6.2. Characterization of TtFBPA Baylis-Hillman Activity**

There has been several research articles aimed at discovering and designing Baylis-Hillman enzymes [Jiang & Yu 2014; Kapoor et al. 2014; Reetz et al. 2007]. However, due to the high degree of similarity between the Baylis-Hillman and aldol reactions and the high levels of substrate promiscuity observed in aldolases, an aldolase enzyme may be a more suitable target for the design of a Baylis-Hillman reaction. In an attempt to understand the underlying reasons, the crystal structures of the complex with the Baylis-Hillman donor substrate HBOP to both WT and W144L TtFBPA were determined. In the case of WT TtFBPA, density appears consistent with a Schiff base formed between HBOP and the catalytic Lys, a species that would not lead to a Baylis-Hillman type reaction. In contrast, certain active sites of the W144L mutant contained density that can be interpreted as a HBOP-Lys adduct or enolate intermediate, similar to that observed in the Baylis-Hillman reaction with DABCO [Tang et al. 2013].

Although crystallographic approaches showed initial ligand binding, unfortunately no Baylis-Hillman activity was observed using NMR. The crystal structures however provides an initial framework for further mutagenesis trials aimed at introducing amino acid residues which allow the enzyme to progress further through the Baylis-Hillman reaction mechanism.
6.3. Determination of ScFruA Biochemical and Structural Properties

The research performed within this thesis and the conclusions drawn provide novel biochemical and structural insight into ScFruA, an aldolase that has so far not been well characterized. It was speculated that ScFruA was a monomer based on *P. gingivalis* FBP aldolase, however, this had never been confirmed. Our data confirmed ScFruA as monomeric, with a much higher retro-aldol activity at 50 °C compared to TtFBPA. This was expected as Brockamp & Kula showed that ScFruA is stable up to 60 °C, while TtFBPA is isolated from a hyperthermophile and operates at an 85 °C optimum temperature [Brockamp & Kula 1990].

The crystal structure for ScFruA was solved in the ligand free form to 2.17Å and bound to DHAP to 1.95 Å. This gave unique insight into the active site structure of a previously unknown aldolase. While initial soaking experiments with DHAP were successful, further trapping experiments conducted using both DHAP and G3P failed. Our crystal structure data will aid further protein engineering studies that make use of rational enzyme design methodologies. Our NMR studies of the ScFruA reaction concluded that ScFruA excerpts little control of the stereochemistry of the reaction, and suggest it is likely that ScFruA catalyzes the formation of both mono and di phosphorylated sugars. In total 4 unique products were identified, one of which was identified as FBP. Of the other three, one was shown to be a di-phosphorylated sugar, thought to be TBP. While the other two are thought to be mono-phosphorylated sugars which could correspond to F1P and F6P.

The conclusions from these studies are that ScFruA lacks substrate specificity, which is typical of aldolase enzymes. However, the full identity of the products could not be identified.
6.4. Implications and Further Investigations

Our data provides a more thorough description of both the TtFBPA and ScFruA aldolase structures and mechanism. The fact both enzymes lack strict stereochemical control could be advantageous for generation of the initial active Baylis-Hillmanases as it suggests a variety of substrates/conformations can be accepted within the active site. In the case of TtFBPA, the enlarged active site pocket of the W144L mutant appears to have enabled formation of the desired HBOP Lys-177 adduct. Therefore, TtFBPA could provide the initial framework for the engineering of a Baylis-Hillmanase, given the fact the initial catalyst-binding step of the reaction can be observed. Unfortunately, a high throughput methodology for screening aldol activity could not be determined, limiting the ability to screen large libraries of mutant enzymes based on the W144L scaffold.

An enzymatic Baylis-Hillman reaction would contain similarities between the aldolase reaction and the chemical MBH reaction. Due to the proximity of Asp-24 in TtFBPA W144L to the HBOP carbonyl, it is likely that this residue is able to act as an acid/base during the reaction mechanism [Lorentzen et al. 2005]. Additionally Tyr-146 may also be able to act as an acid/base for the reaction. A mechanism and active site pocket for an enzymatic Baylis-Hillman reaction using TtFBPA has been proposed (Figure 6.1).

Future studies might include the alteration of the phosphate moiety of HBOP to a Cbz protecting group, in conjunction with further mutagenesis studies. Gutierrez et al. successfully showed that mutagenesis of an E. coli FSA allowed the DHAP phosphate to be substituted for a Cbz protecting group [Gutierrez et al. 2011]. This greatly increases the industrial chemical applications of the product as phosphate groups are rarely required in synthetic chemistry and their removal typically results in the loss of a functional group. Additional studies could also focus on determining whether ScFruA is able to form a
complex with HBOP as time constraints prevented this experimentation during this study.

Figure 6.1: The proposed enzymatic Baylis-Hillman reaction and active site pocket catalyzed by TtFBPA. (A) The proposed reaction mechanism for an enzymatic Baylis-Hillman reaction between HBOP and G3P catalyzed by a TtFBPA enzyme. (B) The active site structure of TtFBPA W144L showing binding to HBOP and proximal amino acid residues. Images made using Pymol version 1.7.4.4.

Straightforward asymmetric entry to highly functionalized 3-substituted 3-hydroxy-
beta-lactams via Baylis-Hillman or bromoallylation reactions. The Journal of organic

analysis and over-expression of the gene for the class II fructose 1,6-bisphosphate

Amarante, G.W., Benassi, M., Milagre, H.M.S., Braga, A. a C., Maseras, F., Eberlin, M.N.,
mechanistic view for thioureas revealed by ESI-MS/(MS) monitoring and DFT
calculations. Chemistry (Weinheim an der Bergstrasse, Germany), 15(45), pp.12460–

determination of methanol with alcohol oxidase. Application to the assay of pectin


specificity, stereospecificity, and catalysis in the class II pyruvate aldolase, Bphl.
Biochemistry, 50(17), pp.3559–3569.


Basavaiah, D., Rao, A.J., & Satyanarayana, T., 2003. Recent advances in the Baylis-

for creativity in chemistry. Chemical Society reviews.

Bednarski, M.D., Simon, E.S., Bischofberger, N., Fessner, W.D., Kim, M.J., Lees, W.,
catalyst in organic synthesis. Journal of the American Chemical Society, 111(2),
pp.627–635. Available at: http://pubs.acs.org/doi/abs/10.1021/ja00184a034

palladium catalysis. Chemical Reviews, 100(8), pp.3009–3066.

Blom, N.S., Tétreault, S., Coulombe, R., & Sygusch, J., 1996. Novel active site in

Bibliography


Lee, C.-H., & Frasch, C.E., 2001. Quantification of Bacterial Polysaccharides by the


Tang, M., Zhang, W., & Kong, Y., 2013. DABCO-promoted synthesis of pyrazoles from tosylhydrazones and nitroalkenes. *Organic & biomolecular chemistry*, 11(37),


8. APPENDIX

8.1. Cloning Vectors

Restriction sites NdeI and XhoI
Total Insert Size 954 bases
Length of Linker 27 bases

Figure 8.1: Vector map for the pET28a vector. Cloning into the pET28a vector utilized the NdeI and XhoI restriction sites, resulting in an N-terminal His6-tag with a 9 amino acid linker region between the protein construct and tag.
Restriction sites
NcoI and XhoI

Total Insert Size
925 bases

Length of Linker
6 bases

Figure 8.2: Vector map for the pET28a vector. Cloning into the pET28b vector utilized the NcoI and XhoI restriction site which resulted in a C-terminal His$_6$-tag with a 2 amino acid linker region between the protein construct and tag.
8.2. Mass Spectrometry References

Figure 8.3: ESI reference spectrum for FBP. ESI mass spectrometry spectra for 25 mM FBP collected using a Bruker Ultraflex II TOF/TOF spectrometer. A major peak at 338.99 Da and minor peaks at 678.98 and 1018.98 Da are observed which correspond to FBP and multiple deprotonated forms of FBP. The expected peak for FBP is 340.116 Da.

Figure 8.4: ESI reference spectrum for DHAP. ESI mass spectrometry spectra for 25 mM DHAP collected using a Bruker Ultraflex II TOF/TOF spectrometer. A major peak at 168.99 Da and minor peaks at 360.97 Da are observed which correspond to DHAP and the double deprotonated form DHAP. The expected peak for DHAP is 170.06 Da.
Figure 8.5: ESI reference spectrum for G3P. ESI mass spectrometry spectra for 25 mM FBP collected using a Bruker Ultraflex II TOF/TOF spectrometer. A major peak contaminant peak at 338.99 Da is observed which corresponds to FBP. A minor peak at 168.99 Da is present which corresponds to G3P. The expected peak for G3P is 170.06 Da.