An Investigation of the Molecular Determinants of Substrate Channeling and Allosteric Activation in Aldolase-Dehydrogenase Complexes

By
Jason Carere

A Thesis
presented to
The University of Guelph

In partial fulfillment of requirements
for the degree of
Doctor of Philosophy
in
Molecular and Cellular Biology

Guelph, Ontario, Canada
© Jason Carere, April, 2013
Abstract

An Investigation of the Molecular Determinants of Substrate Channeling and Allosteric Activation in Aldolase-Dehydrogenase Complexes

Jason Carere
University of Guelph 2013
Advisor: Stephen Seah

The aldolase-dehydrogenase complex catalyzes the last two steps in the microbial meta-cleavage pathway of various aromatic compounds including polychlorinated biphenyls (bph pathway) and cholesterol (hsa pathway). The aldolase, BphI, cleaves 4-hydroxy-2-oxoacids to produce pyruvate and an aldehyde. Linear aldehydes of up to six carbons long and branched isobutyraldehyde were directly channeled to the aldehyde dehydrogenase BphJ, via a molecular tunnel, with greater than 80% efficiency. The molecular tunnel is narrow in positions lined by Gly-322 and Gly-323 in the aldolase. BphI variants G322F, G322L and G323F were found to block aldehyde channeling. The replacement of Asn-170 in BphJ with alanine and aspartate did not substantially alter aldehyde channeling efficiencies, thus disproving a previous hypothesis that hydrogen bonding between the Asn-170 and the nicotinamide cofactor induces the opening of the exit of the tunnel. The H20A and Y290F BphI variants displayed significantly reduced aldehyde channeling efficiencies indicating that these residues control the entry and exit of substrates and products from the aldolase reaction.

The BphI reaction was activated by NADH binding to BphJ in the wild-type enzyme and channel blocked variants. Activation of BphI by BphJ N170A, N170D and I171A was decreased by ≥ 3-fold in the presence of NADH and ≥ 4.5-fold when BphJ was undergoing turnover. These results demonstrate that the dehydrogenase coordinates catalytic activity of BphI through allostery rather than through faster aldehyde release from substrate channeling.
HsaF, an ortholog of BphI from *Mycobacterium tuberculosis* could be expressed as a soluble dimer, however HsaF was inactive in the absence of HsaG, a BphJ ortholog. Acetaldehyde and propionaldehyde were channeled directly to HsaG with similar efficiencies as in the BphI-BphJ system. The HsaF-HsaG complex was crystallized and its structure solved to a resolution of 1.93 Å. Substitution of Ser-41 in HsaG with isoleucine or aspartate resulted in about 35-fold increase in $K_m$ for CoA but only 4-fold increase in $K_m$ for dephospho-CoA, confirming its importance in interacting with the 3’- ribose phosphate of CoA. A second gene annotated as 4-hydroxy-2-oxopentanoic acid aldolase (*Rv3469c*) from *M. tuberculosis* was expressed, purified and found to possess oxaloacetate decarboxylase and not aldolase activity.
Acknowledgements

I would like to thank my family, friends and colleagues at the University. My parents, Ron and Colleen, have always supported me in whatever I set out to do; I have a profound appreciation for all you have done for me. My sister Janice has been an inspiration throughout my degree. To the Boys, you are truly the best friends a guy could ask for; more fun than imaginable and always there when needed. Always remember, be better, go deep, and keep your eyes on the road.

My colleagues at Guelph exemplify what research is about; the dedication they show motivates me to work harder and their insights were regularly helpful. To Amanda R., Chris, Filomena, Amanda P. and Laura, it was great working with you in the lab. Sarah McKenna worked with me as an undergraduate student; her many hours purifying protein and setting trays were key in the success of this project. I owe a huge debt to Dr. Perrin Baker and Scott Mazurkewich, great friends and co-workers. I think the three of us agree that a little healthy competition and taking the time to help each other out made all of us more productive and better scientists over our time here. I would like to thank Dr. Janet Wood, Dr. Rod Merrill and Dr. Steffen Graether for sitting on my advisory committee. Dr. Matthew Kimber was generous enough to provide training in protein crystallography and instrumental in structure determination. I would like to thank the staff at the Canadian Macromolecular Crystallography Facility at the Canadian Light Source for running the Mx Data Collection School.

I would like to thank the University of Guelph and the Department of Molecular and Cellular Biology for providing me with a great education. The Government of Ontario has provided funding during my degree, including an Ontario Graduate Scholarship and the Queen
Elizabeth II Graduate Scholarship in Science and Technology. I would also like to thank all of the other University affiliated organizations which have provided funding in the form of scholarships and grants.

Finally I would like to thank Dr. Stephen Seah; I could not have asked for a better supervisor. Stephen was always around to bounce ideas off of and was never afraid to tell you something you did not want to hear if he thought it would improve your research. He was an extremely adaptive supervisor, his approach varied for different students at different times, to ensure they were as productive as possible while still enjoying their time at the University. Stephen went above and beyond with the amount of time he spent reviewing scholarships, manuscripts and presentations not only making corrections but also teaching me how to become a more effective communicator. I truly believe I would have been less successful without him as a supervisor; thank you Stephen.
Authors Declaration of Work Completed

I declare that unless otherwise stated I have completed all work contained in this thesis.
Table of Contents

Abstract ........................................................................................................................................... ii
Acknowledgements ......................................................................................................................... iv
List of Tables .................................................................................................................................... x
List of Figures .................................................................................................................................. xii
List of Equations .............................................................................................................................. xiv
Abbreviations .................................................................................................................................. xv

Chapter 1: Introduction .................................................................................................................... 1
1.1 meta-Cleavage Pathway ............................................................................................................. 1
1.2 Degradation of PCBs ................................................................................................................. 4
1.3 Cholesterol Degradation .......................................................................................................... 4
1.4 Aldolases ................................................................................................................................... 9
1.5 Dehydrogenases ....................................................................................................................... 10
1.6 Aldolase-Dehydrogenase Complexes in meta-Cleavage Pathways ..................................... 10
1.7 General Structure of DmpG-DmpF ........................................................................................ 17
1.8 Substrate Channeling .............................................................................................................. 23
  1.8.1 Direct Substrate Channeling ............................................................................................... 24
  1.8.2 Substrate Channeling by Proximity .................................................................................... 28
1.9 Evidence of Substrate Channeling in the Aldolase-Dehydrogenase Complex ................... 32
1.10 Justification of Research ........................................................................................................ 33

Chapter 2: Materials and Methods ............................................................................................... 35
2.1 Chemicals and Substrates ....................................................................................................... 35
  2.1.1 Chemicals .......................................................................................................................... 35
  2.1.2 4-hydroxy-2-oxoacid Preparation ......................................................................................... 35
2.2 DNA Manipulation and Cloning .............................................................................................. 36
  2.2.1 Bacterial Strains and Plasmids ............................................................................................ 36
  2.2.2 Cloning of hsaF and hsaG ................................................................................................. 37
  2.2.3 Restriction Enzyme Digestions ............................................................................................ 38
  2.2.4 DNA Ligation ..................................................................................................................... 38
  2.2.5 Agarose Gel Electrophoresis .............................................................................................. 38
  2.2.6 DNA Fragment Purification from Agarose ......................................................................... 38
List of Tables

Table 1.1 Example enzyme systems that undergo substrate channeling ........................................ 29

Table 2.1 Primers for site-directed mutagenesis............................................................................. 40

Table 3.1 Composition of solutions used to elute the various acyl-CoAs from the C18 column. 61
Table 3.2 Channeling efficiency of various aldehydes in the BphI-BphJ enzyme complex ...... 62

Table 4.1 Steady-state kinetic constants of BphI and variants for the aldol cleavage of HOPA and
HOHA ........................................................................................................................................... 77
Table 4.2 Aldehyde channeling efficiencies of BphI variants-BphJ complexes ......................... 78
Table 4.3 Kinetic constants for acetaldehyde and propionaldehyde of wild-type BphJ associated
with the BphI glycine variants ..................................................................................................... 80
Table 4.4 Substrate channeling efficiency in enzyme complexes .............................................. 81
Table 4.5 Channeling efficiencies of Y290F and H20A variants of BphI ................................. 83
Table 4.6 Steady-state kinetic parameters of acetaldehyde for various variants ......................... 84
Table 4.7 Steady-state kinetic parameters of cofactors in dehydrogenase reaction ................. 85
Table 4.8 Acetaldehyde channeling efficiencies in BphI-BphJ variant complexes .................. 86
Table 4.9 Steady-state kinetic constants of BphI and glycine variants for the aldol cleavage of
HOPA in the absence of BphJ cofactors ...................................................................................... 88
Table 4.10 Effects of NADH on steady-state kinetic parameters for the aldol cleavage of HOPA
by BphI .......................................................................................................................................... 89
Table 4.11 Dissociation constants of NAD$^+$ for wild-type and dehydrogenase variants ........ 90
Table 5.1 Metal ion specificity of HsaF ................................................................. 111
Table 5.2 Steady state kinetic parameters of HsaF ...................................................... 112
Table 5.3 Steady state kinetic parameters of HsaG with various substrate and cofactors .... 116
Table 5.4 Steady state kinetics of HsaG as a monomer.................................................. 119
Table 5.5 Data collection, model refinement and final structure statistics for HsaF-HsaG ...... 120
Table 5.6 Steady state kinetics of HsaG variants ............................................................ 132
List of Figures

Figure 1.1 meta-cleavage pathways of various aromatic compounds. ........................................ 2
Figure 1.2 General chemical structure of PCBs ........................................................................... 5
Figure 1.3 Cholesterol degradation in M. tuberculosis................................................................. 7
Figure 1.4 Reaction catalyzed by the aldolase-dehydrogenase complex ..................................... 11
Figure 1.5 Reaction mechanism of BphI ...................................................................................... 14
Figure 1.6 Proposed reaction mechanism of the aldehyde dehydrogenase. ............................... 16
Figure 1.7 Dendrogram of a phylogenetic analysis of aldolases (A) and dehydrogenases (B) ... 18
Figure 1.8 Crystal structure of DmpF-DmpG ........................................................................... 19

Figure 3.1 Structure of DmpG-DmpF and the channel ................................................................. 57
Figure 3.2 Graphical representation of the diameter of the aldehyde channel as determined using
the program MOLE (156) ........................................................................................................... 58
Figure 3.3 Chromatogram of various acyl-CoA esters separated on a HyPURITY C18 (Thermo
Scientific) HPLC column ........................................................................................................ 60

Figure 4.1 Model of 4-hydroxy-2-oxononanoate bound in the DmpG active site ......................... 69
Figure 4.2 Sequence alignment of BphI and orthologs ................................................................. 70
Figure 4.3 Structural changes upon NAD\(^+\) binding in DmpF .................................................... 73
Figure 4.4. SDS-PAGE gel of purified BphI-BphJ wild-type and variants ...................................... 76
Figure 4.5. Tryptophan florescence quenching of BphJ N170D variant ........................................ 91
Figure 4.6 Model of the L90A variant of the aldolase DmpG ........................................................ 96
Figure 4.7 Proposed mechanism of substrate channeling in BphI-BphJ ........................................ 98
Figure 5.1 Reaction catalyzed by HsaF and HsaG ................................................................. 106
Figure 5.2 SDS-PAGE gel of purified HsaF-HsaG, HsaF, HsaG and Rv3469c. ...................... 108
Figure 5.3. Elution volume of purified \textit{M. tuberculosis} enzymes. ....................................... 109
Figure 5.4. Representative kinetic plots of wild-type HsaF-HsaG utilizing various substrates. 113
Figure 5.5 Activity of HsaF when preincubated with varying concentrations of HsaG.......... 114
Figure 5.6. Tryptophan florescence quenching of HsaG when in complex with HsaF.......... 118
Figure 5.7 Structure of HsaF-HsaG .................................................................................. 121
Figure 5.8 Variable length loops on the dehydrogenases. ................................................. 123
Figure 5.9 Homology model of Rv3469c superimposed on HsaF........................................ 124
Figure 5.10 Interaction of S41 with modeled CoA............................................................... 126
Figure 5.11 Sequence alignment of 4-hydroxy-2-oxoacid aldolases with protein orthologs of
Rv3469c. .............................................................................................................................. 128
Figure 5.12 Genetic map of Rv3469c and homologs. .......................................................... 131
List of Equations

Equation 1 Channeling Efficiency..........................................................49
Equation 2 Dissociation Constant (one site)...........................................50
Equation 3 Sigmoidal Binding Model.....................................................50
Abbreviations

ADH, alcohol dehydrogenase
ALDH, aldehyde dehydrogenase
ASADH, aspartate-semialdehyde dehydrogenase
CD, circular dichroism
CHA, 5-carboxy-4-hydroxy-2-oxoadipate
CLS, Canadian Light Source
CMCF, Canadian Macromolecular Crystallography Facility
CPS, carbamoyl phosphate synthetase
CoA, Coenzyme A
DTT, dithiothreitol
EDTA, ethylenedinitrilotetraacetic acid
FSB, frozen storage buffer
G3P, glyceraldehyde-3-phosphate
GAPDH, glyceraldehyde-3-phosphate dehydrogenase
HEPES, N-(2-hydroxyethyl)-piperazine-N’-2-thanesulfonic acid
HMG, 4-hydroxy-4-methyl-2-oxoglurrate
HMG-CoA, 3-hydroxy-3-methylglutaryl-coenzyme A
HOPA, 4-hydroxy-2-oxopentanoate
HOHA, 4-hydroxy-2-oxohexanoate
HPLC, high-performance liquid chromatography
IGP, indole-3-glycerol phosphate
IPTG, isopropyl β-D-thiogalactopyranoside
KDO, 2-Keto-3-deoxyocton
LB, Luria-Bertani
LDH, lactate dehydrogenase
OAA, oxaloacetate
P5C, Δ1-pyrroline-5-carboxylate
P5CDH, Δ1-pyrroline-5-carboxylate dehydrogenase
PRODH, proline dehydrogenase
TCA, tricarboxylic acid
TIM, triose-phosphate isomerase
Tris, tris(hydroxymethyl)aminomethane
WT, wild-type
Chapter 1: Introduction

1.1 meta-Cleavage Pathway

The meta-cleavage pathway is one of the most common pathways utilized by aerobic bacteria to convert aromatic compounds to tricarboxylic acid (TCA) cycle intermediates. This pathway can be divided into two sections, the upper and lower pathways. The upper pathway involves hydroxylation of the aromatic ring which is then cleaved, adjacent to one of the hydroxyl substituents, by extradiol ring-cleavage dioxygenases (1, 2). The lower pathway is responsible for further degradation of the ring cleavage products into TCA cycle intermediates (2). Acetyl CoA and pyruvate are produced from the lower pathways in the catabolism of toluene, biphenyl and phenol (3, 4). Degradation of cholesterol on the other hand leads to formation of propionyl CoA and pyruvate (5). Figure 1.1 shows the pathway for biphenyl (bph pathway), phenol (dmp pathway) and cholesterol (hsa pathway) degradation for comparison (5, 6).

The upper pathway has been well characterized in many species, while many details about the lower pathway have not been elucidated. The aim of the work is to characterize BphI/BphJ and HsaF/HsaG, the aldolase and dehydrogenase complexes of the bph pathway from *Burkholderia xenovorans* LB400 and cholesterol degradation pathway of *Mycobacterium tuberculosis*, respectively. *B. xenovorans* LB400 was isolated from a landfill in New York State containing polychlorinated biphenyls (PCBs) (7). It can degrade more than 20 PCB congeners with up to 6 chlorines on the biphenyl ring (8, 9). *M. tuberculosis*, on the other hand, is the causative agent of tuberculosis, which is a persistent respiratory infection that causes millions of deaths each year. Infection rates are also beginning to climb as the HIV virus spreads, making those with AIDS more susceptible to infection (10). Increasing drug resistance is also playing a
Figure 1.1 *meta*-cleavage pathways of various aromatic compounds.

Biphenyl (A), phenol (B) and cholesterol (C) are transformed via the respective pathways to a 4-hydroxy-2-oxoacid intermediate. This is then cleaved by the aldolase to produce pyruvate and an aldehyde. The aldehyde is transformed to acyl-CoA and NADH by the aldehyde dehydrogenase using Coenzyme A and NAD⁺ as cofactors. (A) *bph* pathway in *Burkholderia xenovorans* LB400 involved in the degradation of biphenyl. (B) *dmp* pathway in *Pseudomonas* strain CF600 involved in the degradation of phenol (C) *hsa* pathway in *Mycobacterium tuberculosis* involved in the degradation of cholesterol. BphI, DmpG and HsaF are homologous aldolases while BphJ, DmpF and HsaG are homologous aldehyde dehydrogenases.
large role in the resurgence of tuberculosis. During chronic infection, *M. tuberculosis* primarily resides within a phagosome-like compartment of macrophages and, as will be discussed in more detail in subsequent sections, cholesterol catabolism appears to be important for its survival within macrophages (11).

1.2 Degradation of PCBs

PCBs are stable aromatic compounds that were produced in large quantities for use as hydraulic fluids, flame retardants and dielectric fluids (12). They are toxic, fat soluble compounds and widespread use has led to environmental contamination (13). Several bacterial strains have been isolated that degrade specific PCB congeners. Due to the large number of different theoretical congeners (209) (Figure 1.2) and the fact that they are manufactured and used as mixtures, multiple bacterial strains are normally required for remediation of PCB contaminated environments (8, 12, 14). Due to its ability to degrade a wide range of PCBs, *Burkholderia xenovorans* LB400 has been the subject of extensive studies (15). *B. xenovorans* LB400 can hydroxylate PCBs by 3,4-dioxygenation or 2,3 dioxygenation allowing it to use a variety of PCB congeners as its sole carbon source (3, 13, 16). *B. xenovorans* LB400 is also interesting due to its ability to also degrade single-ring aromatic hydrocarbons whereas many other PCB degrading strains can only use aromatics with a biphenyl nucleus (17).

1.3 Cholesterol Degradation

During early infection, *M. tuberculosis* can replicate rapidly in the macrophage before the host’s immune response sets in. The release of cytokines including IFN-γ may control the replication of the bacteria, but low levels of the bacteria will remain (18). This results in chronic
Figure 1.2 General chemical structure of PCBs.
Chlorine may replace any hydrogen on the carbon atoms of the biphenyl rings leading to 209 potential congeners.
infection. The recruitment of macrophages, Langhans giant cells and lymphocytes results in the formation of granulomas (tubercles) (19) that can contain the infection, preventing further spread of the bacilli. However, over time this may lead to tissue necrosis, tissue liquefaction and discharge of the bacteria causing further infection. Restricting energy sources in the macrophage may slow down bacterial growth. It has been proposed that *M. tuberculosis* can use host membrane cholesterol as its sole carbon and energy source (20). *M. tuberculosis* can also induce the uptake of low density lipoprotein (LDL) into macrophages resulting in the formation of foamy macrophages which contain lipid bodies and accessible cholesterol esters that could potentially be used as a carbon source for the intracellular pathogen. Thus the cholesterol degradation pathway in *M. tuberculosis* would allow the bacteria to survive for long periods of time in macrophages (21). The enzymes required for cholesterol degradation in *M. tuberculosis* are primarily encoded by a group of 83 genes (22, 23). Inhibition of the cholesterol degradation pathway enzymes in *M. tuberculosis* has been touted as a possible therapeutic strategy to treat tuberculosis (24).

Several other bacteria have evolved to partially or completely degrade cholesterol including *Rhodococcus, Azotobacter, Pseudomonas and Streptomyces* (25-29). Cholesterol degradation is generally separated into two main parts: steroid ring degradation and side chain degradation by β-oxidation (Figure 1.3). The *hsa* and *ksh* pathways are responsible for the degradation of the steroid rings, while the aliphatic side chain is degraded similarly to β-oxidation of fatty acids (30). The *ksh* pathway (from *M. tuberculosis*) is responsible for opening ring B and the aromatization of ring A of the cholesterol molecule, producing 3-hydroxy-9,10-secoandrost-1,3,5(10)-triene-9,17-dione (HSA). This catechol is then partially degraded by the *hsa* pathway, also from *M. tuberculosis*, producing TCA cycle intermediates. Early steps in
Figure 1.3 Cholesterol degradation in *M. tuberculosis*.
A) Cleavage of the steroid rings by the *ksh* and *hsa* pathways. B) Degradation of the side chain of cholesterol.
steroid ring degradation have been characterized (31-34) while most of the lower pathway enzymes have not been investigated. Steps in β-oxidation of the side chain have more recently been elucidated. The genes encoding these enzymes in M. tuberculosis are located in an operon, termed the intercellular growth (igr) operon. Strains of M. tuberculosis lacking the igr operon are unable to grow on cholesterol as a sole carbon source (35). This strain is interesting because growth of the bacteria is inhibited when they are cultured in glycerol and glucose media that contain cholesterol relative to media containing no cholesterol. When a cholesterol importer (Rv3501c) is deleted from the Δigr strain, the sensitivity to cholesterol is removed, indicating that a toxic cholesterol metabolite accumulates in Δigr strains.

Genes involved in cholesterol degradation are up-regulated in macrophages and are essential for growth in mice (36). It has recently been shown that although cholesterol may act as the sole carbon source in vitro, it cannot serve as the sole carbon source in vivo (37). Whether cholesterol is a sole carbon source may vary depending on the cell type and treatment of the cells, but it is evident that cholesterol is important during infection. Disruption of the cholesterol degradation pathway could deprive the organism from a potential carbon and energy source or lead to an accumulation of toxic steroid intermediates (38). Griffin and colleagues demonstrate that hsaF and hsaG, the genes encoding the aldolase and dehydrogenase in the pathway, are not critical to survival on cholesterol in vitro, but the products (pyruvate and propionyl-CoA) are thought to contribute to glyoxylate- and methylcitrate-cycles which are critical to persistent infection (39, 40).

Although not essential for growth on cholesterol, it has been shown by microarray studies that hsaF expression is induced in macrophages compared to 7H9 media (36, 41). One difference between the cholesterol and the PCB degradation pathways is that ring cleavage followed by a
hydration reaction leads to the product 4-hydroxy-2-oxopenatanate and 4-hydroxy-2-oxohexanate, respectively. Thus the aldolases (BphI and HsaF) and dehydrogenases (BphJ and HsaG) that catalyze the subsequent transformation of these products act on substrates that differ in length by one carbon atom. These aldolases and dehydrogenases are the subject of investigation in this thesis and will be discussed in more detail in subsequent sections.

1.4 Aldolases

Aldolases are a broad class of enzymes which catalyze the formation or cleavage of carbon-carbon bonds (42). Class I aldolases contain a catalytic lysine residue that forms a Schiff base intermediate with the substrate in the aldolase’s reaction mechanism (43-45). Class II aldolases utilize a divalent metal cofactor which stabilizes a negatively charged intermediate (46, 47). Class II aldolases are further classified based on the substrate they utilize such as pyruvate, phosphoenolpyruvate or dihydroxyacetone phosphate. The aldolases in the aldolase-dehydrogenase complexes presented here are examples of class II pyruvate aldolases. Class II pyruvate aldolases generally have an octahedrally coordinated divalent metal ion (48). The pyruvyl moiety coordinates the bound metal ion, which is often Mg$^{2+}$ or Mn$^{2+}$, in a bidentate fashion using the C1 and C2 oxygen groups (49, 50). The catalytic mechanism generally proceeds by base abstraction of a proton from a C4 hydroxyl of a 4-hydroxy-2-oxoacid substrate leading to C-C bond cleavage. The pyruvate enolate intermediate is stabilized by coordination to the metal ion and also by a positively charged arginine (51, 52). The enolate is then protonated by a catalytic acid to form pyruvate.
1.5 Dehydrogenases

Aldehyde dehydrogenases (ALDHs) are a broad group of enzymes which metabolize various endogenous and exogenous aldehydes (53). They are found in all forms of life from prokaryotes to eukaryotes. ALDHs have numerous functions including detoxification pathways and central metabolism. ALDHs may have diverged from a common ancestor (54-56). Two groups of ALDHs exist. Phosphorylating and non-phosphorylating ALDHs differ based on their acyl acceptor. Both groups have a cysteine residue that acts as a nucleophile to attack the carbonyl carbon of the aldehyde to form a thiohemiacetal intermediate. A hydride is then transferred to NAD(P)^+ to produce a thioacylenzyme intermediate. Phosphorylating ALDHs use inorganic phosphate as the acyl acceptor resulting in an acyl-phosphate product (57, 58). Non-phosphorylating ALDHs use either water (hydrolytic) or coenzyme A (acylating) as acyl acceptor. Acylating ALDHs exhibit a ping-pong mechanism in which the NAD(P)H and CoA bind the enzyme alternately whereas hydrolytic ALDHs follow a sequential kinetic mechanism with NAD(P)H released last (59-62). The dehydrogenases associated with class two pyruvate aldolases in the meta-cleavage pathway are CoA-dependent or acylating aldehyde dehydrogenases (50).

1.6 Aldolase-Dehydrogenase Complexes in meta-Cleavage Pathways

The aldolase and dehydrogenase from the meta-cleavage pathways form a complex. The aldolase catalyses the cleavage of 4-hydroxy-2-oxoacids producing an aldehyde and pyruvate. The dehydrogenase produces acyl-CoA from aldehyde, NAD^+ and coenzyme A (Figure 1.4).

DmpG/DmpF, from the phenol degradation pathway (dmp pathway), has been purified from the phenol-degrading bacteria Pseudomonas sp. CF600 as follows (63). The aldolase
Figure 1.4 Reaction catalyzed by the aldolase-dehydrogenase complex.

4-hydroxy-2-oxoacid is cleaved by the aldolase to produce pyruvate and aldehyde. The aldehyde is transformed to acyl-CoA using Coenzyme A and NAD\(^+\) as cofactors. R = C\(_x\)H\(_{2x+1}\)
(DmpG) and dehydrogenase (DmpF) activities co-purified in six separation steps that consisted of anion exchange, hydrophobic interaction, gel filtration, blue sepharose, Matrix Gel Green chromatographic steps and ammonium sulfate fractionation. The aldolase activity from the purified protein complex was enhanced in the presence of Mn$^{2+}$ while other divalent cations either inhibited activity or had no effect. The aldolase activity was also found to be stimulated by both NADH and to a lesser extent NAD$^+$. The dehydrogenase, on the other hand, was able to utilize several different aldehydes as substrates, namely acetaldehyde, propionaldehyde, butyraldehyde, isobutyraldehyde and formaldehyde with either NADP$^+$ or NAD$^+$ as cofactors. Structures of DmpG/DmpF have been solved by X-ray crystallography and this will be described in more detail in other sections.

BphI and BphJ, the aldolase and dehydrogenase from the bph pathway of the PCB degrading bacterium, *B. xenovorans* LB400, has also been characterized as follows (64). Attempts to separately express and purify the aldolase and dehydrogenase in recombinant *E. coli* have been unsuccessful, since BphI formed inclusion bodies and BphJ irreversibly precipitated out of solution during and after purification. However *bphI* and *bphJ* can be co-expressed in soluble form in *E. coli* using two compatible plasmids (pBTL-T7 and pET28a). The BphJ protein has an N-terminus poly-His tag encoded by the vector pET28a, which enabled co-purification of the complex by Ni-NTA chromatography. After purification the His-tag can be cleaved using the protease, thrombin.

The class-II pyruvate specific aldolase, BphI was determined to have highest activity using Mn$^{2+}$ as its metal cofactor (64). The kinetic parameters of BphI for various 4-hydroxy-2-oxoacids were determined and specificity constants were similar for 4-hydroxy-2-oxopentanoate and 4-hydroxy-2-oxohexanoate, which differ by one carbon atom in length. The pH dependency
of BphI was measured from pH 6.5 to pH 9.0, and the pKa of a catalytic base was found to be approximately 7.0. A stable pyruvate enolate analogue, oxalate, is a competitive inhibitor of the aldolase reaction with a $K_{ic}$ of $0.93 \pm 0.06 \mu$M \cite{64, 65}. Together the results are consistent with a base-catalyzed abstraction of a proton from the hydroxyl group at carbon 4 of the substrate followed by carbon-carbon bond cleavage to yield the pyruvate enolate (Figure 1.5). Further investigation revealed that BphI is specific for (S) configurations of 4-hydroxy-2-oxpentanoate; this preference can be removed with a substitution of Tyr-290 to phenylalanine \cite{65, 66}. Substrate specificity is governed in part by leucine 89, as substitution of this residue to the smaller alanine led to increased catalytic efficiency for longer aldehyde substrates in the aldol addition reaction \cite{65}. Further modification of the active site by replacing Tyr-290 with phenylalanine and leucine 89 with tryptophan or asparagine led to enzymes with a stereochemical preference opposite to the wild-type enzyme (i.e. from 4(S)-HOPA to 4(R)-HOPA \cite{67}).

A catalytic mechanism for the aldolase has been proposed based on inference from the structure of DmpG in addition to site directed mutagenesis and solvent isotope effect studies on BphI \cite{65} (Figure 1.5). Initially HOPA binds to the active site of BphI where it is stabilized by interactions with a bound Mn$^{2+}$ ion. His-20 then abstracts a proton from the C4 hydroxyl of HOPA, causing a carbon-carbon bond cleavage between the C3 and C4 of HOPA. The resulting pyruvate enolate is again stabilized by Arg-16 and the bound metal ion. The aldehyde is then released from the active site and a water molecule (Water 2) acts as a catalytic acid, donating a proton to the pyruvate enolate. The resulting pyruvate is released from the active site and the catalytic cycle repeats. The alternate hypothesis is that His-20 donates the proton abstracted from HOPA to Water 2, which in turn shuttle the proton to the pyruvate enolate intermediate.
Figure 1.5 Reaction mechanism of BphI.

HOPA binds to the active site of BphI where it is stabilized. Histidine 20 abstracts a proton from the C4 hydroxyl of HOPA, causing a carbon-carbon bond cleavage between the C3 and C4 of HOPA. The resulting pyruvate enolate is stabilized by Arg-16 and the bound metal ion. The aldehyde is then released from the active site and Water 2, stabilized by Tyr-290, binds. This water molecule acts as a catalytic acid, donating a proton to the pyruvate enolate. The resulting pyruvate is released from the active site and the catalytic cycle repeats.
The dehydrogenase, BphJ, a nonphosphorylating acylating aldehyde dehydrogenase, has also been characterized. The catalytic cysteine was identified as Cys-131 as replacement of this residue with alanine or serine abolished catalytic activity. Cys-131 was previously proposed to be activated by Asp-208 based on its proximity to the equivalent cysteine residue in the crystal structure of the homologous DmpF (50, 68). However, substitution of Asp-208 with alanine resulted in less than a two-fold reduction in catalytic efficiency and the pH dependency of the reaction was unaltered, discounting the proposed role of this residue in the catalytic mechanism of the enzyme (69). The substrate specificity of BphJ has also been determined. It is able to utilize aldehydes two to six carbons long as substrates and it has similar specificity constants for acetaldehyde and propionaldehyde (64). Due to sterical obstruction with the side chain of Ile-195, BphJ utilizes short chain aldehydes more efficiently than longer chain aldehydes. When Ile-195 is replaced with alanine the catalytic efficiency of the enzyme for longer aldehydes was increased while replacement with larger residues such as phenylalanine and tryptophan led to reduced specificity for these long aldehydes (69). Like DmpF, BphJ can use both NAD$^+$ and NADP$^+$ as a cofactor, although the $K_m$ is 16-fold lower for NAD$^+$ (64). In addition to using coenzyme A in the deacylation step, BphJ can also use dithiothreitol (DTT) but not free cysteine or glutathione in place of coenzyme A. A proposed reaction mechanism for BphJ is presented in Figure 1.6 (69).

MhpE and MhpF are the aldolase and dehydrogenase involved in degradation of 3-phenylpropionic acid in *Escherichia coli* (70). Recombinant MhpF can be overexpressed in soluble form in the absence of MhpE but MhpE formed inclusion bodies in the absence of MhpF. The authors proposed that MhpF acts as a chaperone for the proper folding of MhpE, although no additional evidence was provided to support this hypothesis.
Figure 1.6 Proposed reaction mechanism of the aldehyde dehydrogenase.

Nucleophilic attack by cysteine on the aldehyde results in a thiohemiacetal intermediate. A thioacylenzyme intermediate is formed after hydride transfer to the NAD$^+$. NADH then leaves the active site allowing for nucleophilic attack by the CoA, resulting in the formation of acyl-CoA.
In addition to being part of degradation pathways, aldolase-dehydrogenase complexes are also found in the nikkomycin biosynthetic pathway. Nikkomycin is an antifungal agent produced by *Streptomyces tendae* Tü901 (NikB-NikA) and *Streptomyces ansochromogenes* 7100 (SanM-SanN) (71, 72). Nikkomycins inhibit chitin synthesis but do not bioaccumulate and display no toxicity to animals, plants or fish. It is therefore a potential antifungal drug (73, 74). NikA or SanN catalyzes the conversion of picolinate-CoA to picolinaldehyde plus CoA and while NikB or SanM condenses picolinaldehyde and 2-oxobutyrate to 4-pyridyl-2-oxo-4-hydroxyisovalerate (75, 76). The physiological direction of these pathways is opposite to that of the aromatic degradation pathways of other homologues. SanN is active when expressed alone while SanM only displays activity in the presence of SanN (76). It is unclear if these enzymes channel the picolinaldehyde intermediate from one active site to another.

A phylogenetic analysis of sequences from various aldolase-dehydrogenase complexes show each complex falls into one of six clades (Figure 1.7). Both the aldolases and corresponding dehydrogenases from each organism appear to cluster similarly (77). Clades I-IV are part of aromatic degradation pathways, clade V contains thermophilic orthologs while members of clade VI are enzymes that are part of biosynthetic pathways.

### 1.7 General Structure of DmpG-DmpF

The crystal structure of DmpG-DmpF is available and the enzyme complex is a tetramer comprised of two DmpG-DmpF heterodimers (Figure 1.8). Two tetramers combine to make up the asymmetric unit of the crystal (50). The DmpG subunits of the biological tetramer contact each other to form the core while the DmpF subunits form the periphery of the tetramer.
Figure 1.7 Dendrogram of a phylogenetic analysis of aldolases (A) and dehydrogenases (B).

The aldolases and dehydrogenases cluster similarly; Clades I-IV are part of aromatic degradation pathways, clade V contains thermophilic orthologs while members of clade VI are enzymes which are part of biosynthetic pathways (77). (Image from Baker et al. 2012)
Figure 1.8 Crystal structure of DmpF-DmpG.
(A) DmpF-DmpG form a heterotetramer consisting of two aldolase protomers (blue) in the centre and two dehydrogenase protomers (green) on the periphery. (B) The individual domains of the aldolase and dehydrogenase. The TIM barrel domain and communication domain of the aldolase shown in blue and slate, respectively. The dehydrogenase consists of a NAD$^+$ binding domain (yellow) and dimerization domain (green and red). The variable length loop, which is not present in BphJ, is highlighted in red.
DmpG (aldolase) has two domains; an N-terminal \((\alpha\beta)_8\) TIM barrel domain and a C-terminal communication domain. The TIM barrel domain has two identical arrangements of \((\beta\alpha)_4\) half-barrels with the active site located at the end of the last \(\beta\)-strand. This structure resembles the HMG-CoA lyase and isopropylmalate synthase, considered to be part of the DRE-TIM metallolyase superfamily that catalyzes diverse carbon-carbon bond cleavage reactions \((78, 79)\). The C-terminal communication domain contains five \(\alpha\)-helices that have several interactions with DmpF. These helices also form part of what is thought to be a channel between the active sites of the DmpG and DmpF subunits. This channel is thought to play an important role in allowing the transit of acetaldehyde directly from the aldolase to the dehydrogenase as will be discussed in further detail below. The DmpG active site which is located at the C-terminus of the \((\alpha\beta)_8\) barrel has a bound metal ion thought to be \(\text{Mn}^{2+}\) since ICP-mass spectrometry detected significant quantities of \(\text{Mn}^{2+}\) in the purified enzyme complex. The \(\text{Mn}^{2+}\) is bound by His-200, His-202 and Asp-18 and to the substrate, pyruvate or the pyruvate enolate analogue, oxalate. This substrate or substrate analogue is also within hydrogen bonding distance of His-21, Tyr-291, Arg-17 and an oxygen atom of the main chain of Gly-52. These residues are conserved among aldolases including BphI and HsaF and are thought to be important for catalysis or substrate binding as described earlier.

DmpF (dehydrogenase) has two domains, an \(\text{NAD}^+\) binding domain (a Rossmann-like fold) and a dimerization domain. The overall structure appears to be similar to the glyceraldehyde-3-phosphate dehydrogenase enzyme family \((80, 81)\). Three of the four subunits of DmpF in the structure were in the holo form, containing bound \(\text{NAD}^+\), while the remaining one was in the apo form. When \(\text{NAD}^+\) is bound, the \(\text{NAD}^+\) binding domain moves toward the dimerization domain resulting in a closed conformation \((50)\). Part of this dimerization domain consists of a
communication loop that is absent in BphJ (Figure 1.8). There was no identifiable separate CoA binding site. Subsequent deuterium exchange experiments suggested the possibility of a shared NAD$^+/CoA$ binding site (82). Therefore the enzyme may have a ping-pong kinetic mechanism, whereby NAD$^+$ first binds the enzyme followed by acetaldehyde, then the NADH produced must leave the binding site before the CoA binds (82). The active site of DmpF lies between the dimerization domain and the NAD$^+$ binding domain. In the apo form, the proposed catalytic cysteine (Cys-132) interacts with a water molecule. This water molecule also forms a hydrogen bond to Asp-209. It was proposed that the Asp-209 activates the water molecule for a base attack on Cys-132, with Cys-132 then changing conformations so that the thiol group comes into close proximity of the NAD$^+$. The thiol group would then be in a position to undergo nucleophilic attack by the aldehyde. As mentioned earlier, it has recently been shown that a D208A (D209 in DmpF) variant in the homologous BphJ has similar steady state kinetic parameters and a similar pH kinetic profile as the wild-type enzyme indicating this residue is not involved in the catalytic mechanism as previously proposed (69).

An investigation of the solvent accessible surface of the DmpG-DmpF complex suggests existence of an intermolecular tunnel between the active sites of DmpG and DmpF. The residues of this 29Å long intermediate channel are mostly hydrophobic and unreactive. The channel appears to be blocked at either end when the enzyme is in its apo form, i.e. in the absence of NAD$^+$. The entrance to the channel from the aldolase is blocked by the tyrosine 291 from the aldolase active site and the exit of the channel to the active site of the dehydrogenase is blocked by the side chains of Ile-172, Ile-196 and Met-198. Tyr-291 was proposed to act as a catalytic acid and also to be gating the access to the tunnel. As Tyr-291 changes conformation to remove a proton from His-21 it is thought to also open ‘the gate’ to the channel so that the aldehyde can
then pass through. After the aldehyde has entered the tunnel, the torsion angle of Tyr-291 changes back to the closed conformation where it can protonate the enolate creating pyruvate. During the process the enolate intermediate is stabilized by the positively charged Arg-17 (Arg-16 in BphI) and Mn$^{2+}$ (50). More recent evidence however suggests that Tyr-291 (Tyr-290 in BphI) is not the catalytic acid and its role in gating remains unclear (65). When the aldehyde travels through the hydrophobic tunnel from DmpG to DmpF, it must enter into the aldehyde dehydrogenase active site through an exit lined by the side chains of Ile-172, Ile-196 and Met-198 which can adopt a variety of different conformations as seen in the crystal structure. Met-198 adopts different conformations in both the apo and holo forms whereas Ile-196 adopts different conformations in the holo form and Ile-172 only adopts different conformations in the apo form. With each of the apo and holo forms having two residues that have two conformations leads to four possible pair combinations. All four of the apo conformations lead to the channel still being blocked while one of the four of the holo forms appeared to open the channel. This open conformation is thought to create a unique conformation where Ile-172 appears to be affected by the interaction of NAD$^+$ with Asn-171. The binding of the cofactor and its interaction with Asn-171 is proposed to enable the open conformation to persist so the aldehyde can exit the channel and enter the active site of DmpF. Without cofactor binding and channel opening the aldehyde becomes trapped in the channel unable to reach the active site of the dehydrogenase. Further analysis of the tunnel connecting the active sites of DmpG and DmpF showed the tunnel is lined predominantly by residues of DmpG. There has not been any reported biochemical evidence that support aldehyde channeling in the DmpG-DmpF complex.
1.8 Substrate Channeling

Substrate channeling is the process by which an intermediate is channeled from one enzyme to another without export to the bulk solvent (83). There are several benefits of substrate channeling: (1) transit time between enzymes can be reduced thereby speeding up reactions (2) chemically unstable intermediates can be protected from the bulk solvent (3) toxic intermediates can be sequestered (4) hydrophobic molecules can be stopped from diffusing out of the cell (5) unfavorable equilibria can be eliminated and (6) intermediates can be kept safe from competing enzymes (83).

Systems that display substrate channeling can be characterized as having one of two general mechanisms, direct channeling or proximity channeling (84). Direct channeling occurs when the intermediate is transferred directly to the second enzyme without diffusion through bulk solvent. Tryptophan synthase, carbamoyl phosphate synthetase and proline/P5C dehydrogenase (PutA) are examples of enzyme systems that display direct channeling. Proximity channeling occurs when the first enzyme is located close to the second. The successive enzymes may be localized on a membrane, a solid surface or in aggregates. Once released from the first enzyme the intermediate may be utilized by the second enzyme without full equilibration with the bulk solvent. A special case of proximity channeling occurs when the intermediate is covalently linked to an enzyme’s prosthetic group that allows for direct transfer from one enzyme active site to another. This ‘swinging arm’ mechanism typically occurs in polyketide or non-ribosomal peptide synthetases where the intermediate is tethered to a long pantetheine prosthetic group or in pyruvate dehydrogenase complex where the intermediate is attached to a lipoamide (85).
1.8.1 Direct Substrate Channeling

One of the most well studied enzyme systems that exhibits direct substrate channeling is tryptophan synthase. This enzyme catalyzes the last two reactions in the synthesis of tryptophan. Indole 3-glycerol phosphate (IGP) is cleaved into indole and glyceraldehyde-3-phosphate (G3P) by the α-subunit and the β-subunit then condenses the indole and L-serine to form tryptophan. The enzyme is a heterotetramer composed of two α and two β subunits (86, 87). The crystal structure of tryptophan synthase from Salmonella typhimurium revealed that the subunits were arranged almost linearly in an α/β/β/α fashion that is approximately 150 Å long (88). Each α-subunit had the familiar (αβ)₈-fold TIM barrel motif. Like other TIM barrel proteins, the active site is located at the C-terminal of the β-barrel. The β-subunit has an N-terminal domain composed of four parallel β-sheets and a C-terminal domain composed of five parallel β-sheets and one antiparallel sheet. The β-subunit also contains a cofactor, pyridoxal phosphate (89) that is located at the interface of the two domains and is linked to Lys-87 (87). The X-ray crystal structure also revealed a tunnel that begins at the active site of the α-subunit and extends through the β-subunit to the center where the pyridoxal phosphate binds. This 25 Å long tunnel is hydrophobic and has the proper dimensions for the transition of an indole molecule (88, 90, 91). The tunnel can accommodate up to four indole molecules as determined from molecular modeling.

Biochemical evidence for substrate channeling in tryptophan synthase was shown by isotope dilution and transient kinetic analysis. When tritium labeled IGP was added to tryptophan synthase with exogenous unlabelled indole, significant radiolabelled tryptophan was produced indicating that the indole produced from the α reaction did not equilibrate with the bulk solvent (92). Transient kinetic analysis using a rapid quench method showed only 1% detectable free
radiolabelled indole when C-14 labeled IGP is mixed with tryptophan synthase on a millisecond time scale. Since the rate of condensation between serine and exogenous indole is slow (40 s$^{-1}$), indole must be channeled from the $\alpha$ to $\beta$ site or else significant accumulation of indole would be observed in the transient kinetic analysis (93). Attempts have been made to block the tunnel by substitution of cysteine 170 to tryptophan (C170W) and phenylalanine (C170F) (94). In single turnover experiments, indole accumulated in both mutant enzymes indicating that passage of indole through the tunnel was blocked or impeded. This is corroborated by the crystal structure of the C170W, mutant where the bulky tryptophan residue was seen to sterically block the tunnel (95).

The activities of the $\alpha$- and $\beta$-subunits in tryptophan synthase appeared to be allosterically regulated. Thus binding of serine to the $\beta$-subunit and its reaction with the pyridoxal phosphate cofactor to form aminoacrylate, enhances the cleavage of IGP at the $\alpha$-subunit by about 30-fold (96). This ensures that serine, the co-substrate of the reaction at the $\beta$ site, is available before indole is produced from the $\alpha$-subunit.

A second well characterized enzyme that exhibits substrate channeling is carbamoyl phosphate synthetase (CPS) which is responsible for the synthesis of pyrimidine nucleotides. This enzyme is heterodimeric containing one small (42 kDa) and one large subunit (118 kDa) as well as three distinct active sites (97). Two tunnels connect the three active sites which span 100 Å from one end of the protein to the other. The first tunnel is responsible for channeling ammonia produced from hydrolysis of glutamine to the N-terminus of the large subunit where it reacts with the carboxy phosphate intermediate to form carbamate. The carbamate is then channeled to the third active site in the C-terminal domain which catalyses the phosphorylation
of carbamate to produce carbamoyl phosphate. The transfer of ammonia through the tunnel has been shown to be a spontaneous and fast process that is ultimately driven by the reaction of ammonia with carboxy phosphate \((98, 99)\). Isotopic labeling studies showed that unlabelled ammonia produced from glutamine in the large subunit is not exchanged with labelled ammonia \((^{15}\text{N})\) in solution \((100)\). Six amino acid residues in the small subunit were replaced by site directed mutagenesis to disrupt the ammonia tunnel \((101)\). Variants with glycine 359 replaced with a bulky phenylalanine or tyrosine residue showed an uncoupling of hydrolysis of glutamine to carbamoyl phosphate formation and there is a lag in the time course for carbamoyl phosphate formation while no lag was observed for the wild-type enzyme. A “leaky” tunnel has also been created by replacements of three residues \((\text{P360A, H361A and R265A})\) \((97)\). In this triple mutant, glutamine hydrolysis is also uncoupled from carbamoyl phosphate formation with hydrolysis of two glutamine molecules required for formation of one molecule of carbamate. This indicates that only 50% of ammonia produced from glutamine hydrolysis is used for formation of carbamate. However, the structure of the mutant enzyme was not solved to ascertain whether the introduced mutations created a “hole” for ammonia to escape from the tunnel.

Proline/P5C dehydrogenase, known as the proline utilization A (PutA) protein, is a multifunctional protein which is involved in the catabolism of proline to glutamate \((102, 103)\). The two enzyme activities are performed by two separate domains, the proline dehydrogenase \((\text{PRODH})\) and Δ¹-pyrroline-5-carboxylate \((\text{P5C})\) dehydrogenase domain \((104)\). The PRODH domain is a FAD-dependent dehydrogenase which oxidizes proline to Δ¹-pyrroline-5-carboxylate. The P5C is then hydrolyzed to glutamate γ-semialdehyde nonenzymatically before the NAD⁺-dependent P5C dehydrogenase \((\text{P5CDH})\) domain oxidizes the intermediate to
glutamate (105, 106). The bifunctional enzyme is restricted to Gram-negative bacteria, as in Gram-positive bacteria and eukaryotes each enzyme activity is found on separate proteins (104).

The crystal structure of PutA from Bradyrhizobium japonicum shows a tunnel linking the two enzyme active sites (105). The active sites are connected by a 41 Å tunnel with a large middle chamber. The hydrophilic tunnel is comprised mainly of charged residues, which has been proposed to create an electrostatic field, altering the $pK_a$ of the pyrrolinium moiety, favouring the formation of glutamate $\gamma$-semialdehyde (105). There is an absence of a lag time between proline utilization and the production of NADH by the second reaction. When a variant lacking PRODH activity was mixed with a variant lacking P5CDH activity a lag phase between the proline utilization and NADH production was however observed (105). A similar trifunctional protein from E. coli and Salmonella typhimurium, also functions as a transcriptional repressor of the put operon (107, 108) and was shown to also display evidence of substrate channeling (109). Using $^{14}$C-labelled proline, it was shown that exogenous P5C cannot compete with endogenously produced P5C from proline.

A substrate has also been shown to channel along an “electrostatic highway” as in bifunctional thymidylate synthase-dihydrofolate reductase from Leishmania major. Thymidylate synthase converts deoxyuridine monophosphate (dUMP) and 5,10-methylenetetrahydrofolate (CH$_2$H$_4$folate) to thymidine monophosphate (dTMP) and dihydrofolate. Dihydrofolate reductase then uses NADPH to convert the dihydrofolate to tetrahydrofolate. Channeling of intermediate, dihydrofolate, was first proposed when no lag time was observed between the addition of CH$_2$H$_4$folate and dUMP and the production of NADP$^+$ (110). The structure of the channel formed by this single polypeptide is quite unlike that of other direct channeling enzymes (111). Rather than a tunnel connecting the active sites, negatively charged dihydrofolate passes along
the positively charged portion of the surface of the protein. Simulations have shown that increasing the ionic strength of the solvent or decreasing the strength of the negative charge of the intermediate leads to reduced channeling efficiency, providing further evidence of the electrostatic nature of this channel (112). The passage of the intermediate is fast (>1000 s⁻¹) as no accumulation is observed in single turnover experiments (113). The addition of a competing enzyme to a reaction mixture containing thymidylate synthase-dihydrofolate reductase does not significantly reduce the amount of product produced, suggesting it is protected from the bulk solvent (114). Like other enzymes that undergo substrate channeling, activation at one active site was observed when cofactors were bound at the other active site (113).

Enzymes that display direct substrate channeling are summarized in Table 1.1.

1.8.2 Substrate Channeling by Proximity

Bacterial microcompartments known as polyhedral bodies may increase the efficiency of enzyme catalyzed transformation of metabolites and also sequester toxic compounds (122). The most commonly known polyhedral body is the carboxysome, which is involved in carbon fixation in cyanobacteria (123). CO₂ and HCO₃⁻ are imported through the plasma membrane and HCO₃⁻ is in turn imported into the carboxysome. HCO₃⁻ is converted to CO₂ by carbonic anhydrase within the carboxysome (124). Ribulose bisphosphate carboxylase/oxygenase (RuBisCO) then fixes the CO₂ by producing two molecules of 3-phosphoglycerate from ribulose bisphosphate (125). RuBisCO is a relatively slow enzyme and its activity is dependent on the concentration of CO₂. Localization to the carboxysome may prevent RuBisCo from undergoing photorespiration with molecular oxygen (126). Localizing substrate and enzymes within the carboxysome acts to improve the efficiency of carbon fixation.
Table 1.1 Example enzyme systems that undergo substrate channeling

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Intermediate</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aldolase-Dehydrogenase Complexes</td>
<td>Aldehydes</td>
<td>(64, 69, 77, 115)</td>
</tr>
<tr>
<td>Tryptophan Synthase</td>
<td>Indole</td>
<td>(88, 90, 92, 94)</td>
</tr>
<tr>
<td>Carbamoyl Phosphate Synthetase (small subunit)</td>
<td>Ammonia</td>
<td>(97, 98, 116, 117)</td>
</tr>
<tr>
<td>Carbamoyl Phosphate Synthetase (large subunit)</td>
<td>Carbamate</td>
<td>(97, 98, 116, 117)</td>
</tr>
<tr>
<td>Glutamine 5'-phosphoribosylpyrophosphate Amidotransferase</td>
<td>Ammonia</td>
<td>(118)</td>
</tr>
<tr>
<td>Proline Utilization A</td>
<td>Glutamate</td>
<td>(104, 105, 109)</td>
</tr>
<tr>
<td></td>
<td>Semialdehyde&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Rat Peroxisomal Multifunctional Enzyme, type 1</td>
<td>Hydroxyacyl CoA</td>
<td>(119)</td>
</tr>
</tbody>
</table>
Thymidylate Synthase-dihydrofolate Reductase

Aspartokinase-homoserine Dehydrogenase $^b$

Bifunctional Dimethylglycine Oxidase $^b$

---

$^a$ This product is the intermediate that undergoes oxidation at the second active site. Δ1-pyrroline-5-carboxylate is released from the first active site and undergoes non-enzymatic hydrolysis before arrival at the second site.

$^b$ These bifunctional enzymes catalyze non-consecutive reactions.
The 1,2-propanediol utilization (*pdu*) operon from *Salmonella enterica* includes 21 genes responsible for the synthesis of a microcompartment which degrades 1,2-propanediol and sequesters the propionaldehyde intermediate (127, 128). These microcompartments only form when 1,2-propanediol is utilized as a carbon source, which may explain why they went unnoticed until relatively recently. The sequestered propionaldehyde is transformed to propionyl-CoA within the compartment’s lumen and is then exported to the cytoplasm of the cell where further degradation takes place (129). Strains which are unable to produce these microcompartments show an accumulation of toxic propionaldehyde, which inhibits growth (130). Phylogenetic evidence indicates that horizontal gene transfer of the *pdu* operon has occurred several times. The 21 gene operon responsible for microcompartment formation can be cloned into an *E. coli* host, which allows it to utilize propanediol and form a microcompartment. This provides evidence that once transferred the machinery can be utilized by a different organism (131).

Pyruvate Dehydrogenase Complex (PDC) is an example of an enzyme complex which utilizes ‘swinging arms’ to transfer intermediates from one enzyme active site to another (132). PDC converts pyruvate to acetyl-CoA and NADH in a process carried out by three separate enzymes: pyruvate decarboxylase, dihydrolipoyl acetyltransferase and dihydrolipoamide/lipoamide dehydrogenase, termed E1, E2 and E3 respectively (133). E2 composes the core of the complex with long fragments containing a lipoamide moiety which protrudes into the surrounding subunits of E1 and E3 which are tethered approximately 11 nm away from the core (134, 135). The entire complex is proposed to consist of 16 dimers of E1, 24 E2 monomers and 8 dimers of E3 (135, 136). E1 first decarboxylates pyruvate using thiamine diphosphate as a cofactor and the resulting acetyl group is transferred to the lipoamide moiety of E2 (137). The lipoamide moiety of E2 then transfers the acetyl group onto CoA, forming acetyl-CoA resulting
in a dihydrolipoamide group on E2 (138). E3, a flavoprotein, reoxidizes the dihydrolipoamide to lipoate passing the electrons to FAD; FADH$_2$ is then oxidized by NAD$^+$ back to FAD producing NADH. The ‘swinging arm’ is highly coordinated in order to protect the thioester intermediate and ensure efficient catalysis (139, 140).

1.9 Evidence of Substrate Channeling in the Aldolase-Dehydrogenase Complex

The possibility that the acetaldehyde product of aldolase, DmpG is channeled to the dehydrogenase, DmpF, has not been shown biochemically, but channeling has been recently demonstrated in the homologous BphI/BphJ complex using two methods (64). First the substrate of BphI, 4-hydroxy-2-oxopentanoate, was added to the BphI/BphJ complex and the BphJ activity was determined spectrophotometrically based on the reduction of NAD$^+$. No lag time was observed in the production of NADH. In addition, the activity of BphJ estimated from the progress curve is 10.4 s$^{-1}$, which is approximately 60% the $k_{cat}$ value for BphJ determined using exogenous acetaldehyde, even though the theoretical maximum concentration of acetaldehyde produced by aldol cleavage of 4-hydroxy-2-oxopentanoate in the assay was only 0.1 mM, 236-fold lower than the apparent $K_m$ of acetaldehyde for BphJ. In fact, the activity of BphJ corresponds to an acetaldehyde concentration 360-fold higher than expected from the maximum acetaldehyde produced from 0.1 mM 4-hydroxy-2-oxopentanoate. The second method for determining substrate channeling took advantage of the similar kinetic parameters for acetaldehyde and propionaldehyde in BphJ. Exogenous propionaldehyde (1 mM) was provided in 10-fold molar excess over 4-hydroxy-2-oxopentanoate in a BphI-BphJ coupled assay where CoA was limiting. If no substrate channeling occurred, the excess propionaldehyde would effectively compete with acetaldehyde released from BphI, resulting in predominantly propionyl CoA formation in the BphJ reaction. However, HPLC analysis of reaction products revealed
95% of the CoA esters produced were acetyl-CoA. Conversely, 99% of the CoA esters produced were propionyl-CoA when an excess of exogenous acetaldehyde was added in an enzyme reaction mixture containing 4-hydroxy-2-oxohexanoate. This demonstrates that propionaldehyde can also be directly channeled from the aldolase to the dehydrogenase.

The $K_m$ values for exogenous aldehydes in BphJ are $> 8$ mM, and coupled to the fact that aldehydes are volatile, reactive and toxic, channeling and sequestering of aldehydes from the cytoplasm of the bacterium may be necessary for efficient conversion to CoA esters and to prevent detrimental effects to the cell. Unique to the aldolase-dehydrogenase system of the aromatic degradation pathway is that the respective enzymes have broad substrate specificities and are capable of utilizing larger substrates (65, 66). While it has been demonstrated that acetaldehyde and propionaldehyde can be channeled by the enzyme complex, it was unclear if longer and branched aldehydes could also be channeled. This has some practical relevance since the meta-cleavage pathway for degradation of certain phenols with alkyl substitutions at the C4 position has been reported in bacteria, yielding aldehydes up to 7 carbon atoms in length (141, 142).

1.10 Justification of Research

There are few examples of substrate channeling in enzyme systems and for most of these systems the molecular details of how efficient channeling is achieved are not well understood. The objective of this research was to characterize channeling in aldolase-dehydrogenase systems, including BphI-BphJ and HsaF-HsaG complex. Both the aldolase and dehydrogenase will be investigated to determine their role in mediating substrate channeling.
Chapter 2 of this thesis details the materials and methods used in the experiments described in this thesis. Chapter 3 focuses on the development of an assay that can accurately quantify substrate channeling and determine the size of aldehyde which can be channeled in the aldolase-dehydrogenase complex. Chapter 4 focuses on a channeling mechanism and what roles the aldolase and dehydrogenase play in channeling and allosteric activation. Whether channeling of aldehydes occurs in chimeric enzyme complexes, made up of aldolase and dehydrogenase derived from different bacteria, was also investigated. The fifth chapter describes the biochemical characterization and crystal structure determination of HsaF-HsaG. Chapter 6 summarizes the pertinent results from this thesis and discuss possible future directions in this area.
Chapter 2: Materials and Methods

2.1 Chemicals and Substrates

2.1.1 Chemicals

Sodium pyruvate, sodium oxalate, NAD+, NADH, alcohol dehydrogenase, (Saccharomyces cervisiae) L-lactate dehydrogenase, (rabbit muscle), Chelex 100 and all aldehydes were from Sigma-Aldrich (Oakville, ON). Restriction enzymes were from Fermentas, T4 DNA ligase and Pfu Turbo polymerase were from Stratagene (Mississauga, ON) or New England Biolabs (Pickering, ON). Ni-NTA Superflow resin was obtained from Qiagen (Mississauga, ON). Aldehyde dehydrogenase from Saccharomyces cervisiae was purchased from Calzyme (San Luis Obispo, CA). All other chemicals were analytical grade and were obtained from Sigma-Aldrich or Fisher Scientific (Nepean, ON).

2.1.2 4-hydroxy-2-oxoacid Preparation

Aldolase substrates were synthesized by either HpaI (a class II aldolase purified as previous described (51)) or BphI. Pyruvate and aldehydes were used to generate various 4-hydroxy-2-oxoacids using the above enzymes as a biocatalyst. 2 M acetaldehyde, propionaldehyde, butyraldehyde, isobutyraldehyde, pentaldehyde and hexaldehyde were combined with 0.4 g pyruvate, 3 mg enzyme (0.5 mg for HpaI) and 1mM Mn²⁺ in a total of 15 mL of 100mM HEPES pH 8.0 to produce 4-hydroxy-2-oxopentanoate, 4-hydroxy-2-oxohexanoate, 4-hydroxy-2-oxoheptanoate, 4-hydroxy-5-methyl-2-oxohexanoate, 4-hydroxy-2-oxooctanoate and 4-hydroxy-2-oxononanoate respectively. The reaction was performed at 25°C for 4 hours with BphI and overnight with HpaI. Generating substrates with BphI generates enantiomerically pure compounds while HpaI generates racemates (66). The reaction was
stopped by addition of 0.2 g Chelex 100 to remove the Mn$^{2+}$. Excess aldehyde was partially removed by heating the reaction mixture to approximately 40°C and vapour were removed by aspiration. After the volatile aldehyde was removed, the enzyme was removed by passing the mixture through an Amicon YM10 filter (Millipore, Nepean, ON). The resultant substrate was then lyophilized and resuspended in 5 mL H$_2$O and acidified to pH 2; resulting in the lactonization of the 4-hydroxy-2-oxoacid. The solution was left overnight at room temperature and subjected to HPLC using an ÄKTA Explorer 100 (Amersham Pharmacia Biotech, Baie d’Urfé, QC) equipped with an Aminex® Fast-acid analysis ion exchange column (100 mm x 7.8 mm). Pure substrates were eluted with a mobile phase of 0.1% formic acid with a flow rate of 0.8 mL/min. Fractions containing substrate were pooled and mixed with 10 mL of 20 mM HEPES at pH 7.5 and neutralized with NaOH. The purified and neutralized substrate was then lyophilized and dissolved in H$_2$O prior to use.

2.2 DNA Manipulation and Cloning

2.2.1 Bacterial Strains and Plasmids

$bphI$, $bphJ$, $TTHB246$ and $TTHB247$ were previously cloned into either pBTL-T7 (64) or pET28a (EMD, Biosciences Inc., San Diego, CA) (64). pTIP QC1 was obtained from the Tamura group at the National Institute of Advanced Industrial Science and Technology (AIST) in Sapporo, Japan. *Rhodococcus jostii* RHA1 was obtained from the Eltis Lab (Department of Microbiology and Immunology, University of British Columbia, BC). pCR™-Blunt II-TOPO® vectors and *Escherichia coli* DH5α were purchased from Invitrogen (Oakville, ON). The *M. tuberculosis* H37Rv genomic DNA is a gift from Dr. Marcel Behr (McGill University, QC).
2.2.2 Cloning of *hsaF* and *hsaG*

The *hsaF* genes were amplified by PCR using the primers

CGCGCATATGCCTTCCAAGGCAAAAGTGGCG

and

CCCCAAGCTTCAATGCGTTACGGCACC. The *hsaF* gene was amplified by PCR using the following primers: GCCGCATATGACCGATATGTGGGACGTC;

CCCCAAGCTTCAATGCGTTACGGCACC and *hsaG* with the following primers:

CGCGCATATGCCTTCCAAGGCAAAAGTGGCG;

GCCCAAGCTTCAATCGCGCTCCTCCTCCTACAAC. Introduced *NdeI* and *HindIII* cleavage sites are underlined. The PCR mixture contained *Mycobacterium tuberculosis* H37Rv genomic DNA treated with *XbaI*, 1 unit of Pfu polymerase (Fermentas), 20 nM of each dNTP, 5% acetamide and 100 pmol of each primer in a final volume of 50 uL. The following touchdown amplification profile was performed: 95°C for 2 minutes, followed by 100 cycles of 95°C for 1 minute, annealing for 1 minute and extension at 72°C for 2 minutes. For every 10 cycles the annealing temperature was reduced by 2°C from an initial temperature of 70°C. The PCR protocol was terminated with a final extension of 72°C for 10 minutes. The amplified fragments were ligated into pCR™-Blunt II-TOPO® (Invitrogen) and transformed into *E. coli* DH5α. The resultant plasmid constructs were digested with *NdeI* and *HindIII* and ligated into pET28a for expression test in *E. coli*. The genes from the plasmid constructs were digested with *NcoI* and *HindIII* and inserted into similarly digested pTIP QC1(143) creating *hsaF* pTIP QC1-His, *hsaF* pTIP QC1-His and *hsaG* pTIP QC1-His. Putative clones were sent to Guelph Molecular Supercenter (University of Guelph) for confirmation of the sequence of the DNA fragment. Plasmids were then transformed into *Rhodococcus jostii* RHA1 according to the protocol described on Section 2.2.14 below.
2.2.3 Restriction Enzyme Digestions

DNA was digested with FastDigest® restriction enzymes from Fermentas. Each digestion contained 2 μL of purified plasmid/genomic DNA, 1× restriction enzyme buffer and 1 U of restriction enzyme in a total volume of 10 μL. Digestions were carried out at 37°C for 1 h.

2.2.4 DNA Ligation

DNA was ligated with 0.2 units of T4 DNA ligase (Invitrogen) in a final volume of 10 μL containing 1× ligase buffer. The ligation was carried out at 15°C for 12 h.

2.2.5 Agarose Gel Electrophoresis

DNA fragments were separated using a 0.9 % agarose gel. Fragment lengths were assessed by comparison to a standard (Fermentas). Samples were mixed with 1× loading buffer prior to loading and electrophoresis was carried out at 110V for 65 minutes in 1× TAE buffer. DNA was stained with ethidium bromide (1 μg/mL) for approximately 10 minutes and observed under UV light using a Geldoc system (Biorad Inc.). Agarose gels containing DNA fragments that are to be used for ligation were stained using GelGreen™ (Biotium Inc.) for 20 minutes and observed under blue light. Bands of interest were excised using a scalpel.

2.2.6 DNA Fragment Purification from Agarose

DNA was purified from agarose gels according to the manufacturer’s recommended protocol using a EZ-10 Gel Extraction spin column from BioBasic (Markham, ON). After purification DNA was eluted from the column using 15 μL of sterile water.

2.2.7 Plasmid Purification

Plasmids were purified using a EZ-10 Spin Column Plasmid Purification Kit from BioBasic (Markham, ON). 1.5 mL of bacterial culture was used when isolating pTIP QC1
plasmids and 4 mL culture was used when isolating all other plasmids. Manufacturer’s recommended procedures were followed for high copy plasmid (pTIP QC1) plasmids and low copy for all other plasmids. Purified plasmids were eluted with 40 μL water or elution buffer when using pTIPQC1 plasmids or 30 μL water for all other plasmids.

2.2.8 Site Directed Mutagenesis

Site-directed mutagenesis was performed according to the Quikchange™ (Stratagene) and modified Quikchange™ methods that use nonoverlapping primers (Table 2.1) (144). The plasmids encoding the following variants were created by previous lab members: bphI Y290F, bphI H20A, bphI L89A, bphJ C131A, bphJ C131S, bphJ I195W (65, 69). Other bphI and bphJ mutants were created using similar PCR protocols of 95°C for 5 minutes, followed by 12 cycles of 95°C for 30 seconds, annealing (T_m of primers minus 5°C) for 30 seconds and 72°C for 6 minutes. The hsaFG mutants were created with a touchdown PCR protocol: 95°C for 5 minutes, followed by 100 cycles of 95°C for 1 minute annealing for 1 minute and 72°C for 2 minutes. For every 10 cycles the annealing temperature was reduced by 2°C from an initial temperature of 70°C. The PCR protocol was terminated with a final extension of 72°C for 10 minutes. Recombinant DNA extracted from transformed E. coli DH5α was sequenced at the Guelph Molecular Supercenter to confirm the presence of desired mutations and the absence of secondary mutations. Primers used for the creation of variants are listed in Table 2.1.

2.2.9 Growth Conditions for Bacteria

E. coli was grown at 37°C on LB agar and in LB broth (1 % tryptone, 0.5 % yeast extract and 10 g/L NaCl). R. jostii RHA1 was grown at 30°C on LB agar and in LB broth. Liquid cultures were continuously shaken at 200 rpm. Where appropriate, antibiotics were added at the
Table 2.1 Primers for site-directed mutagenesis. Substituted codons are underlined.

<table>
<thead>
<tr>
<th>Mutation in BphI</th>
<th>Direction</th>
<th>Sequence 5’ to 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>G322A</td>
<td>Forward</td>
<td>GGCATGGTCGCCGCCAGGAAGACATGATTGAAGACACCCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CCTGGGCGCCGAGCATGCCGCCGGCGGCGCTCCAGTCGACCG</td>
</tr>
<tr>
<td>G322L</td>
<td>Forward</td>
<td>CATGGTCTTCCGCAGGAAGACATGATTGAAGACACCCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CTGGGCGCCGAGCATGCCGCCGGCGGCGCTCCAGTCGACCG</td>
</tr>
<tr>
<td>G322F</td>
<td>Forward</td>
<td>GGCATGGTCGCCGCCAGGAAGACATGATTGAAGACACCCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CCTGGGCGCCGAGCATGCCGCCGGCGGCGCTCCAGTCGACCG</td>
</tr>
<tr>
<td>G323A</td>
<td>Forward</td>
<td>GGCATGGTCGCCGCCAGGAAGACATGATTGAAGACACCCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CCTGGGCGCCGAGCATGCCGCCGGCGGCGCTCCAGTCGACCG</td>
</tr>
<tr>
<td>G323L</td>
<td>Forward</td>
<td>GTCGCGTCCGCGCGCCAGGAAGACATGATTGAAGACACCCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TTCCTGGAGGCGCCACCAGTCGCCGCGGCTCCAGTCGACCG</td>
</tr>
<tr>
<td>G323F</td>
<td>Forward</td>
<td>GTCGCGTCCGCGCGCCAGGAAGACATGATTGAAGACACCCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TTCCTGGAGGCGCCACCAGTCGCCGCGGCTCCAGTCGACCG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutation in BphJ</th>
<th>Direction</th>
<th>Sequence 5’ to 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>I195L</td>
<td>Forward</td>
<td>GGCAAGGCGCATCCTCATCCTCAACCACCCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CGGGTTGAGGAGGATGCGGCTTCGACCG</td>
</tr>
<tr>
<td>I195F</td>
<td>Forward</td>
<td>GGCAAGGCGCATCCTCATCCTCAACCACCCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CGGGTTGAGGAGGATGCGGCTTCGACCG</td>
</tr>
<tr>
<td>N170A</td>
<td>Forward</td>
<td>CGCGCAAGGCGCAATCCCGCAACCACCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CGTCGATGCGGGCGTTCCGGCGGAGCTTGACCG</td>
</tr>
<tr>
<td>N170D</td>
<td>Forward</td>
<td>CGCGCAAGGCGCAATCCCGCAACCACCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CGTCGATGCGGGCGTTCCGGCGGAGCTTGACCG</td>
</tr>
<tr>
<td>I171A</td>
<td>Forward</td>
<td>CAAGAACGCCGAGAATTCAACCACCCGACCACCCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GAATTCGTGAGCCGTTCCGCGGCTTCGACCG</td>
</tr>
<tr>
<td>I171F</td>
<td>Forward</td>
<td>CAAGAACGCCGAGAATTCAACCACCCGACCACCCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GAATTCGTGAGCCGTTCCGCGGCTTCGACCG</td>
</tr>
<tr>
<td>I195A</td>
<td>Forward</td>
<td>GGCATCGCAGCAGTGATGAGCTTCGAGCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GTTGAAGGTGCGATGCGGCTCGGCTCCAGTCG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutation in HsaF</th>
<th>Direction</th>
<th>Sequence 5’ to 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>G321F</td>
<td>Forward</td>
<td>CTCACTGTTCGCCAGGAGAAGACCAGCATCATCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CTGGCAGCAAGGATGCGGCTCCGAGCCG</td>
</tr>
<tr>
<td>G322F</td>
<td>Forward</td>
<td>CTCACTGTTCGCCAGGAGAAGACCAGCATCATCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CTTCCTGGAGGAGGATGCGGCTTCGAGCGGCG</td>
</tr>
<tr>
<td>R15A</td>
<td>Forward</td>
<td>GCTGGCAGGATGCGGCTCCAGCACCACAGGC</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CCATCGCAGCGAGGTGTCGAGTGCTGCAGTCG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutation in HsaG</th>
<th>Direction</th>
<th>Sequence 5’ to 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>C127S</td>
<td>Forward</td>
<td>GATCAACCTCGGGGGGAGAGGCGACCATTCCGATCGT</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TCCCCCGAGGTGATCATGTTGACGTTACGGCGGTCC</td>
</tr>
<tr>
<td>C127Q</td>
<td>Forward</td>
<td>GATCAACCTCGGGGGGAGAGGCGACCATTCCG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CCCCCCTGGGTGATCATGTTGACGTTACGGCGGTCC</td>
</tr>
<tr>
<td></td>
<td>Forward</td>
<td>Reverse</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>S41I</td>
<td>CCGGAGATCGATGGCCTGGCGCGCGGCCAAG</td>
<td>GCCATCGATCTCCCGGCTGGCGCACCATCCAGC</td>
</tr>
<tr>
<td>S41D</td>
<td>CCGAGGACGATGGCCTGGCGCGCGGCC</td>
<td>GCCATCGTCCTCCCGGCTGGCGCACCATCC</td>
</tr>
<tr>
<td></td>
<td><strong>Mutation in</strong></td>
<td><strong>Direction</strong></td>
</tr>
<tr>
<td></td>
<td><strong>TTHB246</strong></td>
<td><strong>Sequence 5’ to 3’</strong></td>
</tr>
<tr>
<td>A324G</td>
<td>Forward</td>
<td>Reverse</td>
</tr>
<tr>
<td></td>
<td>CGGTGGAGGGGCCAGGAGGACTGGATTCTG</td>
<td>CTGCCCTCCCCACCGCTGGCGGCGGCGCG</td>
</tr>
</tbody>
</table>
following concentrations, ampicillin 100 µg/ml, chloramphenicol 25 µg/ml, tetracycline 15 µg/ml, kanamycin 50 µg/ml.

2.2.10 Preparation of Electrocompetent *E. coli* BL21 (λDE3) Cells

Five mL of LB media was inoculated with a single colony of *E. coli* and grown overnight. 0.5 mL of overnight culture was added to 50 mL LB and grown to an optical density between 0.4 and 0.6 at 600 nm. Culture was incubated on ice for 30 minutes and centrifuged at 5000 × g for 15 minutes at 4°C. Supernatant was decanted; the culture was resuspended with 12.5 mL of 10% sterile glycerol and centrifuged at 5000 × g for 15 minutes at 4°C. This step was repeated to further wash the cells. The pellets were then resuspended in 100 µL of GYT media (10 % glycerol, 0.125 % yeast extract, 0.25 % tryptone) and frozen at -80°C in 25 µL aliquots.

2.2.11 Electroporation of *E. coli* BL21 (λDE3) Cells

Electrocompetent cells were thawed on ice. One µL of purified plasmid was added to the cells and the mixture was transferred to an electroporation cuvette (Fisherbrand 1 mm Gap). Cells were electroporated under the following conditions: 25 µF, 200 Ω and 1.25 volts. After electroporation 500 µL of SOC (2 % tryptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl$_2$, 10 mM MgSO$_4$, 20 mM glucose) was mixed with the cells and incubated at 37°C for 1 hour. After the recovery period, cells were plated on LB agar containing the appropriate antibiotics.

2.2.12 Preparation of Chemically Competent *E. coli* Cells

Five mL of LB media was inoculated with a single colony of *E. coli* and grown overnight. 0.5 mL of this overnight culture was then added to 50 mL LB and grown to an optical density of 0.4 at 600 nm. The culture was then incubated on ice for 10 minutes before
centrifugation at 5000 × g for 5 minutes at 4°C. Cells were decanted and resuspended in 10 mL of cold frozen storage buffer (FSB) containing 100 mM KCl, 50 mM CaCl₂, 10 mM potassium acetate and 10 % glycerol (w/v) and incubated for 20 minutes; this step was then repeated. After the second wash the cells were resuspended in 2 mL of FSB, aliquoted at 100 µL and stored at -80°C.

2.2.13 Transformation of Chemically Competent E. coli Cells

Chemically competent cells were thawed and incubated on ice for 30 minutes with DNA mixtures (plasmid or ligation). The mixture was then heat-shocked at 42°C for 45 seconds followed by a 2 minute incubation on ice. The mixture was then resuspended in 900 µL LB and incubated at 37 °C for 1 hour. The cells were then plated on the LB agar containing appropriate antibiotics for selection of recombinants and grown at 37°C.

2.2.14 Preparation of Electrocompetent R. jostii RHA1 Cells

A single colony of R. jostii RHA1 was inoculated into 50 mL of LB media and grown at 30°C for 72 hours. 2 mL of overnight culture was used to inoculate 200 mL LB broth and the culture was grown for approximately 18 to 22 hours until an optical density of 1.0 at 600 nm was reached. Cells were pelleted at 5000 × g for 5 minutes, decanted and resuspended in 10 mL ice cold 10% glycerol. The cells were again pelleted by centrifugation and then resuspended in 10 mL of 10% glycerol. The cells were then aliquoted into 500 µL preparations and stored at -80°C.

2.2.15 Electroporation of R. jostii RHA1 Cells

DNA was added to thawed competent R. jostii RHA1 cells. The mixture was introduced in a 1 mm electroporation cuvette (Fisherbrand). Cells were pulsed at 2.4 kV, 25 µF and 400 Ω. and then diluted with 500 µL LB broth and incubated for 4 hours at 30 °C. After incubation 50
µL was plated onto LB agar containing chloramphenicol colonies were allowed to grow for 3 to 4 days at 30 °C.

2.3 Protein Expression and Purification

2.3.1 Expression of BphI-BphJ in *E. coli* BL21 (λDE3)

Ten mL of recombinant bacterial culture was added to 1 L of LB media containing ampicillin and tetracycline and the culture grown at 37°C to an OD of 0.6 to 0.8 at 600 nm. Afterwards IPTG (isopropyl β-D-thiogalactopyranoside) was added to a final concentration of 1 mM to induce expression of *bphI* and *bphJ*. Cultures were then transferred to 15°C and allowed to grow for a further 20-24 h. Cells were harvested by centrifugation at 12096 × g for 10 minutes.

2.3.2 Expression of HsaF-HsaG in *E. coli* BL21 (λDE3)

The genes encoding HsaF and HsaG were inserted into pBTL-T7 and pET28a expression plasmids, respectively and transformed into *E. coli* BL21 (λDE3) for co-expression. Cultures were grown in LB broth or LB media was supplemented with 1 mM glycine betaine and 1.0 M sorbitol in an attempt to increase the solubility of the protein. Expression of the recombinant genes was induced at different growth temperatures (37 °C and 15 °C) and for different time periods. A (His)$_6$ tagged version of HsaF in pBTL-T7 was also co-expressed with HsaG in pET28a in an attempt to increase the solubility of HsaF.

In an attempt to further increase the solubility of the HsaF-HsaG complex recombinant *E. coli* were also transformed with plasmids encoding the following chaperones: GroES, GroEL and Tig (145).
2.3.3 Expression of HsaF-HsaG in *R. jostii* RHA1

*R. jostii* RHA1 containing the *hsaGF* pTIP QC1-His plasmid construct was propagated at 30°C in 4L Luria-Bertani medium supplemented with 25 µg/mL chloramphenicol. The optimal condition for expression of *hsaFG* was the use of 1 µg/mL of thiostrepton for induction of cell culture at an optical density of 0.6. Upon addition of the inducer, cells were grown at 30°C for 24 hours and harvested by centrifugation at 9605 × g for 8 minutes.

2.3.4 Small Scale Test of Protein Expression

Cells were harvested by centrifugation at 2300 × g for 10 minutes. A 500 µL culture was used for cells grown at 37°C and 1 mL was used for cells grown at 15°C. The cell pellet was resuspended in 50 µL of B-PER II (Pierce) solution. The mixture was vortexed for 2 minutes and the soluble and insoluble fractions were separated by centrifugation at 15800 × g for 5 minutes. The supernatant was transferred to a new tube and mixed with 2X SDS-PAGE loading buffer. The pellet was resuspended with 25 µL of sterile water and 25 µL of 2X SDS-PAGE loading buffer was added. The samples were then boiled for 2 minutes and analyzed by SDS-PAGE.

2.3.5 Purification of His-tagged Proteins

The harvested cell pellet were resuspended in 20 mM sodium HEPES. *E. coli* cells were lysed by passing them through a French press 3 times at 12000 psi and *R. jostii* RHA1 cells were lysed by passing them through a French press 5 times at a pressure of 25000 psi. *E. coli* lysate were centrifuged at 39191 × g for 30 minutes while *R. jostii* RHA1 lysates were centrifuged 3 times at 39191 g for 10 minutes each to remove cellular debris. The supernatant from the centrifugations were passed through a 0.45 µm filter and incubated with Ni²⁺-NTA resin (Qiagen Inc.) for one hour. The mixture was poured into a gravity column and washed with the 50 mM sodium phosphate buffer (pH 8.0), 300 mM sodium chloride (pH 8.0) and 20 mM imidazole. The
His-tagged proteins were eluted with buffer containing 150 mM imidazole. The buffer was changed by dilution in a stirred cell which contained a YM10 filter (Amicon) to 50 mM sodium phosphate pH 7.4. 1 unit of thrombin was added to approximately 1 mg of His-tagged protein with 10% glycerol for 90 minutes to cleave the poly-his tag. After cleavage the thrombin and poly-histidine tag were removed by addition of p-aminobenzamidine-agarose and Ni\textsuperscript{2+}-NTA resin consecutively. The buffer was then changed to 20 mM sodium HEPES (pH 8.5) by dilution in a stirred cell which contained a YM10 (Amicon) filter. Dithiothreitol (DTT) was added to a final concentration of 10 mM and aliquoted enzyme was stored at -80°C.

2.3.6 Assessment of Purity and Concentration of Enzymes

The purity of purified proteins was assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) with BenchMark Protein Ladder (Invitrogen) used as molecular weight markers (10-220 kDa). Gels were then stained with Coomassie Brilliant Blue (146). Protein concentration was assessed by the Bradford method (147).

2.3.7 Determination of the Native Molecular Mass of Enzymes

The native molecular mass of enzymes was calculated using a HiLoad 26/60 Superdex 200 column (GE Healthcare) calibrated with the following proteins: horse heart cytochrome c (12.4 kDa), bovine erythrocyte carbonic anhydrase (29 kDa), bovine serum albumin (66 kDa), yeast alcohol dehydrogenase (150 kDa), and sweet potato β-amylase (200 kDa). Blue dextran (2000 kDa) was used to determine the void volume. Proteins were eluted at 1.5 mL/minute in 20 mM HEPES (pH 7.5) containing 150 mM sodium chloride at 25°C. Size of proteins was estimated from a standard curve.
2.4 Enzyme Assays

2.4.1 Standard Aldolase Assay

Aldolase activity was determined continuously or discontinuously by coupling the production of pyruvate to lactate dehydrogenase (LDH) and monitoring the oxidation of NADH (64). All assays were performed at least in duplicate, at 25°C in a total volume of 1 mL using a Varian Cary 3 spectrophotometer with a temperature controlled cuvette holder. Standard continuous assays contained 0.4 mM NADH, 1 mM MnCl$_2$ and 19.2 Units LDH in 100 mM HEPES buffer (pH 8.0). The concentration of 4-hydroxy-2-oxoacids was varied from 0.1 $K_m$ to 10 $K_m$. Oxidation of NADH was monitored continuously at 340 nm with the extinction coefficient taken as 6200 M$^{-1}$·cm$^{-1}$. To assess the activity of aldolase in the absence of NADH, a discontinuous assay was utilized. Assay mixtures contained 1 mM MnCl$_2$, HOPA or HOHA concentrations varied from 0.1 $K_m$ to 10 $K_m$ and 20-80 μg of enzyme in 100 mM HEPES buffer (pH 8.0). After 10 minutes, the reaction was quenched with 20 mM Na$_2$EDTA (pH 8.5) and 0.4 mM NADH. The amount of pyruvate produced was measured by an end point assay using LDH. Fitting was carried out using nonlinear regression using GraphPad Prism.

2.4.2 Metal Specificity

The metal ion specificity of HsaF was determined by incubating 20 μg of HsaF-HsaG with Na$_2$EDTA for 20 minutes, then removing the EDTA by dilution with Chelex treated buffer. All assay components were treated with Chelex and a standard aldolase assay was performed with the addition of 0.1 mM and 1 mM concentrations of various metal chlorides and 2 mM of 4-hydroxy-2-oxohexanoate.
2.4.3 Aldehyde Dehydrogenase Activity Assays

Substrate specificity of the dehydrogenase was determined using 400 µM NAD⁺, 100 µM coenzyme A for BphJ and 400 µM NAD⁺, 1.5 mM coenzyme A for HsaG unless otherwise stated. Aldehyde concentrations were varied from at least 0.1 $K_m$ to 5 $K_m$. Concentrations of aldehyde stocks were determined by measuring the stoichiometric oxidation of NADH using alcohol dehydrogenase. Cofactor specificities were determined under similar conditions in which the acetaldehyde concentrations were held at 100 mM and concentrations of NAD⁺ and CoA were varied between 0.1 $K_m$ and 5 $K_m$, in the absence of DTT. All assays were performed at least in duplicate, at 25°C in a total volume of 1 mL using a Varian Cary 3 spectrophotometer with a temperature controlled cuvette holder. Production of NADH was monitored continuously at 340 nm with the extinction coefficient taken as 6200 M⁻¹·cm⁻¹. Fitting was carried out using nonlinear regression using GraphPad Prism. Values presented for $K_m$ and $k_{cat}$ are apparent as cofactor concentrations were kept constant.

2.4.4 Test of Substrate Channeling

The amount of aldehyde channeled directly from the aldolase to the dehydrogenase was assessed using an enzyme competition assay. All assays contained NAD⁺ at a concentration of at least 5 $K_m$, coenzyme A at a concentration of least 3 $K_m$, 1 mM MnCl₂, 5 µg BphI-BphJ (200 µg for the H20A and Y290F) or 10 µg of HsaF-HsaG and 20 U of Aldehyde Dehydrogenase from Saccharomyces cerevisiae (ALDH, converts aldehydes to acids and reduces NAD⁺ to NADH) in a total volume of 1 mL in 100 mM HEPES buffer pH 8.0. Reactions were initiated by addition of substrate and the amount of NADH produced was measured spectrophotometrically at 340 nm. HOPA and HOHA concentrations were 25 µM for wild type enzyme (both BphI-BphJ and HsaF-HsaG) and Gly-322, Gly-323 and L89A variants while 200 µM was used for the less active
variants, H20A and Y290F. 100 µM of substrate was used for all channeling experiments with TTHB247-TTHB247 complexes. Other 4-hydroxy-2-oxoacids concentrations were used at concentrations of 100 µM. The reaction was quenched after 5 minutes with 3 N HCl to a final concentration of 70 mM. The sample was centrifuged for 3 minutes at 21000 × g to pellet the denatured enzyme. A 500 µL sample of the reaction mixture was subjected to high performance liquid chromatography (HPLC) using an AKTA Explorer 100 (Amersham Pharmacia Biotech, Baie d’Urfe, QC) equipped with a HyPURITY C18 column (150 × 4.6 mm, Thermo Scientific). Samples were eluted with 50 mM sodium phosphate (pH 5.3)-acetonitrile mixtures at the following ratios: 47:3 for acetyl-CoA and propionyl-CoA, 125:16 for butyryl-CoA and isobutyryl-CoA, 25:4 for pentanoyl-CoA and hexanoyl-CoA. The flow rate was 1.0 mL per min and CoA-esters were detected by UV absorbance at 254 nm. Quantification of the acyl-CoA molecules produced was completed using standard curves of the pure compounds. Pentanoyl-CoA standards are not commercially available and were synthesized using BphJ, pentaldehyde, CoA and NAD\(^+\). The channeling efficiency was calculated by comparing the acyl-CoA produced (determined using HPLC) to the concentration of 4-hydroxy-2-oxoacid utilized by BphI based on the amount of NADH produced by BphJ and ALDH (Equation 1).

Channeling Efficiency = \( \frac{\text{Acyl-CoA produced}}{\text{Total amount of aldehyde produced from the aldol cleavage reaction}} \) \times 100 \% \quad \text{Equation 1}

2.4.5 Determination of NAD\(^+\) Dissociation Constants in Aldehyde Dehydrogenases

The dissociation constant of NAD\(^+\) was assessed by tryptophan fluorescence quenching upon NAD\(^+\) binding. Experiments were completed using a PTI QuantaMaster C-61 steady-state fluorometer (Photon Technology International, London, ON), with a 1 nm bandwidth for
excitation ($\lambda_{\text{ex}} = 290$ nm) and a 2 nm band width for emission ($\lambda_{\text{em}} = 330$ nm for BphI-BphJ, 333 nm for HsaF-HsaG). Titrations were completed in 500 µL of 20 mM HEPES pH 8.5 at 25°C with 100 µg of enzyme. NAD$^+$ was added stepwise to the enzyme mixture. The amount of NAD$^+$ utilized in the assay was below the concentration which would have lead to significant inner filter effects. Fluorescence intensity was corrected for dilution factors and the dissociation constant was calculated by fitting to equation 2, which describes binding to one site:

$$\left(\frac{\Delta F}{F_0}\times 100\right) = \frac{\Delta F_{\text{max}} \times 100 \times [S]}{I_0 \left( K_d + [S] \right)}$$

Equation 2

where ($\Delta F/F_0 \times 100$) is the percent fluorescence quenching (percent change in fluorescence relative to the initial value, $F_0$) following addition of substrate at a concentration $[S]$. Fitting was carried out using nonlinear regression using GraphPad Prism.

2.4.6 Activation of HsaF by HsaG

Activation of HsaF by HsaG was determined by titrating HsaG into a solution of HsaF. Each enzyme was purified separately and quantified. 1.15 nmol of HsaF (40 µL of 28.8 µM) was incubated with various quantities of HsaG in 134 µL of 20 mM HEPES pH 8.5 and 20 mM DTT for two hours. After incubation a standard aldolase assay was performed to assess the activity of HsaF. Data was fit to a sigmoidal binding model, equation 3.

$$Y = \frac{V_{\text{max}} \times X^h}{(K_{\text{prime}} + X^h)}$$

Equation 3

where $Y$ is the specific activity of HsaF, $V_{\text{max}}$= maximum specific activity, $X$ is the molar ratio of HsaG:HsaF, $h$ is the Hill slope, $K_{\text{prime}}$= related to $K_m$, same as $K_m$ if $h=1$
2.4.7 Oxaloacetate Decarboxylase Activity Assay

Oxaloacetate decarboxylase activity was determined continuously by coupling the production of pyruvate to LDH and monitoring the oxidation of NADH (5). All assays were performed at least in duplicate, at 25°C in a total volume of 1 mL using a Varian Cary 3 spectrophotometer with a temperature controlled cuvette holder. Standard continuous assays contained 0.4 mM NADH, 1 mM MnCl₂ and 19.2 Units LDH in 100 mM sodium HEPES buffer (pH 8.0). The concentration of oxaloacetate was varied from 0.1 $K_m$ to 10 $K_m$. Oxidation of NADH was monitored continuously at 340 nm with the extinction coefficient taken as 6200 M⁻¹•cm⁻¹.

2.5 Protein Crystallography and Molecular Modelling

2.5.1 Crystallization of HsaF-HsaG

HsaF-HsaG was concentrated to 2, 5 and 10 mg/mL using a stirred cell containing a YM10 filter and screens were set with the Classics Suite 1 kit (Qiagen) at 15°C and 25°C. Condition 88, containing 0.2 M MgCl₂, 0.1 M Tris-HCl pH 8.5, 30% (w/v) PEG 4000 yielded crystals and was further refined. The pH (6, 7, 8, 8.5 and 9), PEG 4000, (10- 25 %) metal concentration (10 mM to 200 mM), salt concentrations (0 mM to 200 mM) and protein concentration (1 mg/mL to 20 mg/mL) were all varied individually and in combination with other variables. Co-crystallization with substrate or substrate analogues, oxalate (10 mM), pyruvate (20 mM), NAD⁺ (20 mM) NADH (20 mM) and CoA (4 and 10 mM) were attempted. Glycerol (5- 15 %) was added to stabilize the protein.
2.5.2 Structure Solution of HsaF-HsaG

HsaF-HsaG crystals were grown using the sitting drop method at 25°C, using 2 µL of reservoir solution consisting of 18% PEG 4000, 0.2 M MgCl₂, 0.1 M Tris pH 8.5, 10% glycerol and 10 mM sodium oxalate, and 2 µL of 10 mg/mL of HsaF-HsaG. The crystals were soaked in the following solution: 13% PEG 4000, 0.2 M MgCl₂, 0.1 M Tris pH 8.5, 50% glycerol and 10 mM sodium oxalate for 5 minutes and flash frozen in liquid nitrogen. The data was collected to 1.93 Å at the Canadian Light Source, Canadian Macromolecular Crystallography Facility (08ID-1) and processed using X-ray Detection Software (XDS) (148, 149). Crystals were monoclinic with a C₂ spacegroup with dimensions of \(a = 69.69\,\text{Å},\ b = 142.69\,\text{Å}\) and \(c = 148.17\,\text{Å}\). Molecular replacement was completed by Phaser (150) and structure was determined using autobuild in Python-based Hierarchical ENvironment for Integrated Xtallography (PHENIX), searching with a model of HsaF-HsaG on the backbone of DmpF-DmpG (PDB Code: 1NV5) from *Pseudomonas putida* CF600 (151). The structure was rebuilt in COOT (152) and further refined with PHENIX refine (151). The crystal consists of a heterotetramer of two aldolases with residues 1-3 and 342 (343 in chain C)-346 missing and two dehydrogenases with residues 1,2 and 299-303 missing. Table 5.5 contains data collection and final refinement statistics. Figures were generated in PyMol (153). Further refinement of the model was performed by Matthew Kimber of the University of Guelph, Department of Molecular and Cellular Biology.
Chapter 3. Substrate Channeling in BphI-BphJ

Data from this chapter was published in the following journal article:


Statement of Contributions:

Only experiments that were directly contributed by Jason Carere from the above manuscript are included in this chapter. Perrin Baker assisted in creating Figure 3.2.
3.1 Introduction

Substrate channeling is the process by which the product of one enzymatic reaction is transferred directly to the active site of another enzyme, without export to the bulk solvent. Previously, transit time analysis revealed that the BphI-BphJ complex from the PCB degradation pathway of *B. xenovorans* LB400 channels the acetaldehyde from the aldolase to the dehydrogenase. Indeed the high $K_m$ value of BphJ for acetaldehyde (24 mM) would have resulted in a lag phase of over 50 hours if the aldehyde diffused from the aldolase to the dehydrogenase via bulk solvent.

A substrate competition assay was also developed that demonstrated that acetaldehyde and propionaldehyde produced from the BphI-catalyzed cleavage of 4-hydroxy-2-oxopentanoate (HOPA) and 4-hydroxy-2-oxohexanoate (HOHA), respectively, are not released to the bulk solvent but are channeled directly to BphJ (64). Thus, when excess propionaldehyde was added exogenously to a reaction mixture containing 4-hydroxy-2-oxopentanoate, 95% of the CoA esters produced were acetyl-CoA. Conversely, 99% of the CoA esters produced were propionylo-CoA when an excess of exogenous acetaldehyde was added in an enzyme reaction mixture containing 4-hydroxy-2-oxohexanoate. The $K_m$ values for exogenous aldehydes in BphJ are $>$8 mM. In addition, aldehydes are volatile, reactive, and toxic. Therefore channeling and sequestering of aldehydes from the cytoplasm of the bacterium may be necessary for efficient conversion to CoA esters and to prevent detrimental effects to the cell.

The crystal structure of orthologous aldolase-dehydrogenase complex (DmpG-DmpF; 56% and 55% sequence similarity to BphI and BphJ, respectively) from the phenol degradation pathway of *Pseudomonas putida* CF600 depicts a $\sim$21 Å long hydrophobic tunnel between the aldolase and the dehydrogenase active sites, providing a possible route for the direct diffusion of
aldehyde from the aldolase to the dehydrogenase (50). The presence of a conduit for the passage of common intermediates in sequential enzyme reactions has been observed in several other enzyme systems that exhibit substrate channeling, including tryptophan synthase, carbamoyl phosphate synthetase, and glucosamine 6-phosphate synthase (92, 94, 98, 116, 154, 155). Unique to the aldolase-dehydrogenase system of the aromatic degradation pathway is that the respective enzymes have broad substrate specificities and are capable of utilizing larger substrates. While it has been demonstrated that acetaldehyde and propionaldehyde can be channeled by the BphI-BphJ complex, it was unclear if longer and branched aldehydes could also be channeled. This has some practical relevance since the meta-cleavage pathway for degradation of certain phenols with alkyl substitutions at the C4 position has been reported in bacteria, yielding aldehydes up to 7 carbon atoms in length (141, 142).

This chapter describes the development of an assay to quantify the channeling efficiency of linear and branched chain aldehydes in the BphI-BphJ complex.
3.2 Analysis of DmpG-DmpF Crystal Structure

The crystal structure of the BphI-BphJ complex is not available, therefore structural analysis was performed on the crystal structure of the orthologous enzymes, DmpG-DmpF, from *P. putida* CF600 ((PDB 1NVM). The asymmetric unit in the crystal consists of two tetramers, with all aldolases (DmpG) containing bound Mn$^{2+}$ and oxalate, a pyruvate enolate analogue, and three of the four dehydrogenase subunits (DmpF) containing bound NAD$^+$ (Figure 3.1A). The biological unit of the complex is a tetramer consisting of two DmpG subunits in the middle and each of the two DmpF subunits associating with each DmpG subunit at the periphery to form an elongated structure (50). It has previously been determined by gel filtration that BphI-BphJ also forms a tetramer (64). The positions of the tunnel linking the bulk solvent to the active site and the channel linking the aldolase and dehydrogenase were mapped using the software MOLE using the crystal structure of DmpG-DmpF as template (Figure 3.1B). The diameter of the putative aldehyde channel ranges from approximately 2.4 to 5.4 Å (Figure 3.2). 4-hydroxy-2-oxoacid binding in the aldolase presumably requires a rotation about the C$\alpha$ of Tyr-291 to relieve a steric clash between the hydroxyl oxygen of Tyr-291 and C4 of the modelled substrate (65). Movement of Tyr-291 would result in an enlargement in the opening of the channel. Interestingly, when the aldehyde moiety of the 4-hydroxy-2-oxoacids is more than 4 carbons in length, there is no space available in the active site to accommodate the longer alkyl chain unless the distal end of the molecule protrudes into the aldehyde channel.
Figure 3.1 Structure of DmpG-DmpF and the channel.
Structure of DmpG-DmpF showing (A) the tetrameric arrangement of the biological unit with aldolases coloured blue and dehydrogenases coloured green and (B) the location of the tunnel from the aldolase active site to the bulk solvent (blue) and the channel connecting the active sites of the aldolase to the dehydrogenase (magenta). The location of the tunnel and channel were determined using MOLE (156). Images were generated using PyMOL (153).
Figure 3.2 Graphical representation of the diameter of the aldehyde channel as determined using the program MOLE (156).

The tunnel was mapped from the position of the divalent metal cofactor in the aldolase to the putative catalytic cysteine in the dehydrogenase. Positions in the graph corresponding to the aldolase and dehydrogenase active sites are depicted in blue and green, respectively. Computation of the structure when Tyr-291 is in the “closed” conformation as in the crystal structure is shown as a dotted line. The “open” conformation as predicted by the model is a solid line.
3.3 Determination of Aldehyde Channeling Efficiency in the BphI-BphJ Complex.

An enzyme competition assay was developed to obtain a quantitative measure of channeling efficiency. Excess exogenous aldehyde dehydrogenase from *Saccharomyces cerevisiae* was added to the assay so aldehyde released to the bulk solvent would be oxidized to a carboxylic acid. In contrast, aldehydes channeled to BphJ would be converted to acyl-CoA esters. The amount of acyl-CoA produced by the enzyme complex in relation to the total amount of aldehyde produced from the cleavage of 4-hydroxy-2-oxoacid by BphI provided a quantitative measure of channeling efficiency for each aldehyde as indicated by Equation 1 (Pg 49). CoA esters were separated from reaction mixtures by reversed phase high performance liquid chromatography using a C18 column and quantified using standards of known concentrations (Figure 3.3). Elution volumes are presented in Table 3.1.

Channeling efficiency is defined here as the fraction of aldehyde channeled from BphI which was utilized by BphJ to form CoA esters. Therefore it also depends on the degree of coupling between the reactions catalyzed by the aldolase and the dehydrogenase and represents a lower threshold for the true amount of aldehyde which traverses the channel.

Using this assay, channeling efficiency for acetaldehyde and propionaldehyde were determined to be 95 ± 5% and 95 ± 4% respectively. Longer straight chain aldehydes up to six carbons in lengths were also tested and found to be efficiently channeled by the enzyme complex (Table 3.2). Surprisingly, the branched chain aldehyde, isobutyraldehyde, was also channeled with an efficiency similar to that of acetaldehyde, the physiological substrate of BphJ.
Figure 3.3 Chromatogram of various acyl-CoA esters separated on a HyPURITY C18 (Thermo Scientific) HPLC column.

Samples were eluted with 50 mM sodium phosphate (pH 5.3):acetonitrile mixtures (47:3) at a flow rate of 1 mL/min. a) acetyl-CoA standard, b) BphI-BphJ assay mixture with 4-hydroxy-2-oxopentanoate as substrate c) propionyl-CoA standard, d) BphI-BphJ assay mixture with 4-hydroxy-2-oxohexanoate as substrate.
Table 3.1 Composition of solutions used to elute the various acyl-CoAs from the C18 column.

Solutions contained 50 mM sodium phosphate buffer (pH 5.3) and acetonitrile at different ratios (v/v) to separate the various acyl CoA esters. The elution volume of each acyl CoA is indicated in the table.

<table>
<thead>
<tr>
<th>CoA Ester</th>
<th>Carbon Atoms</th>
<th>Ratio (Buffer:Acetonitrile)</th>
<th>Elution volume (mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetyl-CoA</td>
<td>2</td>
<td>47:3</td>
<td>8.2</td>
</tr>
<tr>
<td>Propionyl-CoA</td>
<td>3</td>
<td>47:3</td>
<td>18.2</td>
</tr>
<tr>
<td>Butyryl-CoA</td>
<td>4</td>
<td>125:16</td>
<td>9.35</td>
</tr>
<tr>
<td>Isobutyryl-CoA</td>
<td>4</td>
<td>125:16</td>
<td>9.24</td>
</tr>
<tr>
<td>Pentanoyl-CoA</td>
<td>5</td>
<td>25:4</td>
<td>10.7</td>
</tr>
<tr>
<td>Hexanoyl-CoA</td>
<td>6</td>
<td>25:4</td>
<td>12.5</td>
</tr>
</tbody>
</table>
### Table 3.2 Channeling efficiency of various aldehydes in the BphI-BphJ enzyme complex

<table>
<thead>
<tr>
<th>Channeled Aldehyde</th>
<th>Structure</th>
<th>Channeling Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetaldehyde</td>
<td>CH$_3$CHO</td>
<td>95 ± 5</td>
</tr>
<tr>
<td>Propionaldehyde</td>
<td>CH$_3$CH$_2$CHO</td>
<td>95 ± 4</td>
</tr>
<tr>
<td>Butyraldehyde</td>
<td>CH$_3$(CH$_2$)$_2$CHO</td>
<td>83 ± 2</td>
</tr>
<tr>
<td>Isobutyraldehyde</td>
<td>(CH$_3$)$_2$CHCHO</td>
<td>92 ± 3</td>
</tr>
<tr>
<td>Pentaldehyde</td>
<td>CH$_3$(CH$_2$)$_3$CHO</td>
<td>94 ± 4</td>
</tr>
<tr>
<td>Hexaldehyde</td>
<td>CH$_3$(CH$_2$)$_4$CHO</td>
<td>84 ± 8</td>
</tr>
</tbody>
</table>

$^a$ Assays were performed at 25°C and contained 0.4 mM NAD$^+$, 0.1 mM Coenzyme A, 1 mM MnCl$_2$ and 5 µg BphI-BphJ in the presence of excess ALDH. Substrate concentration was 25 µM HOPA or HOHA and 100 µM for all other 4-hydroxy-2-oxoacids. CoA esters were loaded onto a C18 column and detected at 254 nm by HPLC.
3.4 Discussion

Substrate channeling involves the transfer of a common intermediate from one enzyme to another without release into bulk solvent (83, 157). In the case of the aldolase-dehydrogenase complex in aromatic degradation pathways, sequestering of volatile aldehydes from cellular components is beneficial because it prevents toxicity associated with reactive aldehydes and overcomes the inefficiency of the dehydrogenase for processing exogenous aldehydes due to high $K_m$ values.

Analysis of the DmpG-DmpF crystal structure showed that residues forming the channel that link the active sites of the aldolase and dehydrogenase are predominantly provided by the aldolase and are highly conserved in orthologous complexes. The molecular dimensions of the channel in the aldolase-dehydrogenase complex, with a narrowest diameter of 2.4 Å, appear to be incompatible for passage for the channeled aldehyde (4.3 Å). This suggests that the channel is not rigid but dynamic movement of residues lining the channel is required to allow the intermediate to pass through. While the low B-factors of the channel in the crystal structure of DmpG-DmpF can indicate rigidity, this may be due to crystal packing constraints. Alternatively, since the structure does not contain aldehyde in the channel, it does not represent the complex during substrate channeling. Protein dynamics have also been suggested to occur in other enzyme systems that exhibit substrate channeling, including glutamine-dependent NAD$^+$ synthetase (158) and the flavoenzyme PutA (105). In each case constrictions were observed in the structure of the channel and therefore significant movement of residues in the channels are required to enable the passage of the channeled intermediate.

It has been determined previously that BphI and BphJ have broad substrate specificity (64, 65), although preceding enzymes in the PCBs degradation pathway have not been shown to
utilize substrates that will lead to formation of compounds larger than 4-hydroxy-2-oxohexanoate. It is therefore surprising that aldehydes up to 6 carbons in length are channeled from the aldolase to the dehydrogenase with comparable channeling efficiency to that of acetaldehyde. The fact that the branched chain isobutyraldehyde can also be channeled by the complex indicates that these straight chain aldehydes need not necessarily adopt an extended conformation in order to transverse from the aldolase to the dehydrogenase. Flexibility of the tunnel of carbamoyl phosphate synthetase has also been suggested to account for the ability for the enzyme to channel the larger hydrazine and hydroxylamine in addition to its natural channeled substrate, ammonia (101), although the efficiency of channeling of these larger amines relative to ammonia has not been quantified.

The assay developed to quantify channeling is significantly different from previously described assays. Common assessments of channeling such as transit time analysis or isotope dilution do not account for the amount of intermediate that is actually transformed by the second enzyme. Transit time analysis determines the lag phase expected before the second enzyme reaches steady state after addition of the substrate for the first enzyme. If channeling does not occur, the time required for concentrations of intermediate to accumulate in the solvent can be calculated from the rate constants of the enzymes in the systems (159). Although a lack of lag phase suggests substrate channeling is occurring, it does not offer a quantitative measure of the amount of intermediate channeled. Isotope dilution uses radioactively labeled substrate which produces a radiolabelled intermediate and product. The bulk phase contains high concentrations of unlabelled intermediate; any radiolabelled intermediate released to the solvent will be out competed by the exogenous substrate (93). Comparing the ratio of labelled product to unlabelled product enables a qualitative assessment of channeling efficiency. This ratio does not account for
labelled intermediate which escapes to the bulk solvent. For example, if the $K_m$ of the second enzyme is high, little unlabelled product will be produced regardless of the channeling efficiency. In addition, the presence of exogenous intermediate may complicate analysis, especially if the second enzyme experiences substrate inhibition or regulation. The enzyme competition assay on the other hand provides a quantitative assessment of substrate channeling and the ability to detect small changes in channeling efficiency. The quantification of channeling is an important step in further analysis of the molecular mechanism of substrate channeling in the aldolase-dehydrogenase complex described in subsequent chapters.
Chapter 4: The Molecular Mechanism of Substrate Channeling in BphI-BphJ

Data from this chapter was published in the subsequent journal articles:


† denotes co-first authors


**Statement of Contributions:**

Only experiments that were directly contributed by Jason Carere from the above manuscripts are included in this chapter. Perrin Baker assisted with activation assays. C. Hillis created and purified the Thermus aldolase-dehydrogenase complex and chimeras.
4.1 Introduction

Aldolase-dehydrogenase complexes in bacterial aromatic meta-cleavage pathways are examples of bifunctional enzyme complexes that exhibit substrate channeling. As described in the previous chapter, BphI-BphJ from the PCB degradation pathway in B. xenovorans LB400 can channel aldehyde intermediates, up to six carbons in length, from the active site of BphI to BphJ with an efficiency of >80%. The aldol cleavage reaction performed by BphI is also allosterically activated 15-fold by the dehydrogenase, BphJ, when it undergoes turnover and approximately 5-fold in the presence of the dehydrogenase cofactor, NADH.

In this chapter, using a combination of molecular modeling and site-specific mutagenesis, the contributions of residues in the enzyme complex that govern channeling specificity and efficiency were analyzed. The creation of channel blocked enzyme variants enabled us to test if the observed 5-fold activation of BphI by BphJ when nicotinamide cofactor is bound is due to faster aldehyde release by substrate channeling.
4.2 Analysis of the DmpG-DmpF Crystal Structure

A narrow tunnel connects the active site of the aldolase, DmpG, to the bulk solvent. This tunnel is presumably the entrance for 4-hydroxy-2-oxoacid substrates and the exit for the pyruvate product (Figure 4.1). The narrowest point of the tunnel connecting the active sites of DmpG-DmpF (1.6 Å in diameter) is insufficient in size to allow for passage of 4-hydroxy-2-oxoacids, and therefore rotation of the proposed catalytic base, His-21, is assumed to be required to allow for access of substrates. Upon substrate entry, His-21 is proposed to revert back to its conformation seen in the “closed” position in order to be of optimal distance and geometry relative to the C4-OH of the substrate to be effective as a catalytic base. Substrate binding is presumably also coordinated by a rotation about the Ca of Tyr-291 to relieve a steric clash between the hydroxyl oxygen of Tyr-291 and C4 of the modeled substrate. Movement of Tyr-291 would result in an enlargement in the opening of the channel. Interestingly, when the aldehyde moiety of the 4-hydroxy-2-oxoacids is more than four carbons in length, there is no space available in the active site to accommodate the longer alkyl chain unless the distal end of the molecule protrudes into the aldehyde channel (Figure 4.1).

The channel connecting the active sites of DmpG and DmpF is contributed almost entirely by the aldolase (Figure 4.2). In DmpG, the diameter of the aldehyde channel ranges between approximately 2.4 and 5.4 Å (Figure 3.2). The narrowest point (diameter ~2.4 Å) is lined by residues Gly-323 and Gly-324 on one side and the Cδ1 of Leu-90 on the opposite wall of the tunnel formed by the aldolase, DmpG. Together with the side chain of Tyr-291, the Cδ1 of Leu-90 creates the entrance of the aldehyde channel in the aldolase while the Cδ2 of Leu-90 forms one boundary of the active site of the aldolase (Figure 4.1). Interestingly, in the homologous aldolase (TTHB246) of Thermus thermophilus HB8, the residue corresponding to
Figure 4.1 Model of 4-hydroxy-2-oxononanoate bound in the DmpG active site.

The model was generated by superimposing the pyruvyl moiety of substrate shown in sticks within the active site on the experimentally observed oxalate. The tunnel leading from bulk solvent is shown in blue and the active site and channel connecting the aldolase active site to the dehydrogenase is shown in magenta. His-21 blocks the tunnel leading from the bulk solvent. Residues investigated in this study are shown in sticks. Carbon atoms of Tyr-291 in the crystal structure are shown in yellow lines and the residue rotated 14° clockwise about the Cα to avoid steric clash with the C4 of the modelled substrate is shown in grey as sticks. Leu-90, Gly-323 and Gly-324 line the narrowest point of the tunnel. Corresponding residue numbers in BphI are indicated in brackets. The distal end of 4-hydroxy-2-oxononanoate is too long to fit in the active site and is proposed here to protrude into the channel.
**Figure 4.2 Sequence alignment of BphI and orthologs.**

Amino acid sequences of proteins and their uniprot accession number are as follows: BphI from *B. xenovorans* (P51015), CumF from *Pseudomonas fluorescens* IP01 (P97092), NahM from *Pseudomonas putida* (P51017), DmpG from *Pseudomonas putida* CF600 (P51016), XylK from *Pseudomonas putida* (P51019), MhpE from *Escherichia coli* (P51020), TodH from *Pseudomonas putida* F1 (P51018), BphF from *Pseudomonas sp.* strain KKS102 (P51014), CmtG from *Pseudomonas putida* F1 (Q51983), PhnJ from *Pseudomonas sp.* Strain DJ77 (Q9Z3U6), TTHB246 from *Thermus thermophiles* HB8 (Q53WI0) and HsaF from *Mycobacterium tuberculosis* HRv (P71867). Secondary structures of DmpG are indicated as horizontal arrows (beta-strands) and helices (alpha-helix) above the sequences. Residues that line the aldehyde channel are highlighted in red. Residues in BphI that were replaced by site-directed mutagenesis in this study are indicated by stars at the bottom of the alignment. This alignment was generated using ClustalX (160) and visualized using ESPript (161).
Gly-323 in DmpG (Gly-322 in BphI) is an alanine (Figure 4.2).

The exit of the channel, on the other hand, is formed by three residues of the dehydrogenase DmpF: Ile-172, Ile-196, and Met-198 (Ile-171, Ile-195, and Leu-197, respectively, in BphJ). These residues have been proposed to act as a gate to control the passage of aldehydes that are channeled from the aldolase to the dehydrogenase. This is partly supported by alteration of the rotameric configuration of Ile-172 and Ile-196 between the NAD$^+$-bound and NAD$^+$-free structures mediated by Asn-171 (Asn-170 in BphJ) in DmpF. Asn-171 undergoes a 166° rotation in the presence of NAD$^+$ relative to the apo structure, allowing the carboxamide of Asn-171 to form hydrogen bonds with the ribosyl hydroxyls of NAD$^+$. This movement results in changes to the Ca backbone, which subsequently affects the rotameric configuration of Ile-172 and Ile-196 (Figure 4.3A). In the apo form, Ile-172 is seen in multiple conformations, one of which is more predominant as visualized in the electron density, obtained from the Electron Density Server (Figure 4.3B) (160). The distance between the C$\delta$ atom of Ile-172 and the C$\gamma$2 atom of Ile-196 in the apo structure (observed in two conformations) ranges from 3.4 to 3.7 Å.

Several variants of the BphI-BphJ complex were created to probe the contributions of specific residues in aldehyde channeling. Gly-322 and Gly-323 (Gly-323 and Gly-324 of DmpG) were replaced with bulkier residues alanine, leucine, and phenylalanine. Leu-89 of BphI (Leu-90 in DmpG) was replaced with a smaller alanine residue. The contribution of the proposed gating residues in aldehyde channeling was tested by replacement of His-20 of BphI (His-21 of DmpG) with alanine and Tyr-290 (Tyr-291 in DmpG) with phenylalanine. Ile-195 of BphJ (Ile-196 in DmpF) that forms the exit of the channel was replaced with an alanine (I195A) residue, the isosteric leucine (I195L) residue and phenylalanine (I195F) and tryptophan (I195W) residues. Ile-171 was replaced with alanine or phenylalanine (I171A and I171F, respectively). Asn-170
Figure 4.3 Structural changes upon NAD$^+$ binding in DmpF

(A) Interaction of NAD$^+$ with Asn-171 of DmpF. In the presence of NAD$^+$, Asn-171 rotates 166° and forms hydrogen bonds with the hydroxyls of the nicotinamide ribose ring. This causes movement of the adjacent isoleucine (Ile-172), one of three residues that line the exit of the aldehyde channel leading to the dehydrogenase active site. Carbon atoms in the apo structure are shown in grey and the NAD$^+$-bound structure is depicted in cartoon with carbon atoms shown in blue. (B) Electron density of residues in apo structure showing multiple conformations of Ile-172 and Met-198.
was replaced with alanine and aspartate (N170A and N170D, respectively).

4.3 Site-specific Mutagenesis, Protein Purification and Steady-State Kinetic Analysis.

All variants were successfully created by site specific mutagenesis and purified using the same conditions (64) as wild type enzyme with similar yields and purities achieved (Figure 4.4). Aldolase kinetic parameters of Y290F and H20A variants were previously reported (65) and their catalytic efficiencies were ~3-fold and ~100-fold lower, respectively, relative to the wild-type aldolase. The steady state kinetics for the aldol reaction of the other aldolase variants were determined and the kinetic constants were of the same order of magnitude as the wild-type enzyme (Table 4.1).

4.4 Roles of Internal Channel Residues

4.4.1 Gly-322, Gly-323 and Leu-89 in BphI Channeling Efficiency and Specificity.

Variants G322A and G323A were able to channel acetaldehyde with similar efficiencies to wild-type (Table 4.2). However, channeling efficiencies progressively decreased with increasing aldehyde chain length. The G322A variant was able to channel isobutyraldehyde, albeit at a 30% lower efficiency than the wild-type enzyme, but G323A variant was unable to channel this branched chain aldehyde. Channeling efficiency for acetaldehyde was lower in the G323L variant (63% efficiency) and this variant was unable to channel the larger propionaldehyde. The G322L, G322F and G323F were all unable to channel either acetaldehyde or propionaldehyde. The kinetic parameters of the associated dehydrogenases in these variants were similar to those of the wild-type enzyme, thus the reduced channeling efficiencies were not due to the compromised activity of BphJ (Table 4.3). Creating a channel with a larger diameter
Figure 4.4. SDS-PAGE gel of purified BphI-BphJ wild-type and variants.

Lane 1 contains the molecular weight markers, lane 2 contains wild-type BphI-BphJ, lane 3 I195F variant, lane 4 I195L variant, lane 5 G322F variant, lane 6 G322L variant, lane 7 G323F variant and lane 8 G323L variant.
**Table 4.1** Steady-state kinetic constants of BphI and variants for the aldol cleavage of HOPA and HOHA \(^a\)

<table>
<thead>
<tr>
<th>Substrate</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Enzyme</strong></td>
<td><strong>(K_m) (µM)</strong></td>
<td><strong>(k_{cat} (s^{-1}))</strong></td>
<td><strong>(K_m) (µM)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>HOPA</strong></td>
<td><strong>HOHA</strong></td>
<td><strong>HOPA</strong></td>
<td><strong>HOHA</strong></td>
</tr>
<tr>
<td>WT</td>
<td>89 ± 8</td>
<td>4.07 ± 0.07</td>
<td>117 ± 10</td>
<td>3.9 ± 0.2</td>
</tr>
<tr>
<td>G322L</td>
<td>140 ± 10</td>
<td>1.09 ± 0.05</td>
<td>350 ± 20</td>
<td>0.94 ± 0.02</td>
</tr>
<tr>
<td>G323L</td>
<td>160 ± 20</td>
<td>0.64 ± 0.03</td>
<td>370 ± 70</td>
<td>0.27 ± 0.02</td>
</tr>
<tr>
<td>G322F</td>
<td>90 ± 10</td>
<td>1.06 ± 0.04</td>
<td>300 ± 20</td>
<td>1.05 ± 0.03</td>
</tr>
<tr>
<td>G323F</td>
<td>940 ± 80</td>
<td>4.3 ± 0.2</td>
<td>1100 ± 200</td>
<td>1.3 ± 0.1</td>
</tr>
<tr>
<td>G322A</td>
<td>180 ± 20</td>
<td>2.4 ± 0.1</td>
<td>520 ± 90</td>
<td>1.9 ± 0.1</td>
</tr>
<tr>
<td>G323A</td>
<td>60 ± 10</td>
<td>0.34 ± 0.02</td>
<td>360 ± 30</td>
<td>0.46 ± 0.07</td>
</tr>
</tbody>
</table>

\(^a\) Aldolase assays were performed at 25°C and contained 0.4 mM NADH, 1 mM MnCl\(_2\) and 19.2 Units LDH in 100 mM HEPES buffer (pH 8.0)
### Table 4.2 Aldehyde channeling efficiencies of BphI variants-BphJ complexes

<table>
<thead>
<tr>
<th>BphI variant</th>
<th>Acetaldehyde</th>
<th>Propionaldehyde</th>
<th>Butyraldehyde</th>
<th>Isobutyraldehyde</th>
<th>Pentaldehyde</th>
<th>Hexaldehyde</th>
</tr>
</thead>
<tbody>
<tr>
<td>G322A</td>
<td>89 ± 1</td>
<td>62 ± 7</td>
<td>74 ± 2</td>
<td>60 ± 3</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>G323A</td>
<td>95 ± 10</td>
<td>58 ± 1</td>
<td>29 ± 3</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>G322L</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>G323L</td>
<td>62 ± 1</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>G322F</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>G323F</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>L89A</td>
<td>67 ± 3</td>
<td>68 ± 1</td>
<td>79 ± 2</td>
<td>55 ±2</td>
<td>67 ± 6</td>
<td>17 ± 1</td>
</tr>
</tbody>
</table>

ND- Not Determined
by replacing Leu-89, located on the opposite side of Gly-322 and Gly-323, to alanine led to a surprising reduction in channeling efficiency of 30% for all aldehydes tested. The reduction of channeling efficiency in the variant could be due to the increased size of the channel entrance, subverting the gating mechanism, leading to escape of the aldehyde from the substrate entry/pyruvate exit tunnel.

4.4.2 Substrate Channeling in Thermus thermophilus Aldolase-Dehydrogenase Complex and Chimeric Aldolase-Dehydrogenase Complexes

Sequence alignment revealed that the homologous aldolase from T. thermophilus HB8 (TTHB246) has an alanine residue (Ala-324) at the equivalent position of Gly-322 of BphI. The Thermus enzyme complex was purified and its aldehyde channeling efficiency compared to that of BphI-BphJ. TTHB246-TTHB247 was observed to channel acetaldehyde, the intermediate produced from the cleavage of HOPA, with efficiency similar to that of BphI-BphJ (94 ± 1%). However, channeling efficiency for propionaldehyde was ~57% (Table 4.4). The channeling efficiency for propionaldehyde was similar to that of G322A variant in BphI. Site-specific mutagenesis was used to replace the alanine residue of TTHB246 (A324) with glycine. The resulting variant had a channeling efficiency of 94 ± 1 % for propionaldehyde, thus confirming the importance of glycine at this position for efficient propionaldehyde channeling.

In order to determine if efficient substrate channeling requires partner enzymes from the same species, chimeric aldolase-dehydrogenase complexes made up of wild-type aldolase and dehydrogenase from the B. xenovorans and T. thermophilus enzymes (TTHB-246-BphJ and BphI-TTHB247 complexes) were expressed in E. coli and purified. These chimeric complexes
Table 4.3 Kinetic constants for acetaldehyde and propionaldehyde of wild-type BphJ associated with the BphI glycine variants $^a$

<table>
<thead>
<tr>
<th>Associated Aldolase</th>
<th>Acetaldehyde</th>
<th>Propionaldehyde</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_{m,\text{app}}$ (mM)</td>
<td>$k_{cat,\text{app}}$ (s$^{-1}$)</td>
</tr>
<tr>
<td>WT</td>
<td>24 ± 2</td>
<td>17 ± 1</td>
</tr>
<tr>
<td>G322L</td>
<td>12 ± 2</td>
<td>19 ± 1</td>
</tr>
<tr>
<td>G323L</td>
<td>6.1 ± 0.5</td>
<td>6.3 ± 0.2</td>
</tr>
<tr>
<td>G322F</td>
<td>12 ± 1</td>
<td>7.5 ± 0.2</td>
</tr>
<tr>
<td>G323F</td>
<td>17 ± 3</td>
<td>27 ± 2</td>
</tr>
</tbody>
</table>

$^a$ Dehydrogenase assays were performed at 25°C and contained 0.4 mM NAD$^+$ and 0.1 mM coenzyme A in 100 mM HEPES buffer (pH 8.0).
Table 4.4 Substrate channeling efficiency in enzyme complexes $^a$

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Channeling Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Substrate</td>
</tr>
<tr>
<td>TTHB246-TTHB247</td>
<td>94 ± 1</td>
</tr>
<tr>
<td>BphI-TTHB247</td>
<td>9 ± 2</td>
</tr>
<tr>
<td>TTHB246-BphJ</td>
<td>12 ± 1</td>
</tr>
<tr>
<td>TTHB246–TTHB247 (A324G)</td>
<td>94 ± 4</td>
</tr>
</tbody>
</table>

$^a$ Assays were performed at 25°C and contained 0.4 mM NAD$^+$, 0.1 mM Coenzyme A, 1 mM MnCl$_2$ and 10 µg of enzyme in the presence of excess ALDH. Substrate concentration was 25 µM (S)-HOPA or HOHA. NADH production was measured spectrophotometrically at 340 nm and CoA esters were detected at 254 nm using HPLC.
were found to exhibit low but detectable substrate channeling of ~10 % for acetaldehyde.

4.5 Role of Tyr-290 and His-20 of BphI in Aldehyde Channeling.

Substitutions of the Tyr-290 in BphI, with the smaller phenylalanine residue reduced channeling efficiencies by >30% (Table 4.5). Replacement of His-20, a residue that gates the substrate entry tunnel, with alanine dramatically reduced acetaldehyde and propionaldehyde channeling by more than 70%.

4.6 Role of Residues in the Dehydrogenase in Substrate Channeling

Kinetic parameters for acetaldehyde and NAD$^+$ were determined for BphJ variants and reported in Table 4.6 and Table 4.7. Acetaldehyde channeling efficiencies in N170A and N170D were 83 ± 2 % (Table 4.8). Substitution of Ile-195 in BphJ to a bulkier phenylalanine residue reduced the channeling efficiency of the enzyme complex towards acetaldehyde by only 15%. The I195A variant exhibited similar channeling efficiency to other variants tested (~80%) while I195W displayed a more dramatic decrease in acetaldehyde channeling efficiency (59 ± 1%). Substitution of Ile-171 with either alanine or phenylalanine resulted in 20-30% decrease in acetaldehyde channeling efficiency.

4.7 Allosteric Activation of the Aldolase by the Dehydrogenase

In the wild-type aldolase-dehydrogenase complex, aldol cleavage by BphI was activated 5-fold in the presence of the BphJ cofactor, NADH (64, 115). When both enzymes are undergoing turnover, this activation was increased to 16-fold over the non-activated aldolase. Inspection of the crystal structure of the DmpG-DmpF complex revealed that binding of nicotinamide induces the opening of the exit of the channel in the dehydrogenase, thus facilitating aldehyde channeling. To test if this activation is due to allosteric regulation of BphI
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Acetaldehyde</th>
<th>Propionaldehyde</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y290F (BphI)</td>
<td>57 ± 2</td>
<td>38 ± 1</td>
</tr>
<tr>
<td>H20A (BphI)</td>
<td>11 ± 3</td>
<td>15 ± 1</td>
</tr>
</tbody>
</table>

*Assays were performed at 25°C and contained 0.4 mM NAD+, 0.1 mM Coenzyme A, 1 mM MnCl₂ and 10 µg of enzyme in the presence of excess ALDH. Substrate concentration was 25 µM (S)-HOPA or HOHA. NADH production was measured spectrophotometrically at 340 nm and CoA esters were detected at 254 nm using HPLC.*
Table 4.6 Steady-state kinetic parameters of acetaldehyde for various variants $^a$

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$K_{m,\text{app}}$ (mM)</th>
<th>$k_{\text{cat, app}}$ (s$^{-1}$)</th>
<th>$k_{\text{cat}}/K_{m,\text{app}}$ ($\times 10^3$ M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>24 ± 2</td>
<td>17.2 ± 0.5</td>
<td>730 ± 60</td>
</tr>
<tr>
<td>N170A</td>
<td>36 ± 4</td>
<td>15.9 ± 0.7</td>
<td>0.442 ± 0.053</td>
</tr>
<tr>
<td>N170D</td>
<td>13 ± 1</td>
<td>5.8 ± 0.2</td>
<td>0.446 ± 0.038</td>
</tr>
<tr>
<td>I195L</td>
<td>7.4 ± 0.5</td>
<td>42 ± 2</td>
<td>5.68 ± 0.47</td>
</tr>
<tr>
<td>I195F</td>
<td>35 ± 3</td>
<td>16 ± 1</td>
<td>0.46 ± 0.048</td>
</tr>
</tbody>
</table>

$^a$ Assays were performed at 25 °C and contained 0.5 µg of BphJ variants, 1 mM MnCl$_2$, all other cofactors were at saturating concentrations while the concentration of investigated cofactor was varied from 0.1 $K_m$ to at least 5 $K_m$ in 100 mM HEPES buffer pH 8.0.
Table 4.7 Steady-state kinetic parameters of cofactors in dehydrogenase reaction

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Cofactor</th>
<th>$K_{m, \text{app}}$ (µM)</th>
<th>$K_{m, \text{app}}$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>NAD$^+$</td>
<td>31.2 ± 2.3</td>
<td>30.2 ± 2.7</td>
</tr>
<tr>
<td></td>
<td>CoA</td>
<td>49.4 ± 5.8</td>
<td>12.6 ± 1.3</td>
</tr>
<tr>
<td>N170A</td>
<td></td>
<td>19.2 ± 1.7</td>
<td>19.2 ± 1.7</td>
</tr>
<tr>
<td>N170D</td>
<td></td>
<td>424 ± 31</td>
<td>37.7 ± 3.4</td>
</tr>
<tr>
<td>I171A</td>
<td></td>
<td>64.2 ± 5.6</td>
<td>33.8 ± 3.3</td>
</tr>
<tr>
<td>I171F</td>
<td></td>
<td>61.1 ± 3.6</td>
<td>33.2 ± 3.3</td>
</tr>
<tr>
<td>I195A</td>
<td></td>
<td>25.6 ± 1.8</td>
<td>27.3 ± 2.5</td>
</tr>
<tr>
<td>I195F</td>
<td></td>
<td>42.3 ± 1.8</td>
<td>45.3 ± 2.3</td>
</tr>
<tr>
<td>I195W</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Assays were performed at 25 °C and contained 0.5 µg of BphJ variants, 1 mM MnCl$_2$, all other cofactors were at saturating concentrations while the concentration of investigated cofactor was varied from 0.1 $K_m$ to at least 5 $K_m$ in 100 mM HEPES buffer pH 8.0.
Table 4.8 Acetaldehyde channeling efficiencies in Bphl-BphJ variant complexes

<table>
<thead>
<tr>
<th>BphJ Variants</th>
<th>Channeling Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>95 ± 5</td>
</tr>
<tr>
<td>N170A</td>
<td>83 ± 2</td>
</tr>
<tr>
<td>N170D</td>
<td>83 ± 2</td>
</tr>
<tr>
<td>I171A</td>
<td>84 ± 3</td>
</tr>
<tr>
<td>I171F</td>
<td>77 ± 6</td>
</tr>
<tr>
<td>I195A</td>
<td>84 ± 1</td>
</tr>
<tr>
<td>I195L</td>
<td>79 ± 7</td>
</tr>
<tr>
<td>I195F</td>
<td>80 ± 1</td>
</tr>
<tr>
<td>I195W</td>
<td>59 ± 1</td>
</tr>
</tbody>
</table>

\(^a\) Assays contained NAD\(^+\) at 5 \(K_m\) concentrations, coenzyme A at least 3 \(K_m\) concentrations, 1 mM MnCl\(_2\) and 5 \(\mu\)g of enzyme complex and 20 U aldehyde dehydrogenase (ALDH), which converts acetaldehyde to acetic acids. Reactions were quenched after 5 minutes with 24 \(\mu\)L of 3 N HCl. Acetyl CoA produced were detected at 254 nm by HPLC.
by BphJ or simply faster product (aldehyde) release due to channeling, kinetic parameters in the absence of NADH for the glycine variants that do not channel aldehyde were determined (Table 4.9). The $k_{cat}$ for the aldol cleavage of HOPA by the G322L, G322F and G323F variants, which are unable to channel aldehyde, were still higher by about 3 to 6-fold in the presence of NADH. This indicated that faster aldehyde release in the aldolase due to aldehyde channeling is not solely responsible for the activation of BphI by BphJ.

Allosteric activation of the aldol cleavage reaction in the presence of the NADH in I171A, N170A and N170D BphJ variants was observed to be less than 2-fold (Table 4.10). In variants I171A, N170A and N170D, BphI activity was 2.45, 0.91 and 0.54 (s$^{-1}$) respectively when BphJ was undergoing turnover. This represented a 3.6-, 2.1- and 2.5-fold increase in catalytic rate, respectively compared to unactivated BphI. Thus, these variants resulted in a significant reduction in the activation of the aldol cleavage reaction by BphI, suggesting that the allosteric activation is mediated through these residues in BphJ.

4.8 Dissociation Constant of NAD$^+$ in BphI-BphJ

The wild-type enzyme exhibited a dissociation constant ($K_d$) for NAD$^+$ of 1.8 ± 0.2 µM (Table 4.11). The $K_d$ for NAD$^+$ in the N170D variant increased by only 1.7-fold (Figure 4.5) while substitution with alanine, which is not capable of direct hydrogen bonding to the C2 and C3 hydroxyls of the nicotinamide ribose, led to an increase in $K_d$ of ~4-fold. In comparison, substitution of the adjacent residue, Ile-171, (located ~7 Å from the nicotinamide moiety of NAD$^+$), with alanine led to an increase of $K_d$ for NAD$^+$ by 20-fold.
Table 4.9 Steady-state kinetic constants of BphI and glycine variants for the aldol cleavage of HOPA in the absence of BphJ cofactors

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>( K_m (\mu M) )</th>
<th>( k_{cat} (s^{-1}) )</th>
<th>Fold activation of BphI by NADH ( b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>158 ± 20</td>
<td>0.79 ± 0.06</td>
<td>5.1 ± 0.4</td>
</tr>
<tr>
<td>G322L</td>
<td>81 ± 14</td>
<td>0.286 ± 0.02</td>
<td>3.8 ± 0.3</td>
</tr>
<tr>
<td>G323L</td>
<td>216 ± 7</td>
<td>0.100 ± 0.009</td>
<td>6.4 ± 0.6</td>
</tr>
<tr>
<td>G322F</td>
<td>53 ± 13</td>
<td>0.45 ± 0.07</td>
<td>2.4 ± 0.2</td>
</tr>
<tr>
<td>G323F</td>
<td>130 ± 10</td>
<td>0.72 ± 0.02</td>
<td>6.0 ± 0.2</td>
</tr>
</tbody>
</table>

\( a \) Assays contained 1 mM MnCl\(_2\). HOPA concentrations varied from 0.1 \( K_m \) to 5 \( K_m \) in 100 mM HEPES buffer (pH 8.0) \( b \) Ratio of \( k_{cat} \) of BphI in the presence of NADH to the \( k_{cat} \) in the absence of NADH. BphI variants are in complex with wild type BphJ.
Table 4.10 Effects of NADH on steady-state kinetic parameters for the aldol cleavage of HOPA by BphI $^a$

<table>
<thead>
<tr>
<th>BphJ in the complex</th>
<th>$K_m$ (µM)</th>
<th>$k_{cat, app}$ (s$^{-1}$) ($\times 10^{-2}$)</th>
<th>Fold activation by NADH$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- NADH</td>
<td>158 ± 20</td>
<td>79 ± 6</td>
<td>-</td>
</tr>
<tr>
<td>+ NADH</td>
<td>89 ± 8</td>
<td>407 ± 7</td>
<td>5.1 ± 0.4</td>
</tr>
<tr>
<td>I171A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- NADH</td>
<td>152 ± 22</td>
<td>68 ± 3</td>
<td>-</td>
</tr>
<tr>
<td>+ NADH</td>
<td>28 ± 5</td>
<td>96 ± 3</td>
<td>1.4 ± 0.1</td>
</tr>
<tr>
<td>N170A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- NADH</td>
<td>110 ± 12</td>
<td>44 ± 1</td>
<td>-</td>
</tr>
<tr>
<td>+ NADH</td>
<td>32 ± 5</td>
<td>51 ± 2</td>
<td>1.2 ± 0.1</td>
</tr>
<tr>
<td>N170D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- NADH</td>
<td>192 ± 18</td>
<td>22 ± 1</td>
<td>-</td>
</tr>
<tr>
<td>+ NADH</td>
<td>34 ± 5</td>
<td>36 ± 1</td>
<td>1.6 ± 0.1</td>
</tr>
</tbody>
</table>

$^a$ Assays contained 1 mM MnCl$_2$, HOPA concentrations varied from 0.1 $K_m$ to $10K_m$ in 100 mM HEPES buffer (pH 8.0)

$^b$ Ratio of $k_{cat}$ (s$^{-1}$) of BphI in the presence of NADH (+NADH) and in the absence of NADH (-NADH). + NADH: assays contain NADH to a final concentration of 0.4 mM. –NADH: no NADH added to the assays.
Table 4.11 Dissociation constants of NAD⁺ for wild-type and dehydrogenase variants

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$K_d$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>1.8 ± 0.2</td>
</tr>
<tr>
<td>N170A</td>
<td>8 ± 1</td>
</tr>
<tr>
<td>N170D</td>
<td>3.2 ± 0.4</td>
</tr>
<tr>
<td>I171A</td>
<td>36 ± 4</td>
</tr>
<tr>
<td>C131A</td>
<td>48 ± 6</td>
</tr>
<tr>
<td>C131S</td>
<td>43 ± 7</td>
</tr>
</tbody>
</table>

$^a$ Dissociation constants were determined by monitoring tryptophan fluorescence quenching upon NAD⁺ binding (excitation wavelength of 295 nm and emission wavelength of 330 nm). The titrations were carried out in 500 µL of 20mM HEPES pH 8.5 at 25°C with 100 µg of enzyme, with stepwise addition NAD⁺ aliquots. Mixtures containing no NAD⁺ were used as a control.
Figure 4.5. Tryptophan fluorescence quenching of BphJ N170D variant.

Fluorescence intensity was corrected for dilution factors and the dissociation constant was calculated by fitting to equation 2, which describes binding to one site, and averaging two replicates.
4.9 Discussion

A putative channel within DmpG-DmpF that may allow aldehydes to transverse from the aldolase to the dehydrogenase active sites was identified using the program MOLE with the crystal structure of DmpG-DmpF. This prediction was supported by the effects of replacements of two residues, Gly-322 and Gly-323, in BphI, which were predicted to line the narrowest diameter of the channel. The G322L and G322F variants completely abrogate substrate channeling, suggesting that these residues have effectively blocked the channel while the G322A variant displayed ~20% reduced aldehyde channeling efficiencies relative to the wild-type complex. On the other hand, as the side chain size at position 323 was increased, the size of the aldehyde that these variants could efficiently channel became progressively smaller, supporting the premise that the lower channeling efficiencies observed in the variants is due to steric blockage of the channel. In tryptophan synthase, physical blockage of indole channeling has also been attempted by replacing a cysteine residue (Cys-170) with tryptophan or phenylalanine, resulting in reduced indole channeling rate by at least 10- and 5000-fold, respectively (94). In carbamoyl phosphate synthetase, replacement of Gly-359 with phenylalanine or tyrosine residues blocked ammonia channeling, manifested by uncoupling of hydrolysis of glutamine to carbamoyl phosphate formation and a lag in the time course for carbamoyl phosphate formation (101). A G359S variant in carbamoyl phosphate synthetase exhibited reduced ammonia channeling while preventing hydroxylamine channeling while the G359L variant is defective in channeling both compounds.

The TTHB246-TTHB247 complex is able to channel acetaldehyde efficiently (>90%) but channeling efficiency is only ~57% for the larger propionaldehyde. Significantly, the residue corresponding to Gly-322 of BphI is an alanine in TTHB246. A single A324G substitution in
TTHB246 was able to increase propionaldehyde channeling efficiency to that observed for acetaldehyde further supporting the importance of the glycine residue in this position for channeling aldehydes larger than acetaldehyde.

Previous steady-state kinetic analysis of BphI indicated that the enzyme follows an ordered sequential mechanism whereby the acetaldehyde product must leave the active site first prior to the release of the other product, pyruvate (66). Intuitively this is important to allow the channeling of the aldehyde prior to the release of pyruvate. A reduction in aldehyde channeling efficiency was observed when two residues, His-20 and Tyr-290, present at the substrate entrance and aldehyde channel entrance, were replaced by alanine and phenylalanine, respectively. It appears that these residues may be important in ensuring the coordinated release of products and preventing escape of acetaldehyde from the substrate entrance/pyruvate release site in BphI. When His-20, the proposed catalytic base of the enzyme, is poised for C4-OH proton abstraction of the substrate, the substrate entry/pyruvate release tunnel that leads from the aldolase active site to the bulk solvent is closed. This presumably prevents aldehyde from escaping through this route following C-C bond cleavage of the substrate. The H20A variant is still active, albeit with a 45-fold lowering of $k_{cat}$, due to hydroxide ions replacing the function of the catalytic base in the variant (65). The variant however displayed a marked reduction (>70% reduction) in aldehyde channeling efficiency, possibly due to the fact that the methyl side chain of alanine is insufficient in size to effectively block the escape of aldehydes from the aldolase. In the proline dehydrogenase, PutA, which channels a pyrroline-5-carboxylate intermediate, a tyrosyl residue (Tyr-437) has been similarly observed to block the active site of the dehydrogenase domain from the bulk solvent (161). This residue could potentially function as a substrate entry gate analogous to that of His-20 in BphI.
It was previously demonstrated that the invariant tyrosine residue in the active site of BphI is not the catalytic acid, contrary to a previous suggestion that during proton transfer torsional movement of Tyr-290 results in an opening of the channel entrance allowing for aldehyde channeling to the dehydrogenase to occur (65). Closer inspection of the active site of the aldolase suggests that the binding of the modelled 4-hydroxy-2-oxoacid substrate will result in a steric clash between the C4 of the substrate and the oxygen in the para position of the phenol ring, necessitating the movement of this residue upon substrate binding, which could lead to the opening of the tunnel for aldehyde channeling prior to the aldol cleavage reaction (65). Upon cleavage of the 4-hydroxy-2-oxoacid, the aldehyde is released into the tunnel and Tyr-290 may adopt a closed conformation, as visualized in the crystal structure of DmpG-DmpF, preventing the back diffusion of the aldehyde. In the BphI Y290F variant, channeling efficiency was reduced by about 30%; this could either be due to relief of the steric constraint that is required for tunnel entrance opening or to the fact that the tunnel cannot be closed completely allowing some aldehydes to escape from the aldolase through the substrate entry/pyruvate release site. Aromatic residues have been proposed as gating residues in other enzyme systems that exhibit substrate channeling. For example, the indole side chain of Trp-74 was observed to adopt an alternative rotameric conformation that opens the ammonia tunnel when a glutamine analog was bound in the enzyme (162). In tryptophan synthase movement of Phe-280 in the presence of K⁺ and Cs⁺ results in opening of the indole channel (91).

When Leu-89 in BphI was replaced with the smaller alanine residue, aldehyde channeling efficiencies were also reduced. The delta carbons of Leu-89 form two faces of the tunnel. Cδ1 is directly opposite to Gly-322 and Gly-323 while Cδ2 forms the base of the active site of the aldolase. Replacing Leu-89 with alanine increases the size of the active site of the aldolase and
should broaden the aldehyde channel. Enlargement of the aldolase active site in the L89A variant has been demonstrated to increase the specificities of the aldolase for butyraldehyde, pentaldehyde and hexaldehyde while maintaining similar specificities for acetaldehyde and propionaldehyde (65). However, this modification had negative effects for aldehyde channeling; suggesting that the channel cannot be closed since the entrance to the tunnel is now widened (Figure 4.6).

By inference from the DmpF structure (50) Ile-171, Ile-195 and Leu-197 in BphJ form the exit of the substrate channel connecting the active sites of the aldolase and dehydrogenase. This was confirmed by reduced aldehyde channeling efficiencies when Ile-195 was replaced with bulky residues phenylalanine and tryptophan. A BphJ I195F variant was demonstrated to have a channeling efficiency of 80 ± 1% (115) while I195W exhibited a channeling efficiency of 59 ± 1%. Aldehyde channeling in the BphI-BphJ system is still efficient (83 ± 2 %) when Asn-170 was replaced with alanine, which cannot form hydrogen bond interactions with the ribosyl hydroxyls of the nicotinamide moiety of NAD⁺. Therefore, the previous proposal (50), that substrate channeling is dependent on hydrogen bonding interaction between the side chain of Asn-171 (Asn-170 in BphJ) and NAD⁺, which alters the Cα backbone and rotameric configuration of Ile-172, Ile-196 and Met-198 to enable the exit of the channeled aldehyde, is not supported by the experimental evidence. In comparison, Tyr-290 in BphI partially blocks the entrance of the aldehyde channel. Previous replacement of Tyr-290 compromised channeling efficiency (up to 85% reduction in channeling efficiency). Therefore, Tyr-290 in the aldolase appears to be the important gating residue for aldehyde channeling in BphI-BphJ (115). This implies that significant conformation changes on the channel exit residues can occur to allow for aldehydes to enter the dehydrogenase active site from the channel.
Figure 4.6 Model of the L90A variant of the aldolase DmpG
This figure shows the enlarged substrate binding site and channel. The model was generated by superimposing the pyruvyl moiety of HOPA on the experimentally observed oxalate. The active site and aldehyde channel are shown in magenta. Corresponding residue numbers in BphI are indicated in brackets.
Based on the results presented above, combined with previous work on the catalytic mechanism of the aldolase (65), a mechanism for aldehyde channeling in the aldolase-dehydrogenase complex is proposed (Figure 4.7). In this mechanism the 4-hydroxy-2-oxoacid enters the aldolase through the substrate tunnel where it encounters His-20, which undergoes a conformation change to allow the 4-hydroxy-2-oxoacid to enter the active site of the aldolase. Following substrate entry, Tyr-290 rotates to accommodate the substrate in the active site, thus opening the aldehyde channel. The distal end of the alkyl side chains of aldehydes > 4C in length protrude into the channel. His-20 adopts a “closed” conformation which also enables it to abstract a proton from C4-OH to initiate aldol cleavage. The aldehyde is released into the channel allowing for the torsional movement of Tyr-290 to a “closed” position, preventing the aldehyde from diffusing back into the aldolase active site. His-20 opens the substrate entry tunnel allowing pyruvate to leave and a second molecule of substrate to enter.

In order to determine if efficient aldehyde channeling requires precise interactions between partner aldolase-dehydrogenase from the same organism, chimeric TTHB246-BphJ BphI-TTHB247 enzymes were expressed and purified. While stable chimeric aldolase-dehydrogenase complexes were observed (77) the efficiency of aldehyde channeling in chimeric complexes is low (10 %). From multiple-sequence alignments and the crystal structure of the DmpG-DmpF complex, potential charge-charge repulsions in the TTHB246-BphJ chimera (Arg-207 with Arg-330 and Arg-66 with Arg-176) were identified that may prevent proper alignment of enzyme subunits resulting in aldehydes escaping from the subunit interface. No obvious unfavourable side chain interactions can be seen in the BphI-TTHB247 complex. Alternatively, improper coupling and communication between enzyme subunits could also result in a reduced efficiency of substrate channeling.
Figure 4.7 Proposed mechanism of substrate channeling in BphI-BphJ.
(A) Substrate enters the aldolase through the substrate tunnel where it passes His-20. (B) Upon substrate entry, Tyr-290 rotates to accommodate the substrate in the active site, thus opening the aldehyde channel. His-20 adopts a closed conformation and abstracts a proton from the C4-OH of the substrate to initiate the aldol cleavage reaction (C) Upon aldol cleavage, the aldehyde is channeled to the dehydrogenase via the aldehyde channel and Tyr-290 reverts to a “closed” position, preventing the aldehyde from diffusing back into the aldolase active site. (D) His-20 adopts an “open” conformation allowing pyruvate to leave and a second molecule of substrate to enter.
The asymmetric unit in the crystal of DmpG-DmpF consists of two tetramers, with three of the four DmpF subunits containing bound NAD\(^+\). In the crystal structure of DmpG-DmpF, three residues at the exit of the dehydrogenase adopt different conformations and only in the NAD\(^+\)-bound form was the channel exit opened for aldehyde to enter the dehydrogenase active site. It has been determined that NADH activates the aldolase by 5-fold in the BphI-BphJ complex (64). The BphI G322L, G323L and G323F variants which completely lack the ability to channel acetaldehyde were found to be activated similarly to the wild-type in the presence of NADH. The activation of BphI by BphJ is therefore due to allosteric communication between the two enzymes rather than faster aldehyde release due to channeling.

Replacement of Ile-171 with alanine in BphJ led to a 20-fold increase in NAD\(^+\) dissociation constant for BphJ. In DmpG, Ile-171 is located on the N-terminus on an \(\alpha\)-helix and forms hydrophobic interactions with Ile-196 and Met-198 on an opposing \(\beta\)-sheet. It is envisaged that shortening the side chain in the I171A variant may disrupt this interaction, thereby altering the C\(\alpha\) backbone and affecting the tertiary structure of the NAD\(^+\)-binding domain (Figure 4.1A). Dissociation constants for the Asn-170 variants of BphJ also increased, but by <10-fold indicating that although Asn-170 interacts with NAD\(^+\), it is not critical for cofactor binding. Thus, when Asn-170 is replaced with aspartate or shortened to alanine, hydrogen bonding with NAD\(^+\) can still occur between the carboxylate oxygen and ribosyl hydroxyls in N170D or space may be created for water molecules to interact with NAD\(^+\) in N170A. Indeed, in the structurally similar GAPDH and ASADH, NAD\(^+\) binds in a similar position as DmpF, but ordered water molecules form hydrogen bond interactions with the hydroxyls of the ribose of NAD\(^+\) instead (69, 163, 164). Interestingly, Asn-170 and Ile-171 appear to be important in mediating allosteric activation of the aldol cleavage reaction of BphI upon NAD\(^+\) binding in BphJ. Replacement of
these residues led to attenuation of BphI activation. This allosteric likely ensures that aldehydes released to the solvent in the absence of cofactors are kept to a minimum since the aldol cleavage reaction always occurs in the presence of 4-hydroxy-2-oxoacids. Asn-171 and Ile-172 are located in the α7 and α8 secondary structures of DmpF that interact with the α5 secondary structure of the HMGL-like domain in the aldolase, DmpG. It is possible that the allosteric activation is mediated through these inter-subunit contacts. However, the precise molecular mechanism of this activation is currently unclear since no obvious structural changes in the aldolase can be observed between the crystal structures of the apo- and NAD⁺-bound forms of the DmpF-DmpG complex.

The results presented above elucidate the correlation of substrate channeling with allosteric activation. In the absence of cofactors, the aldolase has lower activity, minimizing the amount of aldehyde produced. This aldehyde would be released to the bulk solvent, where it would need to be detoxified via other means. In the presence of dehydrogenase cofactors, the aldolase activity significantly increases and the aldehyde is efficiently channeled to the dehydrogenase where it is transformed to acyl-CoAs. Allosteric communication is also seen in tryptophan synthase where ligand binding at the α-site influences the binding and reaction rate occurring at the β-site. (165). Although the mechanism of BphI-BphJ activation is unknown, further experiments to decipher the change in structure upon NAD⁺/NADH binding could help elucidate this mechanism.

Analysis of the crystal structure of the aldolase-dehydrogenase complex enabled the identification of a possible route for the direct channeling of aldehyde products from the aldolase to the dehydrogenase. The channel appears to be dynamic and therefore the dimensions of the channel derived from static crystal structure are inadequate in predicting the sizes of aldehydes
that can diffuse through the channel. The coordinated conformational changes of specific gating residues appear to ensure the efficient channeling of aldehydes. Interaction of Asn-170 with the nicotinamide cofactor does not appear to be important for substrate channeling. However, Asn-170 and Ile-171 are important for allosteric activation of BphI. The results presented herein complement the work previously completed and shed further light on the mechanism of substrate channeling and allosteric activation in this enzyme complex.
Chapter 5: Biochemical and Structural Characterization of HsaF-HsaG

Data from this chapter was published in its entirety in the following journal article:


Statement of Contributions:
Sarah McKenna assisted with the purification of some of the enzymes used in these experiments, grew some of the crystals used in crystallography experiments and performed the metal specificity assays under the direct supervision of Jason Carere. Matthew Kimber helped with refinement of the X-ray data and with the idea of the phosphate binding loop reorganization.
5.1 Introduction

*Mycobacterium tuberculosis* is the causative agent of tuberculosis (TB), a bacterial infection that kills nearly 2 million people annually (1). During infection, *M. tuberculosis* survives within the phagosome compartment of macrophages for long periods of time due to its ability to inhibit phagosome-lysosome fusion. Cholesterol, derived from the host’s membrane and lipid bodies within macrophages, has been implicated as an important carbon source for *M. tuberculosis* within the nutrient poor macrophage phagosome-like compartment. Specifically, cholesterol catabolism leads to formation of the central metabolites pyruvate and propionyl-CoA. The latter can be transformed into tricarboxylic acid cycle metabolites, or unique cell wall lipids important for virulence in this bacterium. Cholesterol uptake and degradation have been shown to be essential for intracellular growth and survival of *M. tuberculosis* in animal models (20, 23) and the disruption of certain genes encoding enzymes in the cholesterol degradation pathway leads to reduced virulence and, in some cases, accumulation of toxic dead-end steroid metabolites that are detrimental to the survival of the bacterium growing on cholesterol (35, 166, 167).

Although the ability of diverse bacteria to degrade steroids (168, 169) has been known for some time, the exact catabolic pathway and its associated genes/enzymes is still not fully elucidated. Comparisons of the *M. tuberculosis* genome to those of other steroid degrading bacteria, such as the related actinomycete, *Rhodococcus jostii* RHA1 (5) has allowed the identification of a gene cluster encoding putative cholesterol degrading enzymes. Cholesterol degradation can be divided into two parts: side chain degradation, and the degradation of the steroid rings. The latter is thought to proceed through a pathway analogous to the bacterial meta-cleavage pathway of aromatics, such as phenols and polychlorinated biphenyls (PCBs).
The degradation of diverse aromatic compounds via the meta-cleavage pathway generally leads to the formation of the common intermediate, 4-hydroxy-2-oxopentanoate, as exemplified by the bph pathway responsible for biphenyl/polychlorinated biphenyl (PCBs) degradation. A divalent metal-dependent pyruvate aldolase then catalyzes a retro aldol cleavage of the intermediate, generating pyruvate and acetaldehyde (64). Acetaldehyde is then converted to acetyl-CoA by an acylating aldehyde dehydrogenase that forms a complex with the aldolase. Various aldolase-dehydrogenase complexes including BphI-BphJ from the biphenyl degradation pathway, DmpG-DmpF from the phenol degradation pathway and TTHB246-TTHB247 from Thermus thermophilus have been characterized. Structural and biochemical studies have shown that the acetaldehyde product of the aldolase is not released in the bulk solvent, but rather travels to the aldehyde dehydrogenase via a molecular tunnel connecting the active sites of the two enzymes (50, 115). This process, known as substrate channeling, enables the sequestration of the labile and toxic aldehyde, while ensuring that unwanted products from competing cellular reactions are minimized (83, 87, 170). Analysis of the aldolase and dehydrogenase genes across the genomes of diverse aromatic-degrading bacteria reveals that they are generally adjacent to each other, with the aldehyde dehydrogenase preceding the aldolase gene in an operon. With the exception of a few rare cases, attempts to express each gene separately in a heterologous host such as in E. coli leads to formation of inclusion bodies or unstable proteins (50, 64, 70).

The aldolase-dehydrogenase complexes from steroid degradation pathways have not been previously purified and characterized. Phylogenetic analysis of protein sequences reveals that these aldolases and dehydrogenases (HsaF and HsaG encoded by Rv3534c and Rv3535c respectively in M. tuberculosis H37Rv) form a clade in the phylogenetic tree (Clade I) distinct from aldolases and dehydrogenases from other aromatic degradation pathways (77).
physiological substrates of HsaF-HsaG differ from those of other aromatic degradation pathways as the steroid degradation pathway yields 4-hydroxy-2-oxohexanoate, one carbon longer than 4-hydroxy-2-oxopentanoate, the common metabolite from other aromatic degradation pathways (such as PCBs). Aldol cleavage of 4-hydroxy-2-oxohexanoate produces propionaldehyde rather than acetaldehyde, implying both enzymes act on substrates larger by one methylene group (Figure 5.1).

The genomes of *M. tuberculosis* and related bacteria are also predicted to encode an additional aldolase encoded by Rv3469c, annotated as MhpE, based on the nomenclature of the *E. coli* 4-hydroxy-2-oxopentanoate aldolase in the meta-cleavage pathway of 3-hydroxyphenylpropionate degradation. This gene is neither within the cholesterol degradation gene cluster, nor an aromatic degradation gene cluster, and there is no nearby putative aldehyde dehydrogenase gene. Instead, Rv3469c is within an operon with genes that share homology to dTDP-glucose 4,6-dehydratase, and the large catalytic subunit of acetolactate synthase. Possibly the *M. tuberculosis* genome encodes two putative 4-hydroxy-2-oxoacid aldolases, allowing functional redundancy of enzymes in steroid catabolism.

This chapter describes the heterologous expression, purification and characterization of HsaF-HsaG from *M. tuberculosis* H37Rv. The crystal structure, steady-state kinetic parameters and substrate specificity of HsaF-HsaG were determined and compared with aldolase-dehydrogenase complexes from other aromatic degradation pathways. The additional putative aldolase encoded by Rv3469c is also analysed, and is shown to lack 4-hydroxy-2-oxoacid aldolase activity, but does possess oxaloacetate decarboxylase activity. Detailed sequence analysis reveals that this protein and its orthologs differ in key catalytic site residues, allowing them to be differentiated from the 4-hydroxy-2-oxoacid aldolases.
Figure 5.1 Reaction catalyzed by HsaF and HsaG

4-hydroxy-2-oxohexanoate is cleaved to pyruvate and propionaldehyde by HsaF. The propionaldehyde is then channeled to HsaG where propionyl-CoA is produced using NAD$^+$ and CoA as cofactors.


5.2 Expression and Purification of Recombinant Proteins

The genes *hsaG* and *hsaF*, encoding the dehydrogenase and aldolase of the cholesterol degradation pathway are adjacent within the genome of *M. tuberculosis* and are partially overlapping. The individual genes were PCR amplified from *M. tuberculosis* genomic DNA and separately inserted into two compatible *E. coli* plasmids, pBTLT7 and pET28a, respectively. Since the dehydrogenase, HsaG, was expressed with an N-terminal histidine-tag encoded by the pET28a vector, the HsaF-HsaG complex could be co-purified by Ni$^{2+}$-NTA chromatography. However, the yields of purified protein complex were very low, with about 200 to 400 µg of protein purified per litre of recombinant *E. coli* culture. In an attempt to improve protein yields, the *hsaF*, *hsaG* and *hsaGF* genes were transferred into the Rhodococcal-*E. coli* shuttle plasmid pTIP QC1, for expression of the genes in *R. jostii* RHA1, an Actinomycete that has been previously used successfully to over-express other *M. tuberculosis* cholesterol degrading genes (171). The individual enzymes and enzyme complex were purified by Ni$^{2+}$-NTA column chromatography. The yields of the HsaF, HsaG and HsaF-HsaG complex were about 2 to 3 mg per litre of recombinant *R. jostii* RHA1 culture, a significant improvement from the yields obtained from the *E. coli* expression system. Purified proteins were subjected to SDS-PAGE and their migration in the gel was consistent with the calculated molecular masses of 36.4 kDa and 32 kDa for HsaF and HsaG, respectively (Figure 5.2). The native molecular masses of the enzymes complex were determined by gel filtration to be 159 kDa, consistent with a quaternary structure of two aldolase and two dehydrogenase subunits (Figure 5.3). The individually purified HsaF and HsaG, on the other hand, appear to be dimer (native molecular mass of 67 kDa) and monomer (native molecular mass of 38 kDa), respectively. Similar to the *hsaF* and *hsaG*, the putative *mhpE* was expressed in recombinant *R. jostii* RHA1 to produce an N-terminal
Figure 5.2 SDS-PAGE gel of purified HsaF-HsaG, HsaF, HsaG and Rv3469c.

(A) Lane 1 is the molecular weight markers and lane 2 is the purified HsaF-HsaG complex. (B) Lane 1 is the Molecular weight markers, HsaF is in lane 2, HsaG is in lane 3 and the product of Rv3469c is in lane 4.
Figure 5.3. Elution volume of purified *M. tuberculosis* enzymes.

Elution volume is relative to the elution of Blue Dextrin (117.3 mL). HsaF elutes at 207.7 mL signifying a native molecular mass of 67 kDa. HsaG elutes at 226.0 mL signifying a native molecular mass of 38 kDa. HsaFG elutes at 180.2 mL signifying a native molecular mass of 159 kDa. MhpE elutes at 207.4 mL signifying a native molecular mass of 68 kDa.
histidine tagged product (Figure 5.2). Approximately 2 mg of protein was purified per litre of culture. The native molecular mass of the enzyme was determined by gel filtration to be 68 kDa, consistent with a homodimeric quaternary structure, similar to HsaF.

5.3 Steady State Kinetic Analysis of HsaF

Although HsaF could be expressed independently of HsaG, the purified HsaF had no detectable aldolase activity towards 4-hydroxy-2-oxopentanoate (HOPA) or 4-hydroxy-2-oxohexanoate (HOHA, < 0.0001 s\(^{-1}\)). The enzyme also lacked oxaloacetate decarboxylase activity. This secondary reaction normally catalyzed by metal-dependent pyruvate aldolases that produce a pyruvate enolate intermediate, which is also the key intermediate in the decarboxylase reaction. When HsaF was expressed and co-purified with HsaG, the complex showed aldolase activity. The specific activity of HsaF in the HsaF-HsaG complex was tested using HOHA as a substrate and with various divalent metal ions at concentrations of 0.1 and 1 mM (Table 5.1). The specific activity of HsaF was highest with Mn\(^{2+}\) followed by Co\(^{2+}\) and Ni\(^{2+}\). There was no detectable aldolase activity when the HsaF-HsaG was incubated in the presence of Cd\(^{2+}\), Zn\(^{2+}\) or Cu\(^{2+}\) (<0.0001 s\(^{-1}\)). Using Mn\(^{2+}\) as cofactor, HsaF had similar kinetic parameters for HOPA and HOHA (Table 5.2, Figure 5.4). Like other aldolases from this family, HsaF proved specific for only the 4(S) enantiomer of HOPA.

5.4 Allosteric Activation of HsaF by HsaG

Although HsaF expressed and purified independently from HsaG it had no detectable activity, pre-incubation of the purified protein with HsaG restored aldolase activity. The specific activity of HsaF in response to increasing concentration of HsaG followed a sigmoidal saturation curve (Figure 5.5).
Table 5.1 Metal ion specificity of HsaF

<table>
<thead>
<tr>
<th>Metal Ion</th>
<th>0.1 mM Metal Chloride Salts</th>
<th>1 mM Metal Chloride Salts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mn$^{2+}$</td>
<td>89.5</td>
<td>100</td>
</tr>
<tr>
<td>Ni$^{2+}$</td>
<td>21.1</td>
<td>76</td>
</tr>
<tr>
<td>Co$^{2+}$</td>
<td>30</td>
<td>59.5</td>
</tr>
<tr>
<td>Mg$^{2+}$</td>
<td>26.9</td>
<td>35.1</td>
</tr>
<tr>
<td>Ca$^{2+}$</td>
<td>3.6</td>
<td>15.4</td>
</tr>
<tr>
<td>Cd$^{2+}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Zn$^{2+}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cu$^{2+}$</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

$^a$ The activity obtained with 1 mM Mn$^{2+}$ is taken as 100 %. Assay contained 20 µg of enzyme either 0.1 or 1 mM metal chloride salt, 2 mM 4-hydroxy-2-oxohexanoate, 0.4 mM NADH, 19.2 units of LDH in 100 mM HEPES buffer (pH 8.0) with a total volume of 1 mL. Relative activity is the average activity of duplicated reactions (less than 10 % difference between replicates) compared to the activity of HsaF with 1 mM Mn$^{2+}$.
<table>
<thead>
<tr>
<th>Substrate</th>
<th>$K_m$ (μM)</th>
<th>$k_{cat, app}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_m$ app ($\times 10^4$ M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOPA</td>
<td>4.4 ± 0.5</td>
<td>0.41 ± 0.01</td>
<td>9.31 ± 1.08</td>
</tr>
<tr>
<td>HOHA</td>
<td>4.8 ± 0.6</td>
<td>0.38 ± 0.01</td>
<td>7.91 ± 1.01</td>
</tr>
</tbody>
</table>

Table 5.2 Steady state kinetic parameters of HsaF$^a$

$^a$ Aldolase assays were performed at 25°C and contained 0.4 mM NADH, 1 mM MnCl$_2$ and 19.2 Units LDH in 100 mM HEPES buffer (pH 8.0)
Figure 5.4. Representative kinetic plots of wild-type HsaF-HsaG utilizing various substrates.

Each panel contains an assay using different substrates; data was fit to the Michaelis-Menten equation using non-linear regression in GraphPad Prism. Panel A) HOHA, B) Propionaldehyde, C) Coenzyme A, D) NAD⁺. HsaG assay contained 400 µM NAD⁺, 1.5 mM coenzyme A and aldehyde concentrations were varied from at least 0.1 $K_m$ to 5 $K_m$. HsaG cofactor specificities were determined under similar conditions in which the acetaldehyde concentrations were held at 100 mM and concentrations of either NAD⁺ and CoA were varied between 0.1 $K_m$ and 5 $K_m$. 
Assays were performed at 25°C and contained 0.4 mM NADH, 1 mM MnCl$_2$ and 19.2 Units LDH in 100 mM HEPES buffer (pH 8.0) The data was fit with a sigmoidal model using GraphPad Prism.
Previous work with BphI-BphJ complex showed that the aldolase activity was enhanced by about 5-fold in the presence of the dehydrogenase cofactor, NADH. A similar allosteric activation also occurred in the HsaF-HsaG complex. However, the level of activation was less than 1.5-fold, using a concentration of 0.4 mM NADH (0.25 ± 0.01 µmol•min⁻¹•mg versus 0.33 ± 0.01 µmol•min⁻¹•mg)

5.5 Steady State Kinetic Analysis of the Putative MhpE

The putative MhpE protein encoded by \textit{Rv3469c}, had no detectable aldolase activity with HOPA or HOHA (<0.0001 s⁻¹), even when incubated with HsaG. No activity was detected with other 4-hydroxy-2-oxoacids, namely 2-oxo-3-deoxyoctonate (KDO), 4-hydroxy-4-methyl-2-oxoglutarate (HMG) and 4-carboxy-4-hydroxy-2-oxoadipate (CHA). The protein, however, exhibited oxaloacetate decarboxylase activity with a \(K_m\) value of 3.07 ± 0.07 mM, \(k_{cat}\) value of 0.49 ± 0.04 s⁻¹, and a catalytic efficiency (\(k_{cat}/K_m\)) of 1.60 ± 0.14 ×10² M⁻¹ s⁻¹. This activity was abolished upon incubation of the enzyme with Chelex and EDTA indicating that the oxaloacetate decarboxylase activity was metal ion dependent. The activity of the apo-enzyme was restored upon addition of excess Mn²⁺ in the assay.

5.6 Substrate Specificity of HsaG

Unlike HsaF, HsaG was active both by itself and in complex with HsaF. However, the uncomplexed HsaG exhibited substrate inhibition, with a \(K_i\) of 50 ± 7 and 63 ± 6 mM for acetaldehyde and propionaldehyde, respectively. Kinetic parameters for HsaG within the HsaF-HsaG complex were determined. The enzyme was found to have broad substrate specificity, being able to utilize aldehyde substrates of varying lengths and sizes with similar catalytic efficiencies (Table 5.3, Figure 5.4). The enzyme preferred NAD⁺ over NADP⁺ as cofactor but
Table 5.3 Steady state kinetic parameters of Hsag with various substrate and cofactors

<table>
<thead>
<tr>
<th>Substrate/Cofactor</th>
<th>$K_{m,\text{app}}$ (mM)</th>
<th>$k_{\text{cat, app}}$ (s$^{-1}$)</th>
<th>$k_{\text{cat}}/K_{m,\text{app}}$ ($\times 10^2$ M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetaldehyde$^a$</td>
<td>18 ± 2</td>
<td>9.4 ± 0.3</td>
<td>5.2 ± 0.6</td>
</tr>
<tr>
<td>Propionaldehyde$^a$</td>
<td>15 ± 1</td>
<td>11.1 ± 0.2</td>
<td>7.4 ± 0.5</td>
</tr>
<tr>
<td>Butyraldehyde$^a$</td>
<td>10.6 ± 0.8</td>
<td>7.3 ± 0.2</td>
<td>6.9 ± 0.6</td>
</tr>
<tr>
<td>Isobutyraldehyde$^a$</td>
<td>10 ± 1</td>
<td>5.8 ± 0.3</td>
<td>5.8 ± 0.6</td>
</tr>
<tr>
<td>Pentaldehyde$^a$</td>
<td>20 ± 3</td>
<td>7.8 ± 0.5</td>
<td>3.9 ± 0.6</td>
</tr>
<tr>
<td>Coenzyme A$^b$</td>
<td>0.040 ± 0.003</td>
<td>9.6 ± 0.2</td>
<td>2400 ± 180</td>
</tr>
<tr>
<td>NAD$^+$</td>
<td>0.022 ± 0.002</td>
<td>7.6 ± 0.2</td>
<td>3450 ± 330</td>
</tr>
<tr>
<td>NADP$^+$</td>
<td>1.4 ± 0.1</td>
<td>3.4 ± 0.1</td>
<td>25 ± 2</td>
</tr>
</tbody>
</table>

$^a$ Assays were performed at 25°C and contained 0.4 mM NAD$^+$ and 0.4 mM coenzyme A in 100 mM HEPES buffer (pH 8.0).

$^b$ Assays were performed at 25°C and contained 100 mM propionaldehyde, and 0.4 mM NAD$^+$ in 100 mM HEPES buffer (pH 8.0).

$^c$ Assays were performed at 25°C and contained 100 mM propionaldehyde, and 0.4 mM CoA in 100 mM HEPES buffer (pH 8.0).
exhibited a similar specificity for coenzyme A and dephospho-coenzyme A. The dissociation constant of NAD\(^+\), determined by tryptophan fluorescence quenching titration, was 1.8 ± 0.2 µM (Figure 5.6). The kinetic parameters of HsaG in the absence of HsaF are presented in Table 5.4.

5.7 Substrate Channeling in HsaF-HsaG

HsaF-HsaG was able to channel acetaldehyde and propionaldehyde from the aldol cleavage of HOPA and HOHA, respectively, with similar efficiencies (99 ± 3 % and 98 ± 1 %). Substrate channeling was completely abrogated by substitution of Gly-322 with phenylalanine. Gly-322 lines the tunnel linking the aldolase and dehydrogenase.

5.8 Crystal Structure of HsaF-HsaG

HsaF-HsaG structure was determined at 1.93 Å using a homology model of HsaF-HsaG based on the DmpF-DmpG (1NVM) as model for molecular replacement (refinement statistics, Table 5.5). The protein was co-crystallized with Mn\(^{2+}\) and the pyruvate enolate analogue, oxalate. Electron densities for these ligands were observed in the active site of the aldolase, HsaF. Unfortunately, no electron densities for cofactors in the dehydrogenase were observed even when the protein complex was co-crystallized in the presence of 20 mM coenzyme A, or soaked with 20 mM NAD\(^+\) or coenzyme A.

There are two aldolase and two dehydrogenase protomers in the asymmetric unit, corresponding to the tetrameric biological unit. The two aldolases (chains A and C) form a dimer with each dehydrogenase (Chains B and D) in the periphery forming a BACD elongated structure (Figure 5.7a and b). Analysis using the DALI server (172) revealed a root-mean-square deviation (RMSD) of 1.1 Å between the structure of HsaF and DmpG, the corresponding
Figure 5.6. Tryptophan florescence quenching of HsaG when in complex with HsaF.

Fluorescence intensity was corrected for dilution factors and the dissociation constant was calculated by fitting to equation 2, which describes binding to one site.
### Table 5.4 Steady state kinetics of HsaG as a monomer

<table>
<thead>
<tr>
<th>Cofactor/Substrate</th>
<th>$K_{m, \text{app}}$ (mM)</th>
<th>$k_{cat, \text{app}}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_{m, \text{app}}$ $(\times 10^3$ M$^{-1}$ s$^{-1}$)</th>
<th>$K_i$ (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetaldehyde$^a$</td>
<td>21.3 ± 2.7</td>
<td>19.0 ± 1.5</td>
<td>0.89 ± .13</td>
<td>50 ± 7</td>
</tr>
<tr>
<td>Propionaldehyde$^a$</td>
<td>12.4 ± 1.1</td>
<td>23.8 ± 0.7</td>
<td>1.93 ± 0.18</td>
<td>63 ± 6</td>
</tr>
<tr>
<td>Coenzyme A$^b$</td>
<td>0.415 ± 0.053</td>
<td>15 ± 1</td>
<td>36 ± 5.2</td>
<td>-</td>
</tr>
<tr>
<td>NAD$^+$$^c$</td>
<td>0.022 ± 0.004</td>
<td>3.9 ± 0.2</td>
<td>177 ± 33</td>
<td>-</td>
</tr>
</tbody>
</table>

$^a$ Assays were performed at 25°C and contained 0.4 mM NAD$^+$ and 1.5 mM coenzyme A in 100 mM HEPES buffer (pH 8.0).

$^b$ Assays were performed at 25°C and contained 100 mM propionaldehyde, and 0.4 mM NAD$^+$ in 100 mM HEPES buffer (pH 8.0).

$^c$ Assays were performed at 25°C and contained 100 mM propionaldehyde, and 1.5 mM CoA in 100 mM HEPES buffer (pH 8.0).
Table 5.5 Data collection, model refinement and final structure statistics for HsaF-HsaG

<table>
<thead>
<tr>
<th>Crystallographic data collection statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
</tr>
<tr>
<td>Cell dimensions: a =</td>
</tr>
<tr>
<td>b =</td>
</tr>
<tr>
<td>c =</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>Total observations (reflections)</td>
</tr>
<tr>
<td>Unique observations</td>
</tr>
<tr>
<td>Completeness (last shell)(^a)</td>
</tr>
<tr>
<td>(&lt;I/\sigma(I)&gt; ) (last shell)(^a)</td>
</tr>
<tr>
<td>(R_{\text{sym}} ) (last shell)(^a)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X-ray structure refinement statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymmetric unit contents</td>
</tr>
<tr>
<td>Protein chains</td>
</tr>
<tr>
<td>Water molecules</td>
</tr>
<tr>
<td>Other molecules</td>
</tr>
<tr>
<td>Average B-factor (Å(^2))</td>
</tr>
<tr>
<td>protein</td>
</tr>
<tr>
<td>water</td>
</tr>
<tr>
<td>(R_{\text{cryst}})</td>
</tr>
<tr>
<td>(R_{\text{free}}) (^b)</td>
</tr>
<tr>
<td>r.m.s.d. bond lengths (Å)</td>
</tr>
<tr>
<td>r.m.s.d. bond angles (°)</td>
</tr>
<tr>
<td>Ramachandran favored (%)</td>
</tr>
<tr>
<td>Ramachandran outliers (%)</td>
</tr>
</tbody>
</table>

\(^a\)The last shell includes all reflections between 1.93 and 1.98 Å.

\(^b\)\(R_{\text{free}}\) calculated using 5% of the data which were chosen randomly.
Figure 5.7 Structure of HsaF-HsaG

(A) HsaF-HsaG is organized as a heterotetramer. The HsaF aldolase domains (helices cyan and strands orange) form a homodimer. Each aldolase has an HsaG dehydrogenase (helices blue and strands red) bound to its periphery. (B) Heterotetramer rotated 90° along the 2-fold axis. (C) Tunnel running from the active site of the aldolase to the dehydrogenase as determined by Mole 2.0 (173) is shown in raspberry. Aldolase TIM barrel shown in cyan, communication domain shown in orange, dehydrogenase dimerization domain in blue and Rossmann fold in red. Images were generated in PyMOL (153).
aldolase from the phenol degradation pathway of *P. putida* CF600. The RMSD of the dehydrogenases HsaG and DmpF was however variable (1.6 - 2.2 Å), with the HsaG structure more closely resembling the apo-form of DmpF rather than the NAD\(^+\) bound form.

A significant difference between the two enzyme complexes occurs in a region of the dehydrogenase (residues 252-268 in DmpF and 248-257 in HsaG) that interacts with aldolase (Figure 5.8). HsaG lacks an extra β-sheet in this region in comparison to DmpF (residues 258-260 in DmpF). The topology of the loop in this region is also different whereby the side chain and main chain of Asn-251 of HsaG hydrogen bonds to the main chain oxygen atoms of Phe-58 and Phe-62 on HsaF, respectively. The DmpG loop extends farther away from the aldolase in the dimer interface with the side chains Phe-60 and Phe-64 of DmpG in the loop forming hydrophobic interactions with Leu-258 and Ile-260 of DmpF.

The active site of HsaF contains a Mn\(^{2+}\) ion that is ligated by two histidine residues (His-198 and His-200), two metal bound water molecules and the carboxyl oxygen atoms of oxalate (a pyruvate enolate analog). The guanidinium group of Arg-15 also interacts with the oxalate carboxylate group (Figure 5.9a). The histidine (His-19) and an enzyme bound water molecule previously determined to function as the catalytic base and acid in the homologous BphI are invariant in HsaF and DmpG.

The active sites of HsaF and HsaG are connected by a tunnel. Analysis using MOLE 2.0 (173) (Figure 5.7c) indicates that the tunnel is approximately 20 Å in length, beginning at Tyr-289 of the aldolase and ending at the entrance of the dehydrogenase active site. While most of the tunnel has a radius of approximately 2.5Å, a bottleneck of 1.3 Å wide is located adjacent to residues Gly-321 and Gly-322 of the aldolase. In HsaG, Ile-167, Ile-191 and Leu-193 line the
Figure 5.8 Variable length loops on the dehydrogenases.
The aldolases, HsaF and DmpG share sequence and structural identity and are shown in blue and red, respectively. The loop of HsaG, shown in green, is much shorter and interacts with a different region of the aldolases. The longer loop on DmpF (magenta) forms a small β-sheet.
Figure 5.9 Homology model of Rv3469c superimposed on HsaF.

A) Active site of HsaF with oxalate, a pyruvate enolate analogue and Mn$^{2+}$ (purple) bound. B) Homology model of the product of Rv3469c superimposed on HsaF. The arginine residue (Arg-18), that stabilizes the intermediate and residues which interact with the metal ligands are conserved between the enzymes (white). The catalytic histidine (His-19; yellow) in HsaF and tyrosine in the active site (Tyr-289 pink) are not conserved and are substituted with tyrosine and aspartate respectively in the gene product of Rv3469c. The residues conserved are critical to decarboxylation activity while the residues which are not conserved are involved with aldolase function. Homology model was produced by SWISS-MODEL (174), the two proteins share 27 % identity and 43 % similarity.
exit of the tunnel, however in DmpG, the leucine residue is replaced with a methionine. Superimposition of the three HsaG and DmpF structures showed that leucine and methionine occupy the same space.

HsaG contains two domains, a Rossmann fold domain and a dimerization domain that mediates interactions with the aldolase. These two domains are wider apart than what was observed in the DmpF-DmpG apo and NAD$^+$ bound structures. Unlike DmpG, the thiol of the catalytic cysteine that forms a thioacyl intermediate with the aldehyde substrate adopts only one conformation. HsaG is also related in structure to glyceraldehyde-3-phosphate dehydrogenase, aspartate semi-aldehyde dehydrogenase and methylmalonyl-CoA reductase (175-177). The latter, like HsaG, utilizes the Rossmann fold to bind the nicotinamide cofactor and coenzyme A alternately in its reaction mechanism, although it utilizes NADP$^+$ rather than NAD$^+$ as a cofactor. Although the adenine ribose of NADP$^+$ and coenzyme A both contain a phosphate, the difference in positions of the phosphates (2' in NADP$^+$ and 3' in coenzyme A) requires a subtle alteration in position of the ribose moiety for the phosphates of the two cofactors to bind in a similar position in the enzyme. This change in position causes the α and β phosphates of the NADP$^+$ to shift toward the catalytic cysteine resulting in the α phosphate of NADP$^+$ binding in the same position as the β phosphate of CoA. The movement of the α and β phosphates and ribose allow for the adenine ring and the ribose phosphates in the two cofactors to bind in a similar position in the protein forming hydrogen bonds with the hydroxyl of a serine residue (Ser-43) in methylmalonyl CoA reductase. Coenzyme A was modelled in HsaG by superimposing the adenine ribose with the NAD$^+$ in the crystal structure DmpF. There is an equivalent serine (Ser-41) in HsaG that may mediate similar interactions with the adenine phosphate of coenzyme A (Figure 5.10b). Ser-41 is ~4.3 Å away from the predicted position of the 3' hydroxyl of the ribose ring of NAD$^+$.
Figure 5.10 Interaction of S41 with modeled CoA.

(a) NAD$^+$ was positioned in the HsaG structure by superposition with NAD$^+$ bound DmpG. (b) Proximity of 3'-phosphate of modelled coenzyme A to Ser-41 side chain. The loop containing Ser-41 undergoes remodelling upon CoA binding to accommodate the phosphate.
(Figure 5.10a.) and it appears that, unlike in methylmalonyl CoA reductase, there may be no space to reorient the adenine ribose and the α and β phosphates, potentially accounting for the lower specificity of the HsaG for NADP⁺ than NAD⁺ in contrast to methylmalonyl CoA reductase.

5.9 Analysis of the Sequence Relationship between the Putative MhpE and HsaF

It was determined that although arginine (Arg-15) and two metal binding histidine residues (His-198 and His-200) of HsaF are fully conserved in the gene product of \( \text{Rv3469c} \), the putative MhpE lacks a histidine (His-19 in HsaF) that has been previously implicated for base abstraction of the 4-OH group of 4-hydroxy-2-oxoacids and lacks a tyrosine conserved in HOPA aldolases that helps mediate stereospecificity and gating of substrate channeling (Figure 5.9b). Instead the corresponding positions of these residues in the putative MhpE are replaced by phenylalanine and aspartate. In addition, the C-terminal domain, which interacts with the dehydrogenase and forms the tunnel linking the active sites of the aldolase-dehydrogenase complex sequence in HsaG, is missing in the putative MhpE. Orthologs of the putative MhpE are found in other \( \text{Mycobacterium} \) species, such as \( \text{Mycobacterium canettii} \) CIPT 140010059 (95% sequence identity), which are also misannotated as 4-hydroxy-2-oxopentanoate aldolases. In these \( \text{Mycobacterium} \) species, the related genes are part of an operon in between genes encoding proteins that share homology to dTDP-glucose 4,6-dehydratase and the large catalytic subunit of acetolactate synthase. There is no putative aldehyde dehydrogenase gene near their vicinity in the respective genomes. More distant homologs are found in the archaea \( \text{Nitrosopumilus maritimus} \) SCM1 and \( \text{Methanocorpusculum labreanum} \) Z (sequence similarities of 47% to the gene product of \( \text{Rv3469c} \); Figure 5.11). Like the Mycobacterial enzymes, these proteins contain the catalytic residues for pyruvate enolate formation and stabilization but not the catalytic base for aldolase.
Figure 5.11 Sequence alignment of 4-hydroxy-2-oxoacid aldolases with protein orthologs of Rv3469c.

Amino acid sequences of proteins and their uniprot accession number are as follows: HsaF from M. tuberculosis H37Rv (P51015), 4-hydroxy-2-oxoacid aldolase from T. thermophilus HB (TTHB246; P51016), Rv3569c gene product from M. tuberculosis H37Rv (P51017), Mlab_0987 gene product from Methanocorpusculum labreanum Z (A2SS50), Mcan_34851 gene product from Mycobacterium canetti CIPT 140010059 (G0TIX2) and Nmar_140 Nitrosopumilus maritimus SCM1 (A9A1R9). Secondary structures based on the HsaF crystal structure are indicated as horizontal arrows (beta strands) and helices (alpha helix) above the sequences. Fully conserved residues are highlighted in red. Positions of the histidine catalytic base, tyrosine that is important for substrate binding and the two glycine residues that line the aldehyde channel in 4-hydroxy-2-oxoacid aldolases are indicated as blue stars below the sequences. Conserved arginine and metal binding histidine residues are indicated by a blue triangle. This alignment was generated using ClustalX (178) and visualized using ESPript (179).
activity. It is therefore likely that these archeal gene products possess oxaloacetate decarboxylase activity despite having a different genomic context as the Mycobacterial genes (Figure 5.12).

5.10 Role of Serine 41 of HsaG in Coenzyme A Binding

Serine 41 is proposed to interact with the 3’ ribose phosphate group on the coenzyme A (Figure 5.10b). When Ser-41 was substituted with alanine or aspartate, the specificity constant ($k_{cat}/K_m$) for coenzyme A was reduced between 3 and 4 orders of magnitude (Table 5.6). The $k_{cat}$ values were still reduced for dephospho-CoA, which lacks the ribose phosphate, in the two variants. However the $K_{m,app}$ of variant enzymes with dephospho-CoA was similar to that of wild-type enzyme leading to a comparatively lower reduction in specificity constants for dephospho coenzyme A in the variants of between 44- to 360-fold.
Figure 5.12 Genetic map of Rv3469c and homologs.

Rv3469c and MCAN_34851 are located between the genes encoding putative dTDP-glucose dehydratase (dTDP-glu dehyd) and the large subunit of acetolactate synthase (acetolactate synth) in genomes of *M. tuberculosis* H37Rv and *M. canetti* CIPT 140010059, respectively. Mlab_0987 is adjacent to genes that encode a hypothetical protein (Hyp) and an RNA polymerase Rbp10 (RNA poly) in the genome of *Methanocorpusculum labreanum* Z. Nmar_0140 is upstream of the genes encoding putative AMP dependent synthetase/Ligase (Amp syn/lig) cyclase, hypothetical protein (Hyp) and aminoglycoside N-acetyltransferase (acetyl Transf) in the genome of *Nitrosopumilus maritimus* SCM1.
Table 5.6 Steady state kinetics of HsaG variants

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Cofactor</th>
<th>$K_{m,\text{app}}$ (µM)</th>
<th>$k_{\text{cat, app}}$ (s$^{-1}$)</th>
<th>$k_{\text{cat}}/K_{m,\text{app}}$ (×10$^4$ M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>Coenzyme A</td>
<td>40 ± 3</td>
<td>9.6 ± 0.2</td>
<td>24.0 ± 1.8</td>
</tr>
<tr>
<td></td>
<td>Dephospho-CoA</td>
<td>13 ± 1</td>
<td>9.7 ± 0.4</td>
<td>74.6 ± 6.5</td>
</tr>
<tr>
<td>S41I</td>
<td>Coenzyme A</td>
<td>1475 ± 190</td>
<td>0.054 ± 0.004</td>
<td>0.0036 ± 0.0005</td>
</tr>
<tr>
<td></td>
<td>Dephospho-CoA</td>
<td>64 ± 6</td>
<td>0.131 ± 0.005</td>
<td>0.205 ± 0.021</td>
</tr>
<tr>
<td>S41D</td>
<td>Coenzyme A</td>
<td>1525 ± 244</td>
<td>0.16 ± 0.01</td>
<td>0.0105 ± 0.0018</td>
</tr>
<tr>
<td></td>
<td>Dephospho-CoA</td>
<td>49 ± 5</td>
<td>0.83 ± 0.03</td>
<td>1.69 ± 0.183</td>
</tr>
</tbody>
</table>

$^a$ Assays were performed at 25°C and contained 100 mM propionaldehyde, and 0.4 mM NAD$^+$ in 100 mM HEPES buffer (pH 8.0)
5.11 Discussion

HsaF, HsaG and the HsaF-HsaG complex were successfully expressed and purified using recombinant *R. jostii* RHA1 as host. However HsaF was inactive by itself and complex formation with HsaG is required to induce a catalytically active conformation. This may be necessary to prevent deleterious formation and release of toxic aldehydes in the absence of the partner dehydrogenase. This is analogous to tryptophan synthase, another enzyme that exhibits substrate channeling, whereby complex formation led to 1- to 2-orders of magnitude increase in the activities of the α and β subunits (180). However, unlike the HsaF-HsaG system, the α and β subunits of tryptophan synthase are active when not in complex, albeit at a lower level. The thermophilic ortholog of HsaF from *Thermus thermophilus* (TTHB246, 49 % sequence identity to HsaF) on the other hand, is active without the partner dehydrogenase (TTHB247) and complex formation has negligible effect on $k_{cat}$ value for 4-hydroxy-2-oxopentanoate aldolase activity (77).

Similar to BphI, but unlike TTHB246 which prefers Co$^{2+}$ as cofactor, HsaF had the highest activity with Mn$^{2+}$ as a metal cofactor. HsaF has about 20-fold lower $K_m$ values for 4-hydroxy-2-oxoacids compared to TTHB246 and BphI (64, 77). The $k_{cat}$ values for these substrates are of the same order of magnitude to those determined for TTHB246 but are about 10-fold lower than BphI. Interestingly, the dehydrogenase HsaG has similar kinetic parameters for aldehydes of two to five carbons in length. In contrast, in the BphJ dehydrogenase, specificity constants for butyraldehyde is about half of the specificity constants for acetaldehyde and propionaldehyde. In addition $K_{m \text{ app}}$ value for pentaldehyde was too high to be determined and the specificity constant for this substrate is about 10-fold lower than acetaldehyde or propionaldehyde in BphJ. Previously, it was determined that the alkyl chain of aldehydes extend
towards the back of the active site of BphJ in proximity to Ile-195 (Ile-191 in HsaG) (69). Substitution of Ile-195 with larger tryptophan and phenylalanine residues led to reduced specificity towards long chain aldehydes in BphJ. Inspection of the structure of HsaG and comparison with the BphJ sequence reveals no obvious differences that would account for the ability of HsaG to utilize longer chain aldehydes more efficiently. A comparison of the crystal structure of HsaG and DmpF on the other hand reveals that Met-198 in DmpG extends farther into the tunnel than the equivalent Leu-193 in HsaG. Both of these residues are also at the back of the dehydrogenase active site, in contact with the distal end of the alkyl chain of the aldehyde substrates. Met-198 in DmpG is also more flexible as indicated by the multiple conformations of the side chain in the DmpG structures. This effective volume that Met-198 residue occupies will reduce its specificity for longer chain aldehydes compared to HsaG and BphJ.

HsaF-HsaG was able to channel acetaldehyde and propionaldehyde from the aldolase to the dehydrogenase with more than 90% efficiency. Similar to other orthologs, there appears to be a bottleneck in the tunnel linking the active site, lined by glycine residues (115). Substitution of one of these residues, glycine 322, to the larger alanine residue in HsaF blocks aldehyde channeling. BphI and the T. thermophilus aldolase homolog were shown to be activated by 5.1- and 2-fold respectively in the presence of dehydrogenase cofactors (77). In comparison, the activity of HsaF in the enzyme complex is only activated 1.3-fold in the presence of NADH.

Like HsaF, the gene product of Rv3469c was successfully purified and determined to form a soluble dimer. However unlike hsaF, there are no genes encoding an aldehyde dehydrogenase located in its vicinity in the M. tuberculosis genome. The putative MhpE has OAA decarboxylases activity, but lacks any detectable aldolase activity with the substrates attempted to date. Sequence analysis reveals that the putative MhpE lacks the catalytic histidine
important for C4-OH proton abstraction in the aldolase reaction, which accounts for its lack of 4-hydroxy-2-oxoacid aldolase activity. Instead this histidine is replaced with tyrosine. However the metal cofactor ligands and the positively charged arginine residue previously implicated for stabilization of the pyruvate enolate intermediate are conserved between HsaF and the putative MhpE. Since oxaloacetate decarboxylase proceeds through a pyruvate enolate intermediate, it is not surprising that the gene product of \textit{Rv3469c} possess this activity. These results showed that \textit{Rv3469c} and its orthologs are not 4-hydroxy-2-oxopentanoate aldolases and are misannotated in a number of \textit{Mycobacterium} genomes. Although it is unclear whether the OAA decarboxylase activity is the physiological or secondary activity, the fact that the gene is contiguous with a gene encoding the large catalytic subunits of acetohydroxy acid synthase (which catalyzes the transformation of two pyruvate molecules to form acetylacetate and carbon dioxide) suggests that the OAA decarboxylase activity that forms pyruvate may be the physiologically relevant activity.

Although they interact with a conserved region on the aldolase, the loop in HsaG is quite different than that in DmpF in length and composition. The loop in HsaG contacts the aldolase closer to the interface of the subunits and makes polar contacts with the backbone. The loop on DmpF appears to have both hydrophobic and polar contacts with the aldolase. The side chain of Asn-259 is seen in multiple conformations, some of which orient toward the main chain of the aldolase making polar contacts. The DmpF loop also forms a small β-sheet, with two β-strands interacting with each other at the distal end of the loop just before the turn; this may serve to stabilize the loop and/or its interaction with the aldolase. Sequence alignments and homology modeling have previously suggested that the ortholog BphJ, and other members of its clade, lack this loop entirely (77). Whether these loops, or the lack thereof, play a role in allosteric
activation and communication between the aldolase and dehydrogenase remains to be determined.

It was previously postulated, based on hydrogen-deuterium exchange mass spectrometry analysis, that NAD$^+$ and CoA share a common binding site in DmpF (82). Both NAD$^+$ and CoA have an ADP-ribose moiety. The only difference in this region of the molecules is the 3' hydroxyl of the ribose ring is phosphorylated in CoA, but not in NAD$^+$. Assuming this moiety binds in the same position, there must be sufficient room for the phosphate to bind or a reorganization of the loop containing Ser-42. The energetic cost to reorganize this loop is reduced by the potential hydrogen bonds formed with the 3' phosphate of CoA. As the $K_{m,app}$ value for dephospho-CoA is slightly less than CoA it is possible this energetic cost of loop reorganization is greater than the energy gain from interactions with the phosphate. The significantly increased $K_{m,app}$ of NADP$^+$ compared to NAD$^+$ indicates that the phosphate of NADP$^+$ yields even less binding energy, so the cost of loop rearrangement to now accommodate the 2'-phosphate does not compensate for the energetic gain obtained from phosphate binding. Wild-type HsaF-HsaG was able to utilize dephospho-CoA with similar specificity as CoA. The S41D and S41I variants have about 2000- to 6000-fold reduced catalytic efficiencies for CoA compared to the wild-type. While the specificity constants for dephospho-CoA were reduced by 44- to 360-fold in comparison, with similar $K_{m,app}$ values as the wild-type enzyme. This suggests that Ser-41 is involved in binding of the 3’ phosphate of the CoA.

In conclusion, HsaF and HsaG differ from previously characterized aldolase-dehydrogenase complexes from other aromatic pathways in terms of its substrate specificity, degree of allosteric activation, and solubility of the aldolase. Like its distant homolog methylmalonyl-CoA reductase, HsaG contains an equivalent serine residue that we propose
interacts with the 3’ phosphate of CoA adenine ribose. However, the difference in specificities for NAD$^+$ and NADP$^+$ in the evolutionarily related methylmalonyl-CoA reductase and HsaG indicates that this residue is not the molecular determinant for nicotinamide coenzyme discrimination.
Chapter 6: Conclusion and Future Directions

6.1 Summary and Conclusion

BphI-BphJ is an aldolase-dehydrogenase complex from *B. xenovorans* from the PCBs degradation pathway. The aldehyde produced from BphI is channeled directly to the active site of BphJ through an intermolecular tunnel with more than 90% efficiency as determined by a novel enzyme competition assay. Analysis of a crystal structure of an orthologous aldolase-dehydrogenase complex (DmpG-DmpF) enabled the identification of a route for the direct channeling of aldehyde products from the aldolase to the dehydrogenase. This route was confirmed by substituting residues located at a bottleneck in the channel with larger amino acids which sterically blocked the channel. Substitution of a glycine residue at this location limited the passage of larger aldehydes but not smaller aldehydes. Similarly, in the *Thermus* ortholog an alanine at this location led to more efficient acetaldehyde than propionaldehyde channeling. Replacement of this residue with glycine restored it ability to channel both aldehydes with high efficiencies. Chimeric *Bph-Thermus* enzyme complexes are still able to channel aldehyde, albeit at a drastically lower efficiency, suggesting that efficient aldehyde channeling requires the cognate partner enzymes.

The channel appears to be dynamic and therefore the dimensions of the channel derived from static crystal structure are inadequate in predicting the sizes of aldehydes that can diffuse through the channel. The coordinated conformational changes of specific gating residues, His-20 and Tyr-290, appear to ensure the efficient channeling of aldehydes. Substituting residues located at the exit of the channel in the aldehyde dehydrogenase with larger residues did not significantly impact channeling efficiently, implying these residues do not function as a gate. Interaction of Asn-170 with the nicotinamide cofactor in BphJ does not appear to be important.
for substrate channeling. However, Asn-170 and Ile-171 are important for allosteric activation of BphI.

The 3D-structure of the HsaF-HsaG complex was solved using X-ray crystallography. HsaF and HsaG differ from previously characterized aldolase-dehydrogenase complexes from other aromatic pathways in terms of its substrate specificity, allosteric activation and solubility of the aldolase. Like its distant homolog methylmalonyl-CoA reductase, HsaG contains an equivalent serine residue that is postulated to interact with the 3’ phosphate of CoA adenine ribose. This was confirmed by substituting the serine with larger or negatively charged residues resulting in 2000- to 6000-fold reduced catalytic efficiencies for CoA compared to the wild-type enzyme. However, the difference in specificities for NAD$^+$ and NADP$^+$ in the evolutionarily related methylmalonyl-CoA reductase and HsaG indicates that this residue is not the molecular determinant for nicotinamide coenzyme discrimination.

The results presented in this thesis elucidated the molecular determinates of substrate channeling in aldolase-dehydrogenase complexes, shed light on allosteric activation and characterized an ortholog from a previously unstudied clade.

6.2 Future Directions

Currently crystal structures from only two orthologs are available. Solving the crystal structure of a complex from each clade may provide structural evidence for the varying biochemical properties. In addition, the crystallization of each subunit individually would determine how the structure of the enzymes changes upon binding with a partner enzyme. These crystallographic studies may provide insight into why some complexes are active when expressed individually while others are not.
It is unclear what roles, if any, do the variable length loop in the dehydrogenases play in allosteric activation. Construction of variants with modified, extended or shortened loops may help clarify this question.

The ability of the BphI-BphJ complex to channel aldehydes of different lengths and sizes is another potential area that warrants further study. Judging from the size of the aldolase active site, the distal ends of long 4-hydroxy-2-oxoacids most likely protrude into the channel. Therefore it is possible that the aldehydes are being ‘pre-loaded’ into the channel and once released they would pass into the channel into the dehydrogenase. Solving a structure with a long chain 4-hydroxy-2-oxoacid bound in the active site of the aldolase will be an important step towards verification of this mechanism. Similarly, structures of aldehyde trapped in the channel may aid in elucidating how protein dynamics could alter the size of the channel and therefore accommodate the large aldehydes.

Similarly crystallographic structure analysis of the aldehyde dehydrogenase with bound nicotinamide and CoA and its comparison to the related methylmalonyl CoA reductase may shed light on how the Rossmann fold could bind CoA and yet discriminate between NAD⁺ and NADP⁺.

Finally, structure-function studies on Rv3469c, the mycobacterial enzyme that is misannotated as a 4-hydro-2-oxoacid aldolase may provide a basis to understand its relationship with the aldolases. For example, it will be interesting to substitute Tyr-22 with histidine to determine if this will be sufficient for aldolase activity.
References


gene cluster and substrate specificities of phenol hydroxylase and catechol 2,3-dioxygenase, *Microbiology* 149, 3265-3277.


