

Title Exploring the B. hominis TPI-GAPDH fusion enzyme glycolytic activity and stability

Name: Daniel McNerney

This is a digitised version of a dissertation submitted to the University of Bedfordshire.

It is available to view only.

This item is subject to copyright.

Exploring the *B. hominis* TPI-GAPDH fusion enzyme glycolytic activity and stability



Daniel McNerney

University of Bedfordshire

A thesis submitted to the University of Bedfordshire, in fulfilment of the requirements for the degree of Master of Science by Research

December 2018

Author's Declaration

I, Daniel McNerney declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University;
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- Where I have drawn on or cited the published work of others, this is always clearly attributed;
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- Where the thesis or any part of it is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- None of this work has been published before submission.

Abstract

Protist parasites cause deadly diseases and huge financial losses to crops. Eukaryotic organisms share similar cell structure and function. This makes novel drug targets difficult to find, and treating protist diseases very challenging. Studies have demonstrated a novel mitochondrial targeted glycolytically active TPI-GAPDH fusion enzyme as a potential drug target. This enzyme has been located in a variety of pathogenic Stramenopiles, including *B. hominis* and *P. infestans.* Additionally, the mitochondrial targeting of the enzyme suggests a retention of glycolysis originating from the mitochondrial precursor endosymbiont.

This study explores the glycolytic function and enzyme stability of the TPI-GAPDH fusion protein when compared to its segregated subunits. Enzymatic assays demonstrated that GAPDH functions more effectively as a part of the intact fusion enzyme, and that GAPDH may compete for GAP with TPI. This is contrary to the canonical understanding that TPI is a perfect diffusion rate-controlled enzyme that is not rate-limiting. Additional assay data suggests that L-GAP may interact with GAPDH as a competitive inhibitor of glycolytic function. Higher initial rates of enzymatic activity were observed by GAPDH and TPI-GAPDH when isomerically pure D-GAP was used as substrate compared to DL-GAP. Lastly, thermal shift data suggests that the TPI-GAPDH fusion enzyme may be far less stable than its subunits. TPI may have a destabilising effect on the GAPDH moiety of the fusion enzyme resulting in dissociation at approximately 25°C lower temperature than GAPDH by itself. Incubation of GAPDH and TPI subunits appeared to stabilise TPI and show potential evidence of protein-protein binding.

Acknowledgements

I would like to thank my supervisor Dr. Robin Maytum,

for investing a great deal of time with me,

and for giving me the opportunity to learn so much from him.

Table of Contents

Author's Declaration……...2

Abstract…….....3

Acknowledgements……...…4

Table of Contents……..…..…5

Chapter 1: Introduction……...9

1.1 Protist Parasites ……..9

1.1.1 Protist Classification……..…….9

1.1.2 Treating Protist Diseases……..……11

1.1.3 Protist Metabolism……12

1.2 TPI Structure and Function……..……..……14

1.2.1 TPI Structure…….……14

1.2.2 TPI Function…….……15

1.3 GAPDH Structure and Function……17

1.3.1 GAPDH Structure……17

1.3.2 GAPDH in Glycolysis……..……18

1.4 TPI-GAPDH Fusion Enzyme……...……19

1.4.1 Structure and Function of the Fusion Enzyme……...…19

1.4.2 Evolutionary Implications……21

1.5 Aims…….…22

Chapter 2: Methodology and Methods……..……23

2.1 Production of *S. cerevisiae* Proteins……23

2.1.1 Competent Cells……23

2.1.2 Plasmid DNA Purification……23

2.1.3 *S.cerevisiae* Genomic DNA Purification……24

2.1.4 Preparation of Gene Inserts……25

2.1.5 Digestion of Vector and Inserts……27

2.1.6 Ligation of Inserts into pJC20 Vector……27

2.1.7 Transformations……27

2.1.8 Screening of Clones……28

2.1.9 Sequencing of Clones……28

2.2 Overexpression and Preparation of Proteins……29

2.2.1 Overexpression of Protein in Recombinant *E. coli*……29

2.2.2 Protein Extraction……29

2.2.3 Purification of Polyhistidine-Tagged Proteins……29

2.2.4 Preservation of Proteins……30

2.3 Enzyme Assays……30

2.3.1 Preparation of Protein for Spectrophotometer Analysis……30

2.3.2 Adjusting Enzyme Concentration……31

2.3.3 Enzyme Assays……31

2.3.4 Equilibration Assays……32

2.3.5 D-GAP Isomer Assays……32

2.3.6 Active Site Inhibition Assays……32

2.4 Enzyme Stability……33

2.4.1 Thermal Shift Assays……33

2.4.2 MMTS Thermal Shift Assays……33

Chapter 3: Results……35

3.1 Production of Proteins……35

3.1.1 Production of Clones……35

3.1.2 Protein Expression and Purification……37

3.1.3 Analysis of Molecular Weights……42

3.2 Enzyme Assays……45

3.2.1 Calculating Enzyme Concentration……45

3.2.2 Establishing Enzyme Rates……46

3.2.3 Calculating V_{max} and K_m……51

3.2.4 Point Rate Comparisons……56

3.2.5 DL-GAP vs D-GAP Point Comparison……61

3.3 Thermal Shift Assays……64

3.3.1 TPI and GAPDH ……64

3.3.2 Fusion Protein and Subunit Combination……67

3.3.3 NAD^+ Stabilises the FusionProtein……68

3.4 Active Site Cysteine Inactivation……70

3.4.1 MMTS Inactivation Assays……70

3.4.2 MMTS Thermal Shift Assays……73

Chapter 4: Discussion……77

4.1 *B. hominis* TPI-GAPDH fusion enzyme catalyses the same glycolytic reactions as more common Enzyme subunits……77

4.2 Dividing *B. hominis* GAPDH and TPI moieties results in decreased catalytic activity……78

4.3 TPI competes with GAPDH for GAP……79

4.4 L-GAP may competitively inhibit GAPDH glycolytic function……80

4.5 TPI destabilises the TPI-GAPDH fusion protein……81

4.6 GAPDH appears to stabilise TPI and may show evidence of enzyme interaction……81

4.7 NAD^+ stabilises TPI-GAPDH fusion protein……81

4.8 MMTS destabilises but does not inactivate TPI proteins containing a cysteine at dimer interface……82

4.9 Future Work……83

Bibliography……85

Appendices……90

Chapter 1: Introduction

1.1 Protist Parasites

1.1.1 Protist Classification

The classic term “protist” refers to a polyphyletic group of eukaryotic organisms which do not belong to any of the three well-defined monophyletic kingdoms. This grouping has been artificially created as a convenience simply to refer to those eukaryotic organisms which are not plants, animals, or fungi (Wiser, 2011). As a result, the term “protist” covers an enormous diversity of life. Organisms within this group tend to have extensive differences from one another and are often more closely related to members of other kingdoms than fellow protists. Organisms in this loosely-defined kingdom can be unicellular or simple multicellular, auto or heterotrophic, mobile or stationary, asexual, sexual or both. Diversity of life ranging from kelp and algae, slime moulds, and plasmodium parasites which infect human red blood cells have historically all fallen under this umbrella-oddball kingdom (Embley and Martin, 2006). Many modern classification systems now separate protists into super groups along the eukaryote branch, abandoning the name “protist” as a scientific identity but continuing its use for convenience (Wiser, 2011).

Protist phylogeny has always been highly problematic. This has largely been as a result of determining relatedness of species strictly by morphological observation, which has many limitations (Tan, 2008; Embley and Martin, 2006). Reconstructing evolutionary events involves a considerable number of character changes, many of which are likely under selective pressure and subject to convergence (Sleator, 2013). When unrelated species develop analogous

structures, they can be misinterpreted as being shared due to common ancestry. There are many examples of non-related species repeating unique evolutionary innovations some of which include; the development of echolocation in whales and bats, the silk producing abilities of spiders, silk worms, silk moths, and weaver ants, and the development of wings in birds, bats, and butterflies.

The phenomena of gaining or losing similar evolutionary adaptations by unrelated species is known as homoplasy. When determining phylogenies homoplasies may result in Long Branch Attraction. This is a form of systematic error which causes distantly related lineages to be incorrectly interpreted as being much more closely related. When a number of molecular or morphological changes accumulate within a lineage, LBA can occur causing one lineage to appear similar and closely related to another lineage. Modern phylogenetics utilizes whole genome sequencing. As a result, protist phylogenies are constantly being updated and altered (Embley and Martin, 2006; Roger and Hug, 2006; Tamura et al., 2012).

The maturation of molecular phylogenetics has resulted in the abandonment of past erroneous morphological classification and given rise to the creation of new and more accurate phylogenies based on whole genome data. Modern classification now separates protists of the past in seven supergroups which also include plants, animals, and fungi. These are: Excavata, Stramenopiles, Alveolata, Rhizaria, Archaeplastida (includes plants), Amoebozoa, Opisthokonta (includes fungi and animals) (Roger and Hug, 2006; Tamura et al., 2012).

1.1.2 Treating Protist Diseases

Treating bacterial infections is relatively easy when compared to treating protist infections. Structures unique to prokaryotic organisms offer a range of viable drug targets which if targeted would not simultaneously harm the cells of a eukaryotic host (Madigan et al., 2009). Often times bacterial infections are treated with antibiotics which target these unique prokaryotic functions directly. Penicillin inhibits cell wall synthesis, chloramphenicol and tetracyclines inhibit protein synthesis, and quinolones inhibit nucleic acid synthesis.

Treating diseases caused by protists can be challenging as a result of similarities in eukaryotic cell structure and function to a host. Drugs which target and disrupt protist structures or biological functions will likely harm the host as well. Because of this, high toxicity is a major problem in treating diseases caused by protists in humans. Two common examples of protists causing disease are *Blastocystis hominis* and *Giardia lamblia*. These pathogenic organisms can cause severe gastrointestinal problems and giardia is capable of killing a human host (Omar, 2006; Tan, 2008). *B. hominis* can be visualised in the stool smples of an infected host as seen in figure 1.

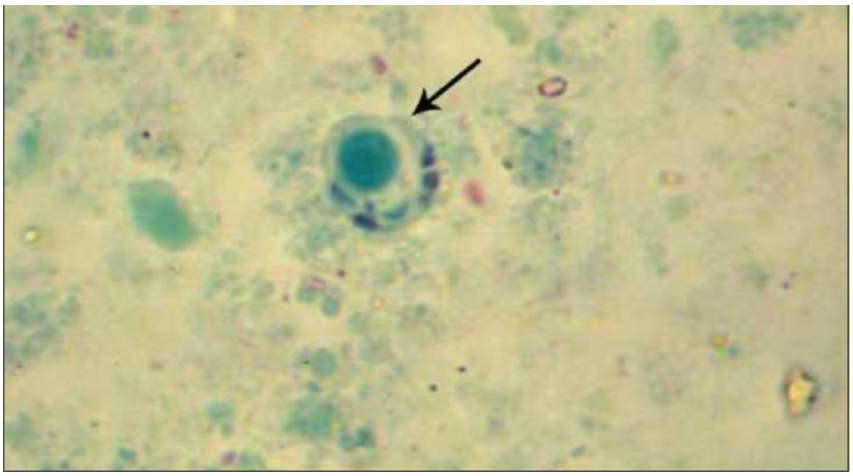


Figure 1: *Blastocystis* spp. stained with trichrome by ×1000 objective in a stool sample (Elghareeb, 2015)

This presents difficulty in finding good drug targets specific to these organisms which would not also harm the host. In comparison, there are a large variety of antibiotics to treat bacterial disease. These drugs commonly target structures specific to prokaryotes, allowing the majority of infections to be cured without causing toxic side effects or harm to the eukaryotic host (Madigan et al, 2009).

1.1.3 Protist Metabolism

The majority of eukaryotic organisms share most of the same basic metabolic pathways. Eukaryotic energy metabolism has always revolved around some form of carbon metabolism (Mulller et al., 2012). Compared to prokaryotes, very little has changed in the basic biochemistry of core metabolic pathways. Prokaryotes however are able to fit nearly any environmental niche, and as a result have evolved additional metabolic pathways in order to metabolize a far wider range of nutrients which are unavailable to eukaryotes (Der Giezen, 2011; Embley and Martin, 2006).

The widespread similarity of eukaryotic metabolism suggests that it is not normally a good drug target. However, understanding the metabolism of a pathogenic organism is always important as even a small difference can provide a potential effective drug target. Metabolic organelles which can be found in parasitic protists but are absent in multicellular eukaryotes have the potential to be viable drug targets. Hydrogenosomes are a reduced form of mitochondria use fermentation and substrate level phosphorylation to produce ATP (Jacobs et al., 2011; Michels et al., 2006). Research exploring the targeting of hydrogenosomes has determined that resveratrol inhibits growth and is cytotoxic to the cells (Mallo et al., 2013).

It has been discovered that within the protists, the group of stramenopiles contains organisms that possess a novel mitochondrial targeted glycolytically active TPI-GAPDH fusion enzyme. Included among these organisms are the pathogenic organisms *Blastocystis hominis*, which causes diarrhea in humans, and *Phytophthora infestans,* the causative agent of potato blight that was responsible for the Irish potato famine (Omar, 2006; Tan, 2008). This fusion enzyme is hugely significant in two ways. Firstly, it has implications for the origins of eukaryotic glycolysis. The mitochondrial location suggests a retention of glycolysis originating from the endosymbiotic alpha proteobacteria rather than the archael host. Secondly, the novel structure of this enzyme makes it a potential selective drug target as it differs from the separate enzymes found in higher eukaryotes.

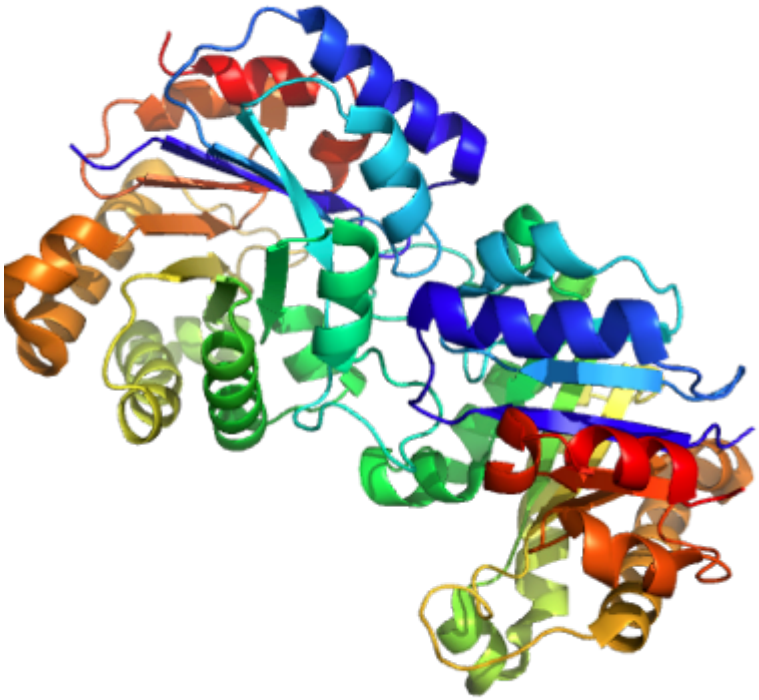
1.2 Structure and Function of TPI

1.2.1 TPI Structure

Triosephosphate isomerase (TPI) is a dimeric glycolytic enzyme which contains two subunits of 28kDa with the catalytic site located at the dimer interface. This enzyme is critical for glycolysis, a biological process present in virtually all living things. As such, it has been studied extensively since the 1930's (Chuang et al., 2005).

A reactive site cysteine residue found in several parasites at the TPI protein dimer interface has been suggested as a potentially viable drug target. Some examples of pathogens containing this cysteine residue include *Leishmania mexicana, Plasmodium falciparum, Trypanosoma brucei and Trypanosoma cruzi.* Derivatization with thiol-specific reagents have been shown to induce enzyme inactivation and aggregation for organisms with TPI enzymes containing this residue. More importantly, TPI varieties which are lacking this residue are mostly unaffected (Garza-Ramos et al., 1996). TPI function is critical for net production of ATP in glycolysis. This enzyme therefore may be a promising target for glycolytic pathogens.

Methyl Methanethiosulfonate (MMTS) has been shown to induce exposure of hydrophobic surfaces and cationic groups in TPI. Research has demonstrated that MMTS can inactivate TPI proteins by acting on this cysteine residue. MMTS reacts rapidly with exposed cysteine thiol groups inducing dissociation of the TPI dimer and inducing the formation of monomers (Ramos et al., 1996).



- Figure 2: Structure of *S. cerevisiae* TPI dimer at 1.9-angstroms resolution
DOI: 10.2210/pdb1YPI/pdb

1.2.2 TPI Function

During glycolysis, the enzyme aldolase splits fructose-1,6-bisphosphate into two sugars which are isomers of one another; Glyceraldehyde 3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP). TPI catalyses the rapid interconversion between DHAP and GAP molecules at a thermodynamic equilibrium of 22:1. TPI is considered to be a “perfect” diffusion rate-controlled enzyme and as such is not rate-limiting. This means that it converts the GAP produced in the previous step into DHAP before the GAPDH enzyme is able to utilise it (Knowles and Albery,1977; Wierenga et al., 2010). DHAP is not catalysed by GAPDH, but only the GAP molecules continue in the glycolytic process. TPI is emphasized as a vital enzyme by the fact that it is the only glycolytic enzyme for which a deficiency is lethal (Orosz et al., 2009).

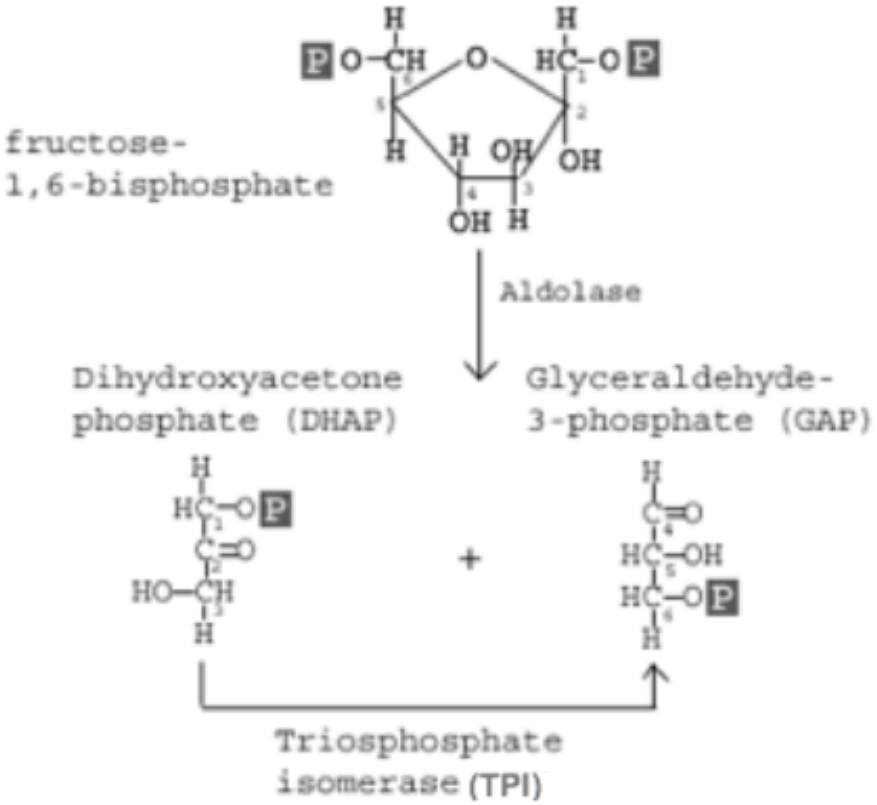


Figure 3: Aldolase splits fructose 1,6 bisphosphate into two sugars which are isomers of one another; DHAP and GAP. TPI rapidly inter-converts DHAP and GAP molecules. Only GAP is used in the next step of glycolysis (Aryal, S 2015).

1.3 GAPDH Structure and Function

1.3.1 GAPDH Structure

GAPDH is a highly conserved enzyme of around 335 amino acids in length. It contains two major domains: The NAD^+ binding domain (1-150), and catalytic or GAP domain (151-335) which contains the active site cysteine (Sirover, 2014).

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) has been classically defined as a glycolytic enzyme located in the cytosol of cells (Oesper, 1954; Trentham, 1968). Under normal conditions GAPDH exists primarily as a homotetramer, an oligomer formed from four identical monomers. It is composed of four identical 36kDa subunits each containing a single catalytic site which is divided into two binding domains for two separate molecules: the substrate glyceraldehyde-3-phosphate (GAP) and the coenzyme Nicotinamide adenine dinucleotide (NAD^+) (Figure 4b). Together the two binding domains form a single catalytic site (Moras et al., 1975).

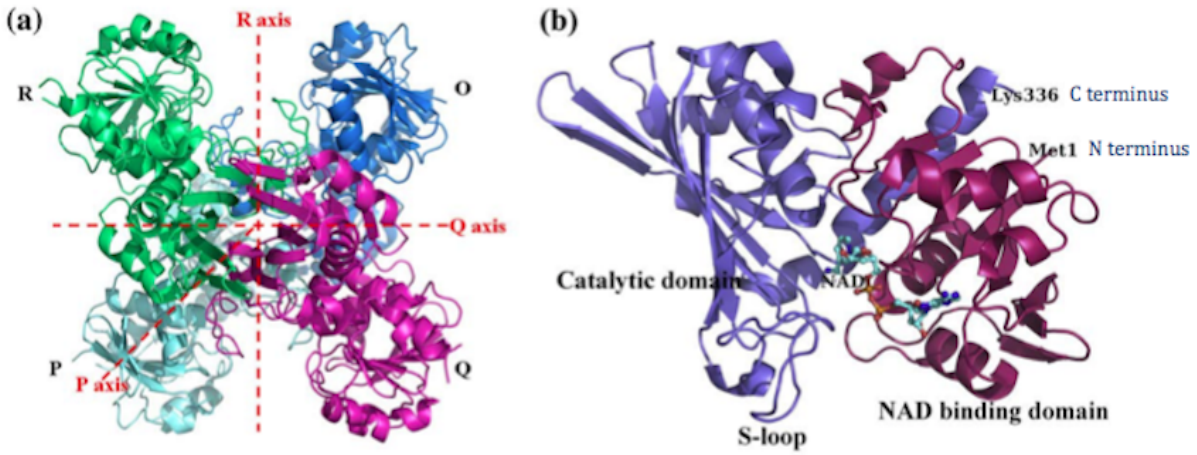


Figure 4: (a) Crystal Structure of tetrameric GAPDH from MRSA. (b) The catalytic domain (purple) binds the substrate and forms the catalytic site with the NAD^+ binding domain (pink) The N terminus at the methionine and the C terminus at the lysine are also highlighted. From Mukherjee *et al*. 2010.

1.3.2 GAPDH in Glycolysis

Along the glycolytic pathway GAPDH converts GAP into 1,3 bisphosphoglycerate (1,3 BPG) in a redox reaction. This step is critical as the product 1,3 BPG will lose a phosphate group to phosphorylate an ADP molecule, yielding the first ATP along the pathway. GAPDH also produces two NADH + H^+ from NAD^+ for each molecule of glucose. In aerobic respiration, this NADH is transported to the mitochondria where it takes part in more ATP synthesis by oxidative phosphorylation and is oxidized again to NAD^+ for use again. However, in organisms that are amitochondriate, have reduced mitochondria remnants, or go through periods of oxygen

starvation; NAD^+ is regained through fermentation reaction that oxidizes NADH + H^+ into NAD^+ (Karp, 2009).

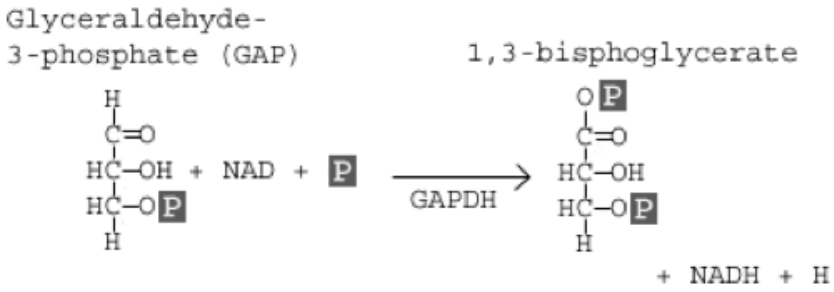


Figure 5: GAPDH adds an inorganic phosphate to GAP producing 1, 3 BPG. (Aryan,S 2015)

1.4 TPI-GAPDH Fusion Enzyme

1.4.1 Structure and Function of the Fusion Enzyme

The structure of the TPI-GAPDH fusion protein has recently been determined using x-ray crystallography. As shown in Figure 5 below, the tetrameric GAPDH core is made from two fused subunits bound together by beta sheet interactions (highlighted in white), then by comparatively weaker interactions binding the two dimers into a tetramer. The TPI dimers make up two subunits at the top and bottom (Liaud et al., 2000).

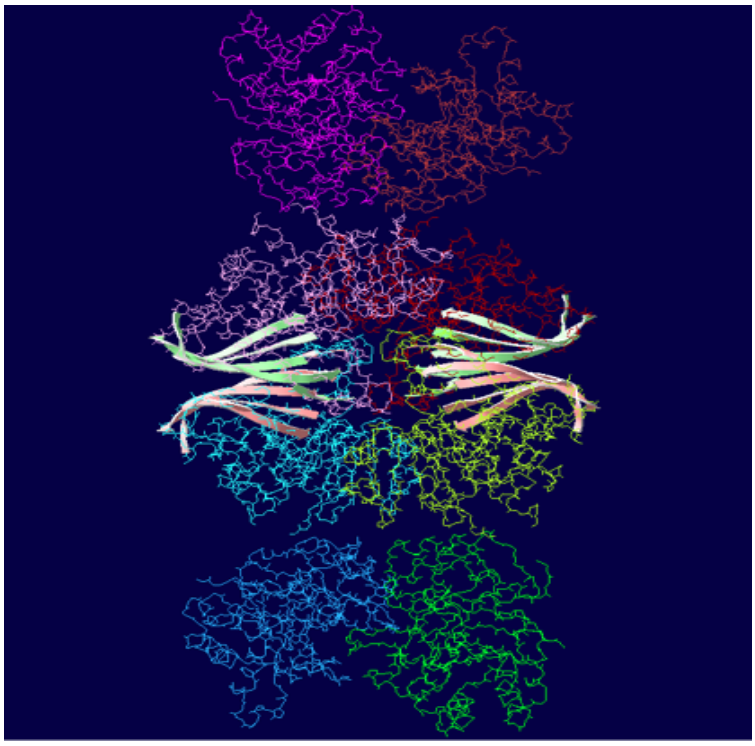


Figure 6: Structure of TPI-GAPDH fusion from *B. hominis* (van derGiezen, Maytum unpublished data). Fused TPI-GAPDH subunits are shown in pseudo-matching colours, with pairs of TPI dimers made from two subunits at the top and bottom. The GAPDH core is then made from two fused subunits bound together by the beta-sheets highlighted in white. The C-termini of the TPI dimers align perfectly with the N-termini of the matching GAPDH subunits.

B. hominis contains complete TPI moieties upstream of all three GAPDH coding regions (Nakayama et al., 2012). Interestingly, genomic analysis has also uncovered a mitochondrial targeting pre-sequence as part of the fusion protein that has more sequence homology to the *alpha-proteobacterial* sequence than to anything in the cytosol. It is believed that TPI-GAPDH undertakes the same function of TPI and GAPDH in glycolysis (Liaud et al., 2000)

This fusion-protein structure evidences that the individual TPI and GAPDH components are highly homologous to known structures of the proteins in other eukaryotes ranging from single celled *S. cerevisiae* to *H. sapiens.* The core GAPDH tetramer and the TPI dimers are essentially

identical to the structures known for individual non-fusion enzymes. In some regards this is unsurprising as these vital glycolytic enzymes must have evolved from early life and are highly conserved. However, it is astounding that the resolved structure of the fusion protein evidences perfect structural alignment of the TPI dimers to the GDPDH core, in which the C-termini of the TPI dimers align perfectly with the N-termini of the matching GAPDH subunits (Der Giezen, 2011; Liaud et al., 2000).

1.4.2 Evolutionary Implications

The mitochondrial targeting signals found upstream of the TPI-GAPDH fusion enzyme suggest a possible mitochondrial origin for glycolysis. Some phosphoglycerate kinase enzymes in *Stramenopiles* have also been discovered with mitochondrial targeting signals (Nakayama et al., 2012). The broad distribution of these signals in *Stramenopiles* suggest that the TPI-GAPDH fusion protein would have been present in a common ancestral species. This may suggest the loss of the singular fusion enzyme to becoming two separate proteins over time. This has significant evolutionary implications as it suggests that all TPI and GAPDH proteins are derived from an ancestral fusion protein. It has been suggested that the origin for aerobic respiration in eukaryotes is mitochondrial, and therefore alpha-proteobacterial. These genes could have been transferred to the host after endosymbiosis. Phylogenetic analysis performed on glycolytic enzymes in *Stramenopiles* has discovered that they are more closely resembled to the alpha-proteobacteria than any other organism. Traditionally it has been believed that glycolysis belonged to the host organism, however this information would suggest that it belonged to the endosymbiont (Der Giezen, 2011; Embley and Martin, 2006; Liaud et al., 2000).

1.5 Aims

The TPI-GAPDH fusion protein is a novel enzyme which has not been studied much to date. As a result of this little has been done to explore the structural effect of the TPI-GAPDH fusion protein on glycolytic function and protein stability when compared to its segregated subunits. The aims of this research are:

- To compare the glycolytic activity of the fusion protein to that of its individual subunits
- To clone and express *S. cerevisiae* TPI and GAPDH for use in diagnostic studies
- To further understand the structure / function relationship between the TPI and GAPDH fusion enzyme moieties

Chapter 2: Methodology and Methods

2.1 Production of Proteins

2.1.1 Competent Cells

BL21 DE3 pLysS *Escherichia coli* and XL-1 Blue *E. coli* competent cell stock preparation was performed as outlined by (Nishimura et al., 1990).

2.1.2 Plasmid DNA Purification

A 50ml stock of XL-1 Blue *E. coli* containing the pJC20 plasmid was grown overnight in 0.1mg/ml LB-Amp and centrifuged at 5000g for 10 minutes, the supernatant was discarded. The pellet was resuspended in 250µl resuspension buffer (30mM Tris HCl, 10mM EDTA, 100ug/ml Rnase A, pH 8.0), and transferred to a 2ml tube. 250µl lysis buffer (0.2M NaOH, 1% SDS) was added, the tube gently inverted 5 times, and left on the the bench for 3 minutes. 350µl neutralisation buffer (4M guanidine hydrochloride, 0.5M potassium acetate, pH 4.2) was added, and the tube immediately gently inverted 5 times. The mix was centrifuged at 16,500g for 10 minutes, and the supernatant transferred into a plasmid purification miniprep spin column (NBSBio). This was centrifuged at 5000g for 1 minute and the supernatant discarded. 500µl wash buffer (20mM NaCl, 2mM Tris HCl, 70% ethanol, pH 7.5) was added to the column and centrifuged for 1 minute at 5000g, and the supernatant discarded. This step was then repeated. The empty tube was spun at 16500g for 2 minutes to dry. The column was placed in a microcentrifuge tube to elute. 50µl ultrapure water (UPW) was added to the column, allowed to stand for 1 minute, and centrifuged at 5000g for 1 minute. The DNA was measured using a

nanodrop spectrophotometer and run on a 1% agarose gel containing 1µg/ml ethidium bromide at 70 volts to visualise concentration.

2.1.3 *S. cerevisiae* Genomic DNA Purification

100µl *S. cerevisiae* overnight culture was pelleted at 5000g for 2 minutes and the supernatant removed. The pellet was resuspended in 200µl lysis buffer (10mM Tris pH 8.0, 1mM EDTA, 100mM NaCl, 1% SDS, 2% Triton X-100). 300µl glass beads and 200µl 1:1 phenol:chloroform was added. Cells were vortexed for 3 minutes. 200µl TE (10mM Tris pH 8.0, 1mM EDTA) was added. The mixture was centrifuged at 15000g for 10 minutes. The supernatant was transferred to a fresh tube. 900µl 95% ethanol was added, then centrifuged at 15000 for 10 minutes. The pellet was washed with 70% ethanol and centrifuged at 15000g for another 10 minutes. The supernatant was removed and the tube was dried. The pellet was dissolved in 100µl TE (Lõoke et al., 2011).

2.1.4 Preparation of Gene Inserts

The primers used to generate gene inserts of interest are listed in the table below.

Table 1: Primers for amplification of *S. cerevisiae* TPI1, TDH1, TDH2, TDH3 genes

Primer No.	Description	Length	Sequence
RMM117	Sc TPI 5'	29	ggccatATGGCTAGAACTTTCTTTGTCGG
RMM118	Sc TPI 3'	37	gcgggatccTTAGTTTCTAGAGTTGATGATATCAACA
RMM119	Sc TDH1 5'	33	ggccatATGATCAGAATTGCTATTAACGGTTTC
RMM120	Sc TDH1 3'	31	gcgggatccTTAAGCCTTGGCAACATATTCG
RMM121	Sc TDH2 5'	32	ggccatATGGTTAGAGTTGCTATTAACGGTTT
RMM122	Sc TDH2 3'	29	gcgggatccTTAAGCCTTGGCAACGTGTT
RMM123	Sc TDH3 5'	33	ggccatATGGTTAGAGTTGCTATTAACGGTTTC
RMM124	Sc TDH3 3'	29	gcgggatccTTAAGCCTTGGCAACGTGTT

A 400µl PCR master-mix was created as follows: 350µl UPW, 40µl Taq Buffer, 12µl dNTPs, 8µl 1µg/ul *S. cerevisiae* gDNA . 100µl of the mix was transferred into four PCR tubes and 3µl of Forward and Reverse primer added to respective tubes at a final concentration of 10µm. The PCR parameters were as follows:

Table 2: PCR parameters for amplification of *S. cerevisiae* gene inserts

Cycle Step	Temperature (°C)	Time	Number of cycles
Initial Denaturation	95	4 minutes	1
Denaturation Anneal Extension	95 50 72	20 seconds 20 seconds 40 seconds	30 cycles
Final Extension	72	4 minutes	1

The amplified inserts were run on a 1% agarose gel, visualised under UV light, and the bands at 750bp (TPI1), and 999bp (TDH1-3) cut out and purified using perfect prep gel cleanup kit as follows: The gel fragments were weighed and dissolved in 3x weight gel cleanup buffer. 1x isopropanol was added and the solution was vortexed. Each mixture was spun at 5000g through gel cleanup columns with 700µl at a time, the flow through was discarded. When complete the tubes were spun at 16500g for 2 minutes to dry. The columns were transferred into microcentrifuge tubes and 30µl UPW added to the centre of the columns, let stand for 1 min, spun at 5000g for 1 min, another 20µl added to the centre of the columns, let stand for another minute, and spin again at 5000g for 1 min.

2.1.5 Digestion of Vector and Inserts

A double restriction digest was performed on both the pJC20 vector and *S. cerevisiae* gene inserts using BamH1 and Nde1. Reactions were performed with 50µl template plasmid DNA, 2µl BamH1, 2µl Nde1, and 6µl FD green buffer were spun down in microfuge tubes, stirred, then incubated at 37°C for 30 min. The digested inserts and vector were gel purified, cut out, and eluted as outlined in section 2.1.4.

2.1.6 Ligation of Inserts into pJC20 Vector

The positive control was performed using an unknown gene contained within the pJC20 vector prior to digest and purification.. The Nde1 and BamH1 plasmid double digest ensured that the ends of the insert were correctly cut and compatible for ligation. The negative control was performed by omitting DNA insert from the ligation mix. Upon transformation this would evidence if the empty vector ligated to itself.

2.1.7 Transformations

10µl ligation mix was added to 100µl competent cells on ice and mixed. The mixture was incubated on ice for 20 minutes, heat shocked at 42.5°C for 90 sec, then placed back on ice for 5 min. 900µl LB was added and the cells were incubated at 37°C for 60 minutes. The tubes were spun at 500g for 2 minutes and the supernatant was poured off except for the last drop. The cells were gently resuspended and plated onto 0.1mg/ml LB-Ampicillin (Sigma) agar and grown overnight at 37°C.

2.1.8 Screening of Clones

Clones were screened by PCR. Colonies expressing ampicillin resistance were used as a PCR template and T7 primers were used to duplicate the DNA present between the T7 promoter and the T7 terminator. The PCR product was analysed using 1% agarose gel electrophoresis and bands visualised under UV light. A background self-ligated plasmid negative would generate an insert of approximately 150bp.

2.1.9 Sequencing of Clones

A potentially positive clone which expressed ampicillin resistance was grown overnight and purified according to the plasmid DNA purification protocol outlined in section 2.1.2. These plasmids were digested and run on gel-electrophoresis to visualise insert size. Additionally clones were screened by PCR. Clones which appeared to contain a gene of interest were to be sequenced were sent off to the University of Cambridge Department of Biochemistry for DNA sequencing. The clone Samples were eluted in UPW and prepared at a concentration of 25ng/µl. A Sanger DNA Sequencing Request Form was completed with Abs 260/280 information and sent off along with 5µl 10µm forward and reverse primers. The sequencing results were returned in a FastA format and analysed.

2.2 Overexpression and Preparation of Proteins

2.2.1 Overexpression of Protein in Recombinant *E. coli*

1L of LB media was supplemented with 0.1mg/ml ampiliccin and inoculated with a colony of recombinant BL21 DE3 and incubated at 30°C until A600 OD 0.6 was reached. The culture was induced by the addition of 0.1mg/ml IPTG (isopropyl-β-D-thiogalactoside) and incubated at 30°C for 14 hours overnight. The culture was split into two 500ml portions and centrifuged at 8000rpm for 10 minutes to pellet the cells. The supernatant was discarded. Pellets were aliquoted into falcon tubes and frozen for later use.

2.2.2 Protein Extraction

Frozen cells were thawed and resuspended in 10ml cold LEW buffer (50mM NaH2PO4, 300mM NaCl, pH 8.0). The solution was put on ice and lysed by sonication at 6.25nm cycling 90 seconds on and 60 seconds off, for 4 cycles. Lysed cells were centrifuged at 15000rpm for 10 minutes at 4°C.

2.2.3 Purification of Polyhistidine-Tagged Proteins

Soluble protein was purified from the supernatant by affinity chromatography using Protino Ni-IDA Resin columns. 2ml LEW was filtered through the column for equilibration. The supernatant was passed through the column at a rate of less than 1 ml/min and collected in 2ml fractions. The column was washed with 4ml LEW buffer collected in two fractions. The protein was eluted with 5ml of elution buffer (LEW buffer, 250mM imidazole, pH 8.0) collected in five 1 ml fractions.

The elutions were immediately placed on ice and tested for presence and concentration of GAPDH by running 10µl of each on a 13.5% SDS-PAGE gel.

2.2.4 Preservation of Proteins

Protein in the elution fractions was salted out of solution by the addition of ammonium sulphate to a final concentration of 0.4g/ml. Precipitated protein stocks were stored at 4°C until needed.

2.3 Enzyme Assays

2.3.1 Preparation of Protein for Spectrophotometer Analysis

500µl protein stock was transferred into an eppendorf tube, and pelleted at 8500g for 2 minutes. The supernatant was carefully removed and discarded. The protein pellet was dissolved in 150µl cold buffer (15mM Sodium Phosphate pH 8.5) by vortexing. The solution was transferred into a micro-dialysis cap and covered with a single layer of dialysis tubing. Any remaining ammonium sulphate in the protein solution was dialysed out of solution for a period of 6 hours in 500ml of buffer (15mM Sodium Phosphate pH 8.5). The micro-dialysis sample was placed in 500ml fresh buffer for another 6 hours. Once dialysed, the protein solution was transferred into a fresh eppendorf tube. Insoluble protein was pelleted from the solution by centrifugation at 10,000g for 5 minutes. The supernatant containing usable dissolved soluble protein was transferred into a fresh tube.

2.3.2 Adjusting Enzyme Concentration

Protein concentration was calculated by measuring the OD between 220nm and 350nm using a nanodrop spectrophotometer. The absorbance at 280nm value was used in order to calculate concentration using the following equation: Concentration(mM) = (A÷E÷MW) x 10^6 Where A = absorbance at 280nm, E = extinction coefficient and MW = molecular weight (as predicted by ProtParam). (http://web.expasy.org/protparam/), the data for each protein can be found in Appendices. All enzyme concentrations were adjusted to 15μM for use in assays by diluting with arsenate buffer.

2.3.3 Enzyme Assays

Rates of enzyme activity were established by enzymatic assays. When using GAP as substrate the assay mix contained 260μl 15mM sodium pyrophosphate arsenate buffer, 10μl 7.5mM NAD^+ (Sigma N7004), 10μl 0.1M dithiothreitol (DTT) (Calbiochem) and 10μl 15μM enzyme. A blank of the assay mix was established at 340nm on a spectrophotometer prior to starting. The assays were set up for 3 minutes of data collection taking readings at 0.25-second intervals. The assays were started by addition of 10μl 7.5mM DL-GAP (Sigma G5251) at the stated concentrations.

When using DHAP as substrate the assay mix contained 250μl-260μl 15mM sodium pyrophosphate arsenate buffer depending on the volume of protein. The assay mix contained 10μl 7.5mM NAD^+ (Sigma N7004), 10μl 0.1M dithiothreitol (DTT) (Calbiochem) and 10μl of each 15μM enzyme- Either10μl Fusion protein or 10μl GAPDH His C and 10μl TPI, for a total assay volume of 300μl. A blank of the assay mix was established at 340nm on a spectrophotometer

prior to starting. The assays were started by addition of 10µl DHAP (Sigma 51269) at stated concentrations.

2.3.4 Equilibration Assays

Equilibration assays were performed as outlined in section 2.3.3 with the exception of beginning the reaction with NAD^+. The assay was started by addition of 10µl 7.5mM NAD^+ (Sigma N7004-1G).

2.3.5 D-GAP Isomer Assays

Assays were performed similarly to the equilibration assays as outlined in section 2.3.4. All D-GAP assays were performed at stated concentrations and compared to initial rates under the same condition when using DL-GAP as substrate.

2.3.6 Active Site Inhibition Assays

30µl of 15µM enzyme stock was mixed with 3µl 10mM MMTS for a final concentration of 1mM MMTS. The stocks were incubated in the fridge at 4°C for 120 minutes, then assayed for any change in catalytic activity. Assays were performed as outlined in section 2.3.3 and measured against controls.

2.4 Enzyme Stability

2.4.1 Thermal Shift Assays

22µl reactions were set up in triplicates in PCR tube strips containing 19µl of 15mM protein, 1µl 15x SYPRO Orange dye, and either 2µl UPW or 2µl 7.5mM NAD^+, and spun down to remove bubbles. Proteins were cycled in a PrimeQ Real-time Nucleic Acid Detection System using the FCO3 filter (excitation at 520nm and emission at 560nm) according to the table below:

Table 3: Thermal shift assay parameters

Step #	Temp°C	Cycles	Hold time	Read Time	Increments
1	20-90	141	30sec	30sec	0.5°C

2.4.2 MMTS Thermal Shift Assays

Reactions were set up as outlined in section 2.4.1 without the addition of NAD^+. Samples contained either 2µl 10mM MMTS (1mM final concentration) or 2µl UPW for a total reaction volume of 22µl. Table 4 illustrates a longer hold time for these experiments in order to more gradually dissociate the enzymes, resulting in higher resolution data.

Table 4: PrimeQ PCR Fluorescence Spectroscopy MMTS Assay Parameters

Step #	Temp°C	Cycles	Hold time	Read Time	Increments
1	20	10	60sec	30sec	N/A
2	30	60	60sec	30sec	N/A
3	30-90	121	30sec	30sec	0.5°C

Section 3: Results

3.1 Production of Proteins

3.1.1 Production of Clones

There were several clones generated which expressed ampicillin resistance and demonstrated corresponding band weights when digested and run on a polyacrylamide gel. However, PCR and sequence analysis of plasmids revealed large concatemers had formed during the ligation process, and none of the four *S. cerevisiae* clones were ultimately generated. This was disappointing as the more common naturally occurring independent TPI and GAPDH enzymes were intended for a comparative study with the segregated components of a fusion protein.

25ng/µl samples of the DNA inserts were sent to the Department of Biochemistry at Cambridge University for Sanger DNA sequencing. The results showed the expected sequence suggesting that the inserts were fit for cloning. Ligation attempts were attempted using several protocols with varying concentrations of vector, insert and ligase. Transformation attempts always produced viable clones for the positive controls (40-400 clones), evidencing an adequate transformation protocol.

Clones expressing ampicillin resistance were grown on a small scale and induced with IPTG for protein expression, but no potential transformants expressed. On several occasions these non-expressing clones were sequenced and evidenced concatomers and multiple vector portions ligated together. Dephosphorylation of vector prior to ligations did not return successful production of clones.

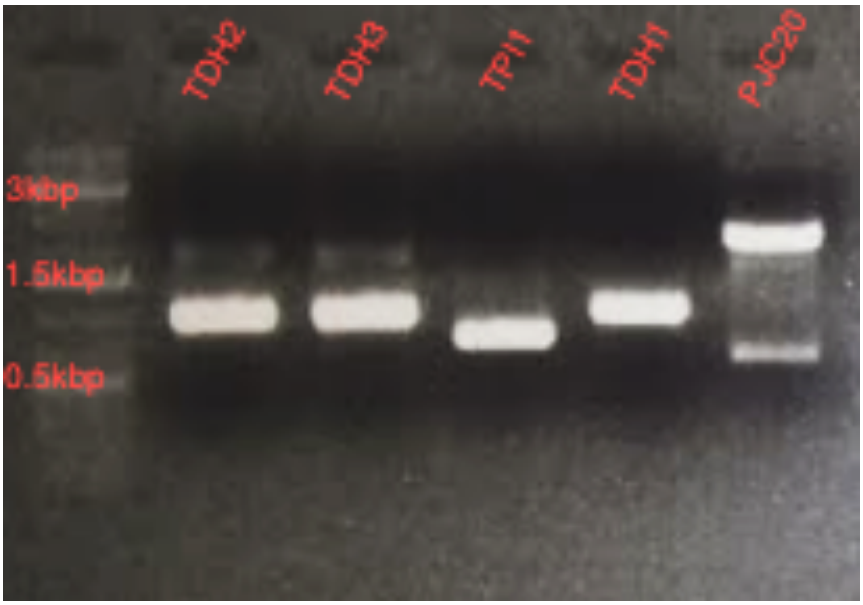


Figure 7: Gene inserts run on an agarose gel represent the correct sizes expected for the genes. The TPI1 gene insert is approximately 747bp, the TDH 1,2, 3 genes are approximately 999bp in length. Here the digested pJC20 plasmid band can be clearly seen at approximately 2.3 kbp and the old gene is seen removed at about 700bp.

Cloning of the *S. cerevisiae* TPI1 and TDH1,2,3 genes was unsuccessful. As a result the Saccharomyces TPI protein was purchased through Sigma (product code **T2507**). Stocks of BL21 DE3 pLysS clones containing *B. hominis* TPI, GAPDH, and Fusion protein inserts were available from a previous study, and were retrieved from a -80C freezer.

3.1.2 Protein Expression and Purification

Large-scale protein expression was performed upon successful small-scale expression. The pJC20 and PET14b vectors both utilise the T7 promoter system in order to regulate protein expression (figure 8). The gene of interest is ligated into the vector between the T7 promoter region and the T7 terminator region. This gene is maximally expressed when T7 RNA polymerase is produced upon induction of the BL21 DE3 plysS. The λDE3 lysogen which carries the gene encoding for T7 RNA polymerase is under the control of the lacUV5 promoter (Studier et al. 1990). The addition of IPTG (a non-hydrolysable lactose analogue) induces expression of T7 RNA polymerase, which binds to the T7 promoter region and transcribes the target gene. T7 promoter systems are notorious for being “leaky”. They occasionally minimally express genes under the control of the T7 promoter without intentional induction. As a countermeasure the pLysS plasmid carries the gene encoding T7 lysozyme as well, which lowers background expression levels (Marbach and Bettenbrock, 2012).



Figure 8: pJC20 vector map. DNA inserts were ligated between Nde1 and BamH1 restriction sites (Evans, 2014)

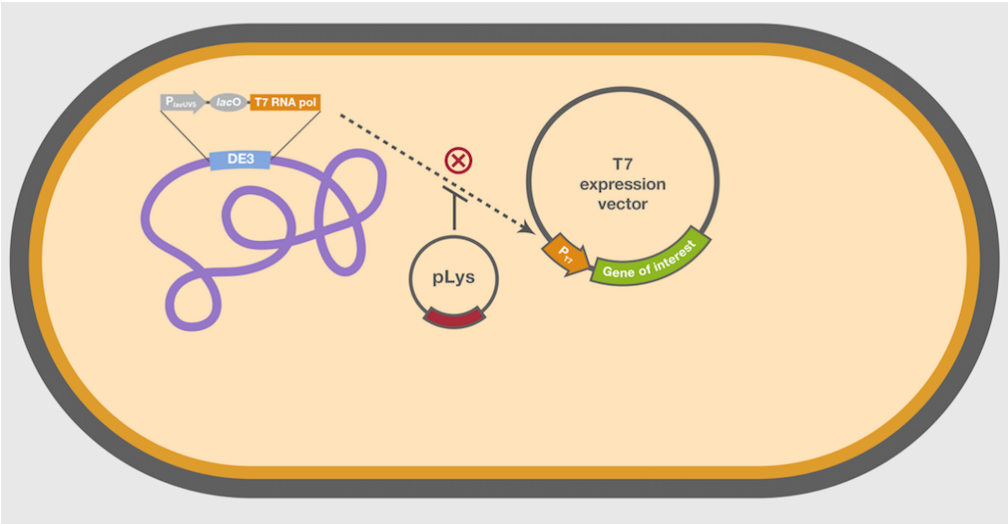


Figure 9: BL21(DE3) *E.coli* cell containing the pLysS plasmid (ThermoFisher).

Proteins of interest were isolated and purified from cell lysates utilising nickel affinity chromatography. The polyhistidine- tag on the proteins allowed them to bind to the Protino Ni-IDA Resin column when the lysate was run through. The imidazole side chain of histidine has a specific binding affinity for metal ions, in this case Ni^{2+} contained in the dry silica resin of the column. As a result, the proteins of interest were bound tightly to the resin while the rest of the lysate flowed through into waste fractions. The column was washed with LEW buffer in order to flush anything remaining in the column which had not bound. The desired proteins were eluted with a buffer containing imidazole which binds to the metal ions and displaces the proteins.

Fractions were collected at each stage of the expression and purification processes and were analysed on a 13.5% SDS PAGE gel (Figure 10). A band is visible in the induced cell sample between the 26 kDa and 36 kDa markers which is not present in the uninduced cell sample. This band represents the correct size of the Blastocystis TPI protein expected. The same band is much more pronounced in the lysate sample taken prior to passing through the affinity column. The lysate and flow-through samples represent the lysate before and after passing through the affinity purification column. Once the lysate has passed through the column the protein of interest is no longer present, but all other protein bands remain constant. This evidences that the protein has bound to the column while all other undesired proteins contained in the lysate have not. The wash flushed any remaining unbound proteins from the column. The protein that had bound to the column reappears in the elution fractions which also contain far less contaminants. The progression from a large to small band in the elution fractions confirm that the polyhistidine-tagged protein has been fully displaced and eluted.

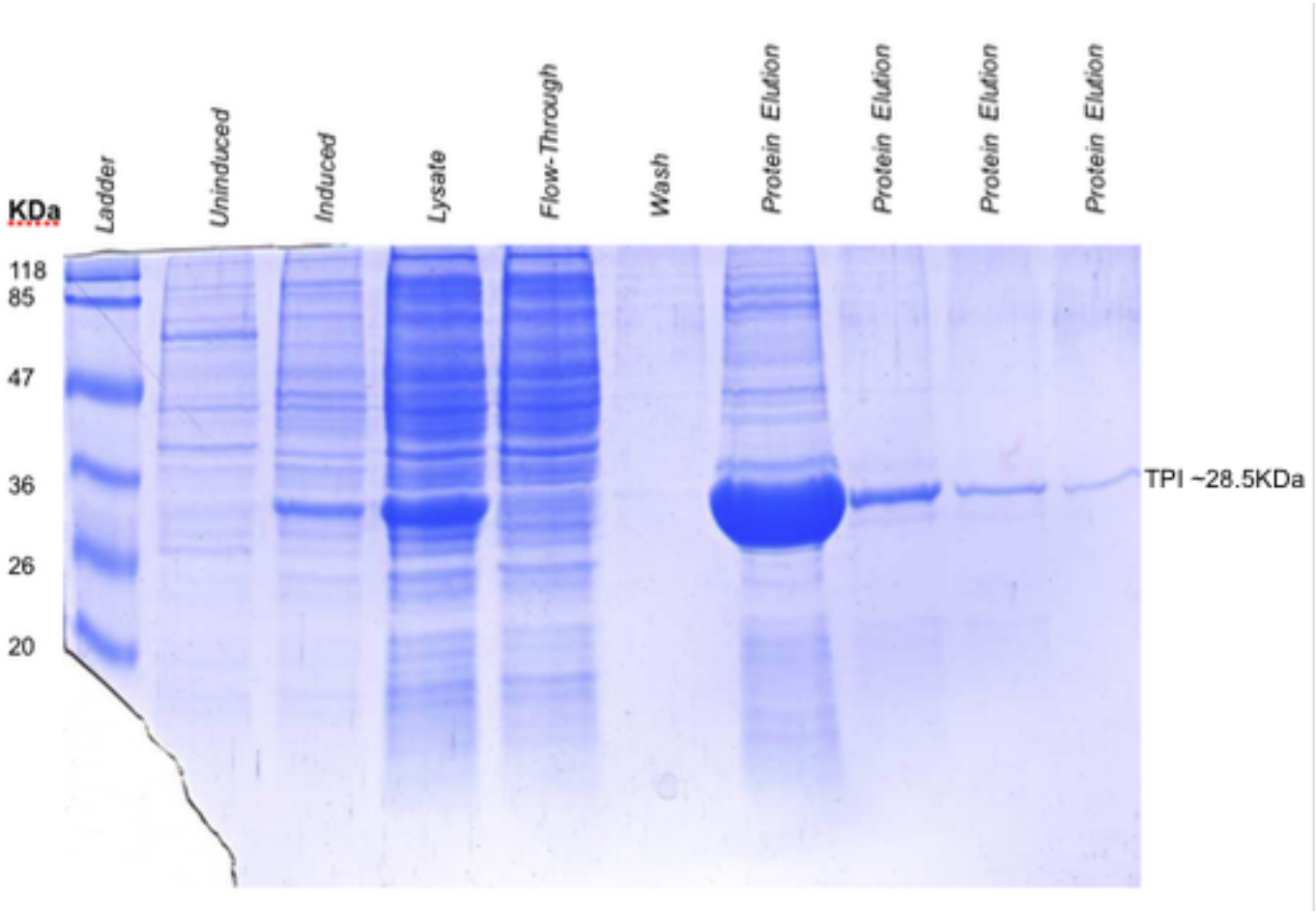


Figure 10: A 13.5% agarose SDS-PAGE gel showing each step in the protein expression and purification process. The uninduced sample represents normal protein levels. The IPTG induced sample contains the *B. hominis* TPI protein demonstrated by the new band of around ~28.5KDa.

All other proteins were expressed and purified in the same fashion. Figure 11 depicts a 13.5% SDS PAGE gel performed with crude cell lysates which contain a large quantity of protein of the respective expected sizes.

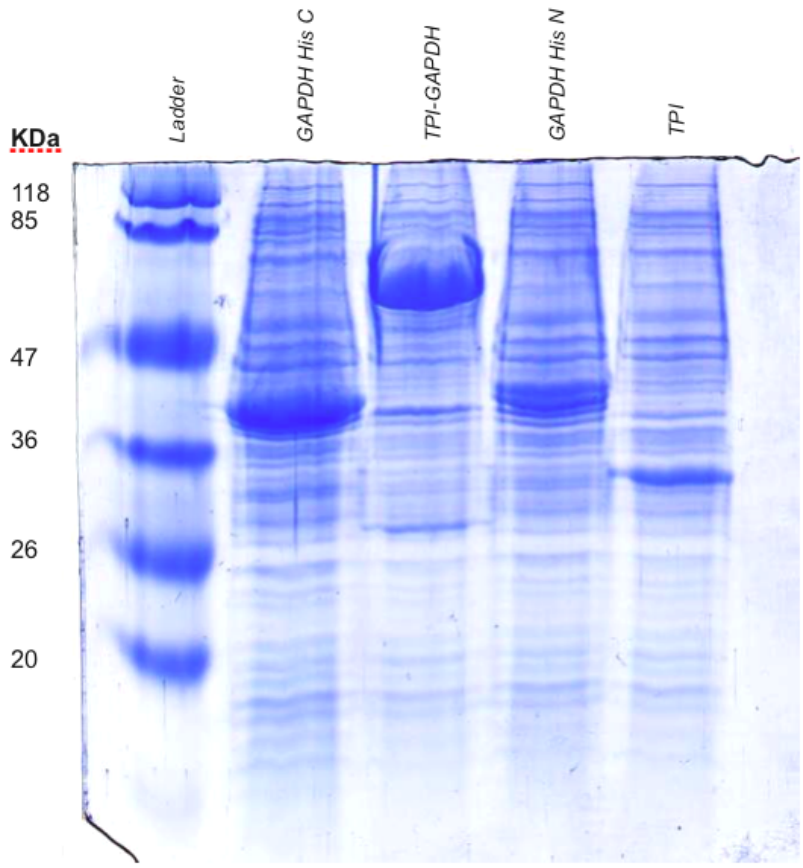


Figure 11: 13.5% SDS-PAGE Gel showing lysates containing proteins of interest. Bands of expected sizes for GAPDH His C = 38.7kDa, TPI-GAPDH= 65kDa, GAPDH His N= 38.7kDa, and TPI= 28.5kDa are pronounced in each respective sample. Plenty of unwanted protein from cell debris is also visible.

3.1.3 Analysis of Molecular Weights

Ammonium sulphate preparations were re-dissolved in 15mM tetrasodium pyrophosphate buffer and run on a 13.5% SDS PAGE GEL. ImageJ was used to analyse the distances travelled by the denatured protein samples in order to determine their molecular weight. Data points were generated by measuring the distance travelled by each protein sample. The Fisher EZ-Run Protein Ladder provides bands of known sizes which can be analysed and plotted as a function of distance travelled. The equation for a line of best fit was generated and the data points for each unknown were plugged in.

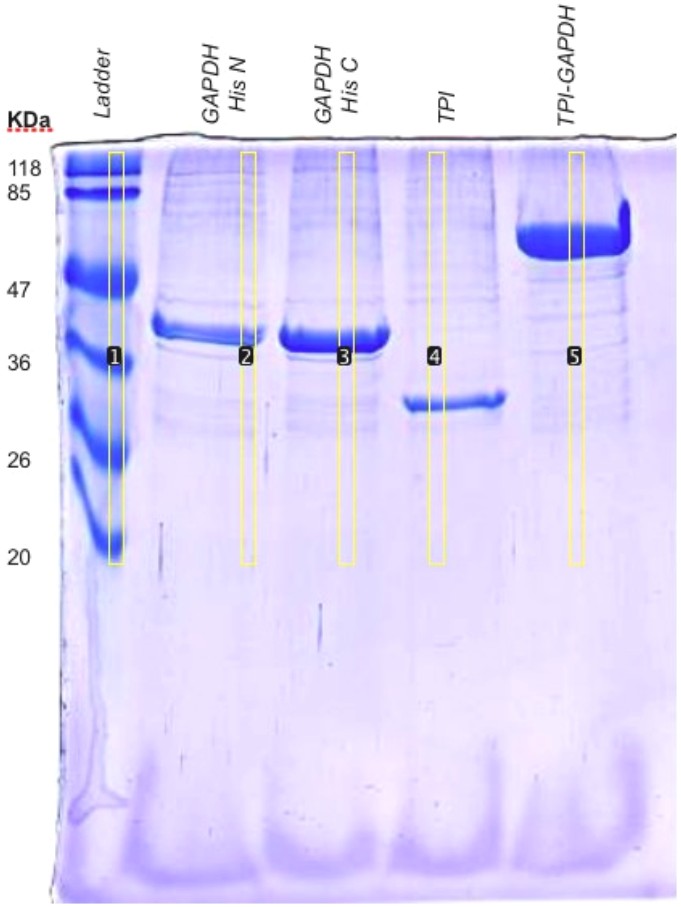


Figure 12: ImageJ Analysis of a 13.5% SDS PAGE gel showing re-dissolved ammonium sulphate preparations.

Log (MW) ladder values were plotted against the distance travelled. An equation line was generated and the distances travelled by each sample were calculated in order to generate a Log (MW) value for proteins of interest. This number was multiplied by $10^{Log(distance)}$ to determine the approximate molecular weight of samples.

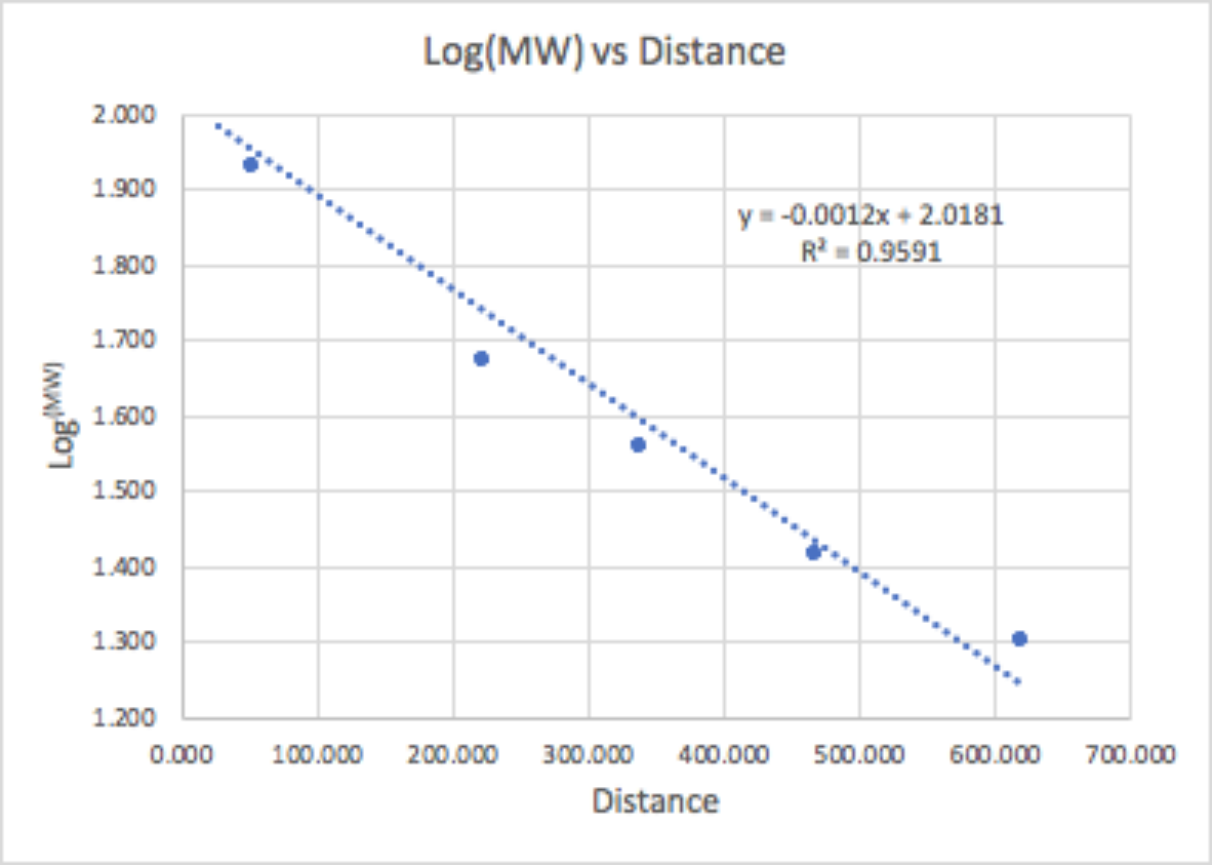


Figure 13: Log (MW) values of known protein ladder sizes plotted against distance travelled. The equation for the line of best fit was used to determine the approximate sizes of the protein samples produced.

Table 5: ImageJ data to determine MW as a function of distance travelled

Protein Elution Gel	Known MW (kDa)	Measured MW (kDa)	Distance Travelled	Log (MW)
Ladder	118.000	x	28.000	2.072
	85.000	x	54.000	1.929
	47.000	x	224.000	1.672
	36.000	x	338.000	1.556
	26.000	x	468.000	1.415
	20.000	x	621.000	1.301
GAPDH HisC	38.7 kDa	40.973	338.000	1.613
GAPDH His N	38.7 kDa	41.314	335.000	1.616
TPI	28.6 kDa	31.427	434.000	1.497
Fusion	65.0 kDa	69.263	148.000	1.841

Table 5: ImageJ data for producing molecular weight values of protein samples. Log (MW) ladder values were plotted against the distance travelled. An equation line was generated, and the distances travelled by the samples were plugged in to generate a Log (MW) value for proteins of interest. This number was multiplied by $10^{Log(distance)}$ to determine approximate molecular weight of samples.

3.2 Enzyme Assays

3.2.1 Calculating Enzyme Concentration

Figure 14 shows the absorbance spectrum of the enzymes measured between 220nm and 350nm using a nanodrop spectrophotometer. Amino acids with aromatic rings in their side chain absorb light at 280nm.

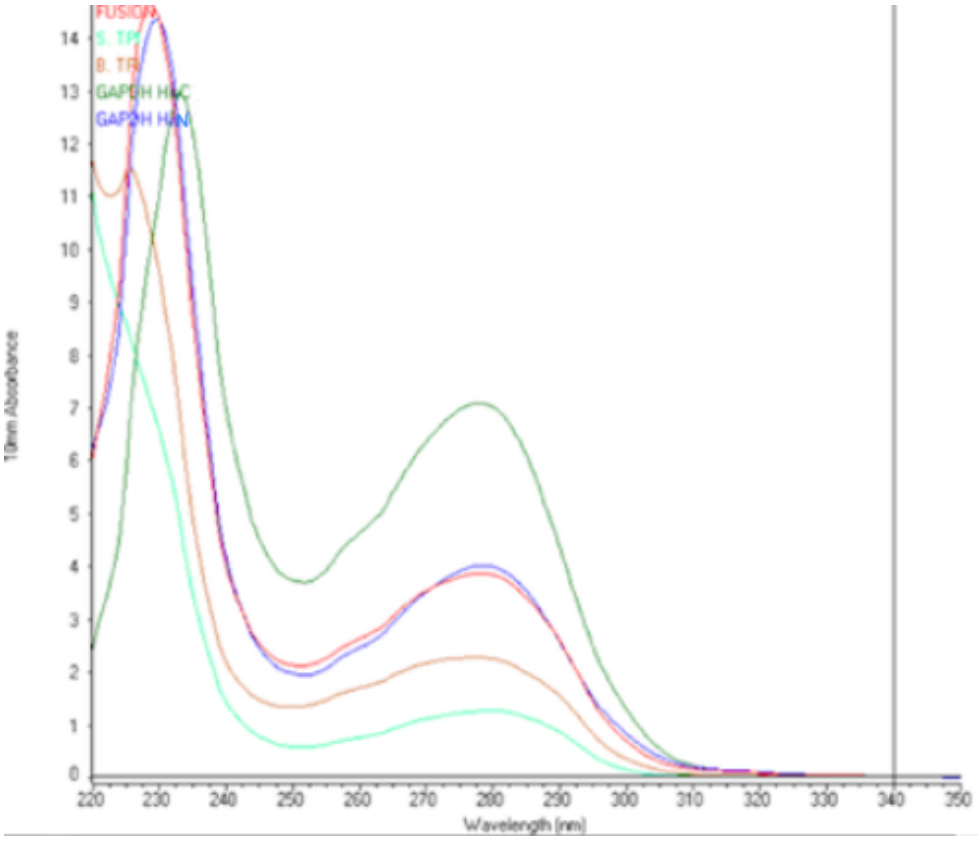


Figure 14: Absorbance spectrum of the TPI-GAPDH fusion enzyme, *S. cerevisiae* TPI, *B. hominis* TPI, GAPDH His C, GAPDH His N from 220nm to 350nm, showing large peaks at 280nm.

The absorbance values determined at 280nm could then be used to determine enzyme concentration using the following equation:

$$\text{Concentration}(\mu M) = (A \div E \div MW) \times 10^6$$

Where A = absorbance at 280nm, E = extinction coefficient and MW = molecular weight (as predicted by ProtParam). All enzyme concentrations were adjusted to 15μM for use in assays.

3.2.2 Establishing Enzyme Rates

Reaction rate is determined by a range of factors including enzyme concentration [E], substrate concentration [S], temperature, the maximum catalytic rate of the enzyme (V_{max}), and the [S] required to reach half V_{max} known as the Michaelis Constant (K_m). Slight variation in [E], [S], or temperature can have significant impacts on rate, which would ultimately result in flawed analysis of enzyme activity. Assays were performed in pseudo first order conditions at 20°C in order to produce comparable Michaelis- Menten plots.

Depletion of available NAD^+ would result in a decreased enzyme activity. Therefore, an excess of NAD is critical to the precise determination of enzymatic rates. The conversion of NAD^+ (260nm) to NADH (340nm) is also the detectable reaction on which all assays performed in this research rely on. GAP is oxidised by the coenzyme NAD^+, and then phosphorylated by the addition of a free phosphate group. GAPDH catalyses the removal of a hydrogen atom from the GAP by NAD^+, resulting the conversion of NAD^+ to NADH and GAP to 1,3 BPG. Figure 15

below validates the quantification of NADH as a function of enzyme activity and also evidences that the concentration of NAD^+ remains effectively constant through the reaction.

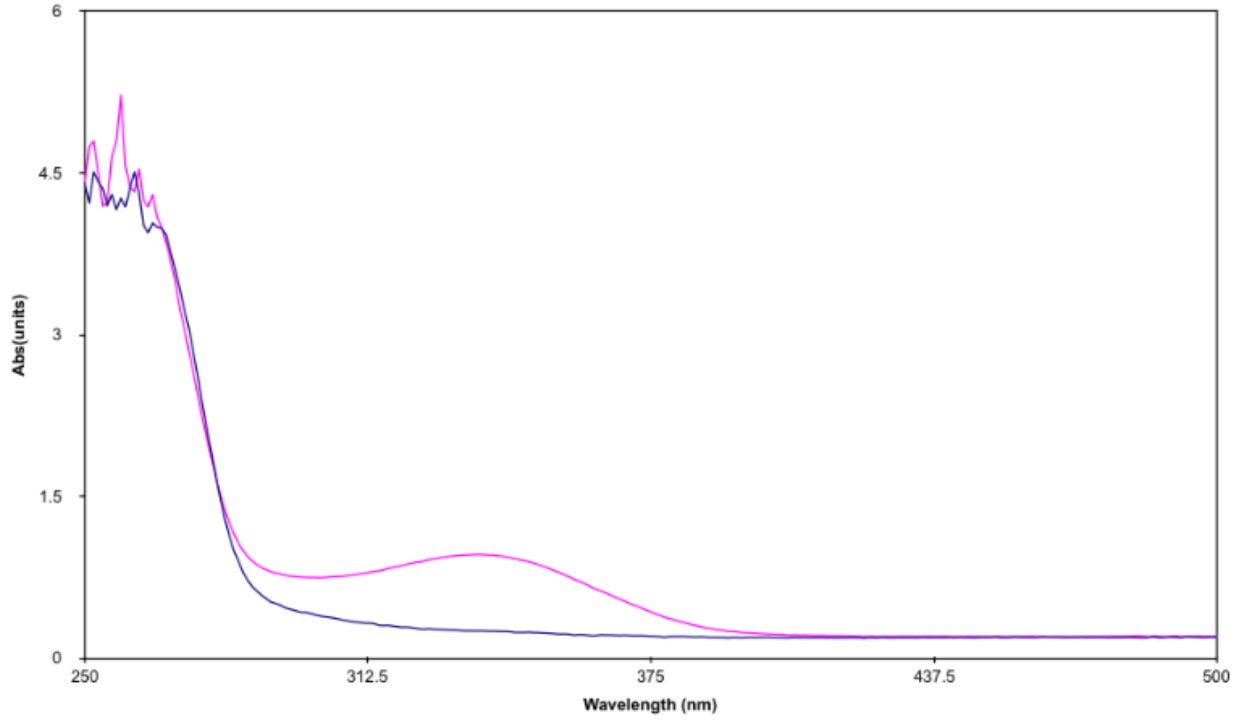


Figure 15: Absorbance of an assay solution before (blue) and after (purple) GAPDH catalysis of GAP. The high NAD^+ concentration (260nm) remains equal in both samples while a significant quantity of NADH (340nm) can be visualised after 30 minutes of activity.

Using a constant enzyme concentration [E] while manipulating the concentration of a substrate [S] in a reaction allows us to produce variations in initial rate (V_0) as a function of [S]. These values will be used produce a curve in order to determine the V_{max} and K_m of the enzymes according to the first order Michaelis-Menten equation:

$$V_0 = V_{max}\,[S] \div K_m + [S]$$

Figure 16 below represents a plot of rate against [S] depicting a Michaelis-Menten curve. The V_{max} line represents maximal rate of reaction due to substrate saturation of enzyme. The amount of substrate needed to achieve half of the value of V_{max} is labelled K_m. A linear period is highlighted in red where rate is approximately proportional to [S]. In this region increasing the [S] by 1 results in a proportional increase of rate by 1. As the rate approaches $V_{max,}$ a hyperbolic curve is produced representing an increasingly diminished rate return on [S] used.

]

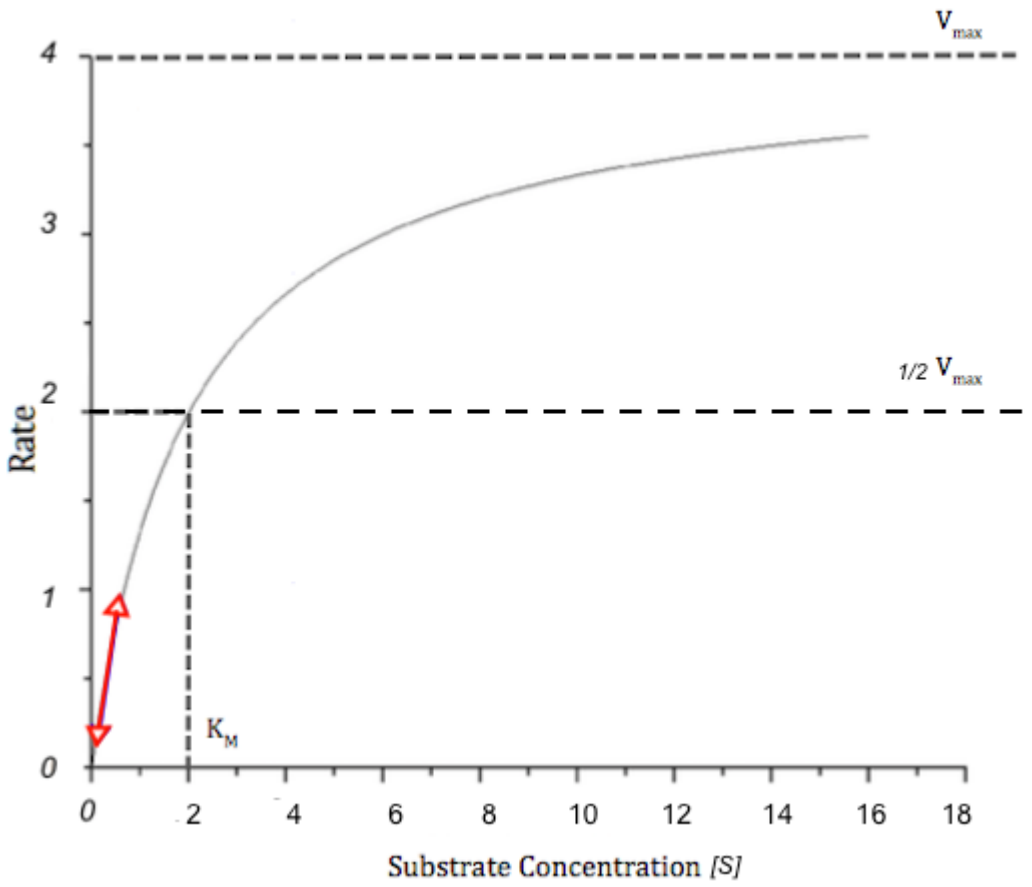


Figure 16: A Michaelis-Menten curve representative of rate against [S]. This curve demonstrates a V_{max} of 4, and a K_m of 2. The linear region in red represents the region where rate is proportional to [S]

Figure 17 represents primary data from enzyme assays plotting the absolute absorbance at 340nm over time. NADH is produced as a result of GAPDH dehydrogenase activity and quantified as a measure of enzyme activity. GAP concentrations plotted are 0.03mM, 0.06mM, 0.13mM, 0.25mM, 0.5mM and 1mM. A significant increase in initial rate can be easily visualized from the readings which suggests that the concentrations of GAP used are in the linear region of Michaelis-Menten curve. This verifies the concentrations of substrate used for the assay.

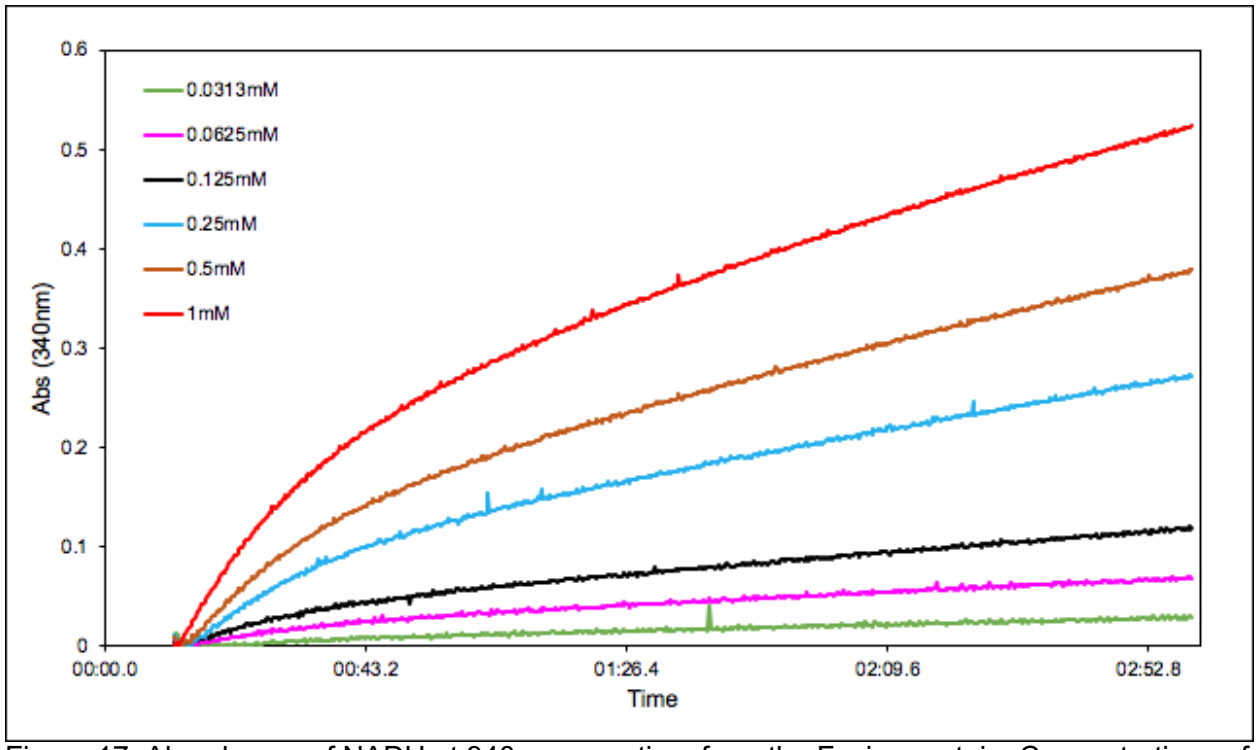


Figure 17: Absorbance of NADH at 340 nm over time from the Fusion protein. Concentrations of 0.0313mM, 0.0625mM, 0.125mM, 0.25mM, 0.5mM and 1mM GAP are shown.

A variety of combinations of enzyme or enzyme combinations were tested with various concentrations of GAP or DHAP as substrate. The rate can be understood by the following equation:

$$R = \Delta Concentration \div \Delta Time$$

The initial rate of each enzymatic reaction performed was calculated using the VisionPro spectrophotometer software. All initial rate data can be found in appendix 7.

3.2.3 Calculating V_{max} and K_m

The average rate for each set of assays was calculated using Excel and then plotted into IC50.tk. Using IC50.tk a curve was fitted to the data points according to the Michaelis-Menten equation in order to estimate V_{max} and K_m values. This simulated data allowed the curve to be projected over a far greater range of [S] than could be tested. Figure 18 illustrates the IC50.tk-generated curve for GAPDH His C, GAPDH His N, TPI-GAPDH Fusion Protein, and GAPDH His C + B.TPI with GAP as the substrate. The data points show the actual data gathered.

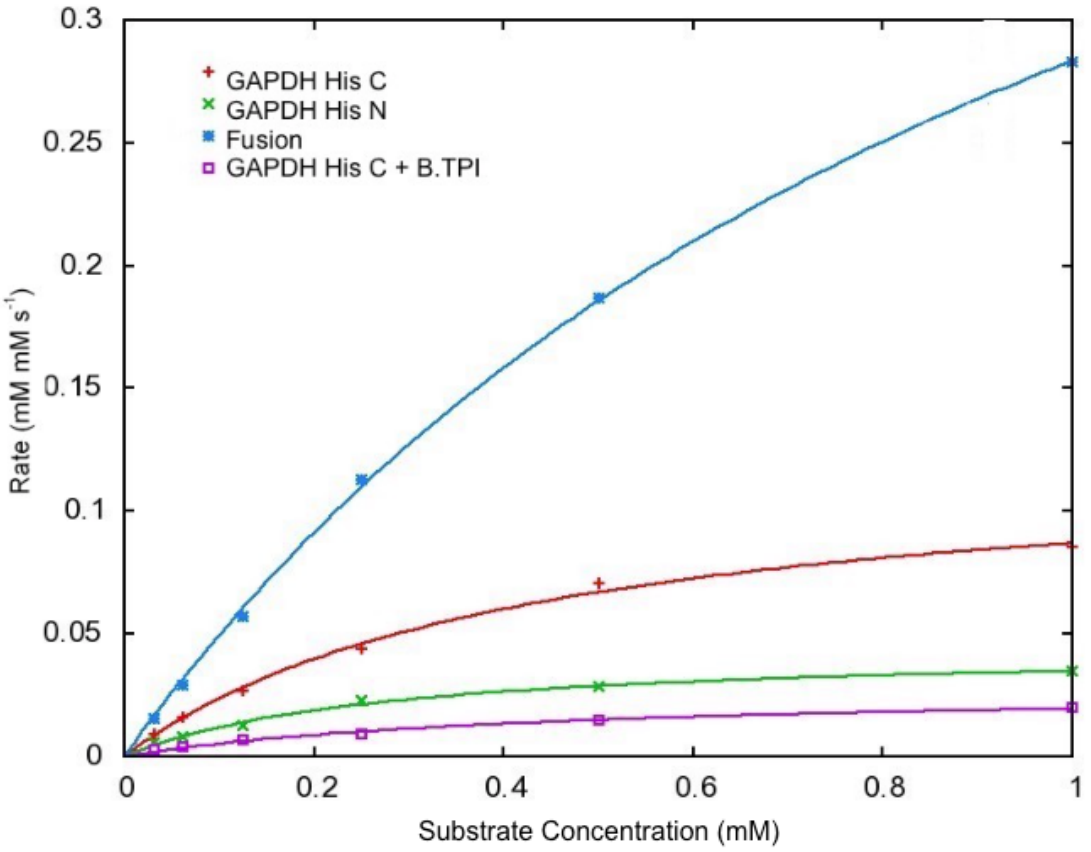


Figure 18: Fitted curves of GAPDH His C (red), GAPDH His N (green), Fusion Protein (blue) and GAPDH His C + B.TPI (purple) with GAP as the substrate. The coloured data points represent the average rates over a range of substrate concentrations. The V_{max} and Km values were simulated with IC50.tk using the values collected.

The simulated V_{max} and K_m values are shown in table 6 below. A higher Km means that a higher concentration of substrate in needed to reach half of V_{max}. The V_{max} for the blastocystis TPI/GAPDH fusion enzyme appears to be significantly higher than blastocystis GAPDH on its own. This is surprising as the TPI portion of the fusion protein should be rapidly removing available GAP by inter-converting GAP to DHAP at a 1:22 ratio. A much slower rate therefore is expected for the fusion protein construct compared to GAPDH for equal [S]. The V_{max} of the His C GAPDH variant is significantly higher than the His N, and His C with B.TPI. For GAPDH His C with and without Blastocystis TPI the Vmax should be identical. The variation causes difficulty in determining the actual K_m of the enzyme. The V_{max} and K_m values produced for the GAPDH + B.TPI samples are unrealistically low and therefore unreliable.

Table 6: V_{max} and K_m simulated values for enzymes using GAP as the substrate

Enzyme	V_{max} (mM mM^{-1} s^{-1})	K_m (mM)
Fusion	0.600	1.117
GAPDH His C	0.124	0.426
GAPDH His N	0.045	0.275
GAPDH His C + B.TPI	0.028	0.444

The fusion protein V_{max} and K_m are significantly higher than the GAPDH enzymes. However, discrepancies in GAPDH HisC and GAPDH HisC + B.TPI V_{max} and K_m bring into question the reliability of these results. The fusion enzyme has a V_{max} 4.84 times higher than GAPDH His C when GAP is used as substrate.

Another set of curves was fitted for enzymes tested with DHAP as substrate (figure 19). Average rates of triplicates were calculated using VisionPro spectrophotometer software and a curve was fitted to the data points generated using IC50.tk in order to determine estimated V_{max} and K_m values. The TPI/GAPDH fusion, GAPDH His C + B.TPI, GAPDH His C + S.TPI all demonstrated higher rates of enzyme activity compared to when DL-GAP was used as substrate. This is surprising as the introduction of pure GAP substrate should, based on what is currently understood, provide a rate up to 22 times higher than an equivalent amount of DHAP substrate. This preference of DHAP to DL-GAP suggests either a divergence in the TPI-GAPDH relationship as currently understood or could be potential evidence of L-GAP isomer inhibition to GAPDH.

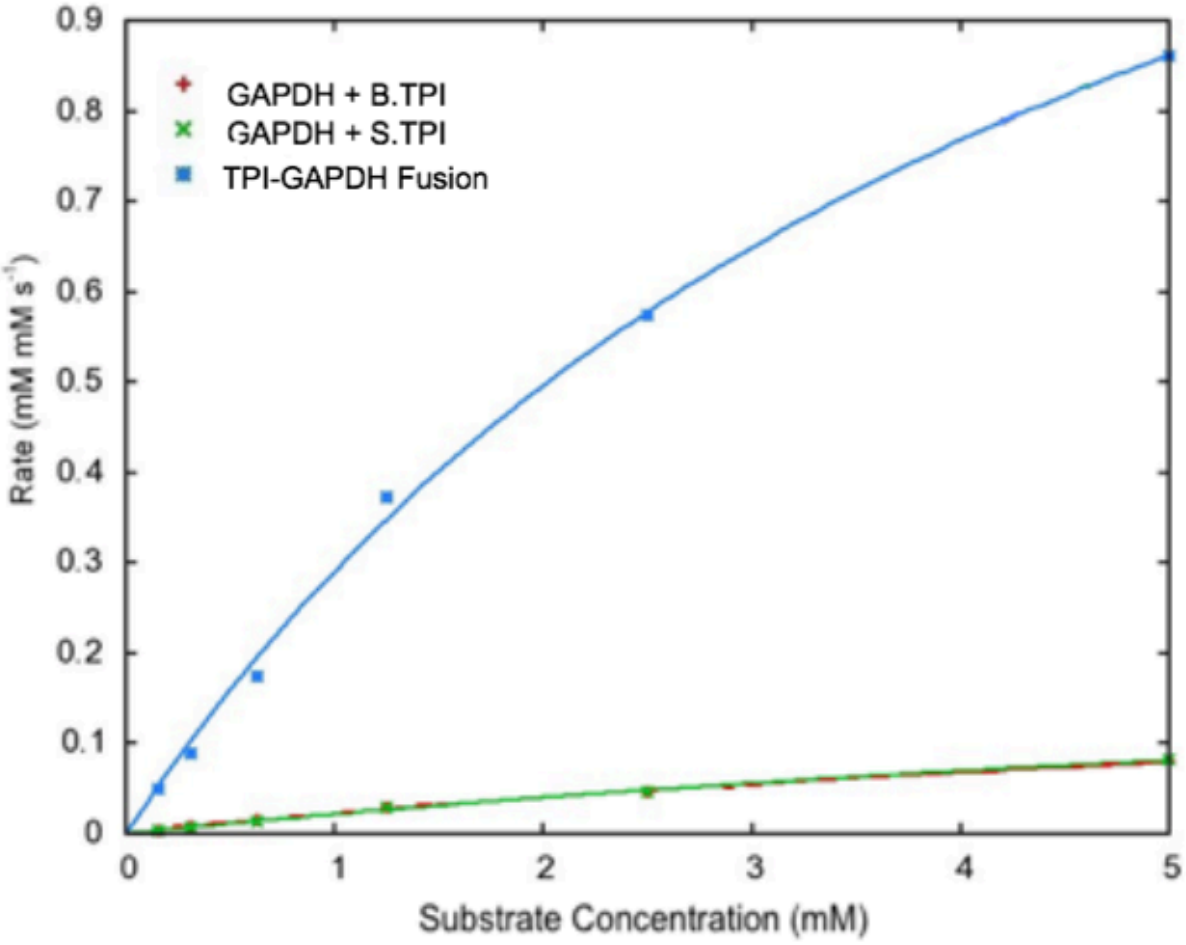


Figure 19: Fitted curves of GAPDH + B. TPI (red), GAPDH + S.TPI (green), Fusion Protein (blue) using DHAP as substrate. The coloured data points represent the average rates over a range of substrate concentrations. The B.TPI curve is semi-hidden behind the S.TPI curve as the rates are nearly identical. The V_{max} and K_m values were simulated with IC50.tk using the values collected.

Simulated V_{max} and K_m values are shown in table 7 below. Note that the V_{max} for the blastocystis fusion enzyme is 2.83 times higher with DHAP as substrate compared to GAP. Additionally, the B.TPI + GAPDH combination produced a simulated V_{max} 7.9 times higher with DHAP as a substrate, although the V_{max} produced with GAP is unreliable. The Km values increased significantly when DHAP was used as substrate. The fusion protein reflects a Km 4.33 times higher, and the B.TPI + GAPDH 20.2 times higher. A higher Km illustrates a lower affinity for DHAP as substrate, requiring a greater [S] to achieve V_{max}.

Table 7: V_{max} and K_m simulated values for enzymes using DHAP as the substrate

Enzyme	V_{max} (mM mM^{-1} s^{-1})	K_m (mM)
Fusion	1.696	4.839
B.TPI + GAPDH	0.221	8.974
S.TPI + GAPDH	0.264	11.177

3.2.4 Point Rate Comparisons

Initial rates were determined in triplicates for each enzyme or enzyme pair using varying concentrations of either GAP or DHAP as substrate. NAD^+ concentration remained constant at 0.25mM of the total reaction volume. The rates produced were then compared across the range of conditions tested. The reactions were begun by the addition of either GAP, DHAP or NAD^+. Initiating a reaction with NAD^+ allowed time for the TPI enzyme to equilibrate the GAP/DHAP concentration prior to the reaction starting. The TPI enzyme should equilibrate the GAP:DHAP concentrations to the canaonoical ratio of 1 GAP: 22 DHAP, resulting in a significantly lower rate. These comparisons were performed in order to better visualise the equilibrating effects of the TPI enzyme.

Figure 20 below represents initial rate comparisons when reactions were begun by the addition of GAP or allowed to equilibrate and begun with NAD^+. The enzyme or enzyme pairs tested were GAPDH His C, GAPDH + Blastocystis TPI, GAPDH + Saccharomyces TPI, and the TPI/GAPDH fusion enzyme. GAPDH initial rates were effectively equal and unaffected by a GAP or NAD start (P=0.733). The rates for both B.TPI + GAPDH and S.TPI + GAPDH were very similar showing a 36% and 33% decrease in enzyme activity when allowed to equilibrate (P=0.021, and P=0.008) The fusion enzyme demonstrated a 35% decrease in initial rate when allowed to equilibrate(P=0.001). A 22-fold reduction in GAP would result in a significantly slower rate. This suggests that the GAP available is not being equilibrated to the expected 1:22 ratio.

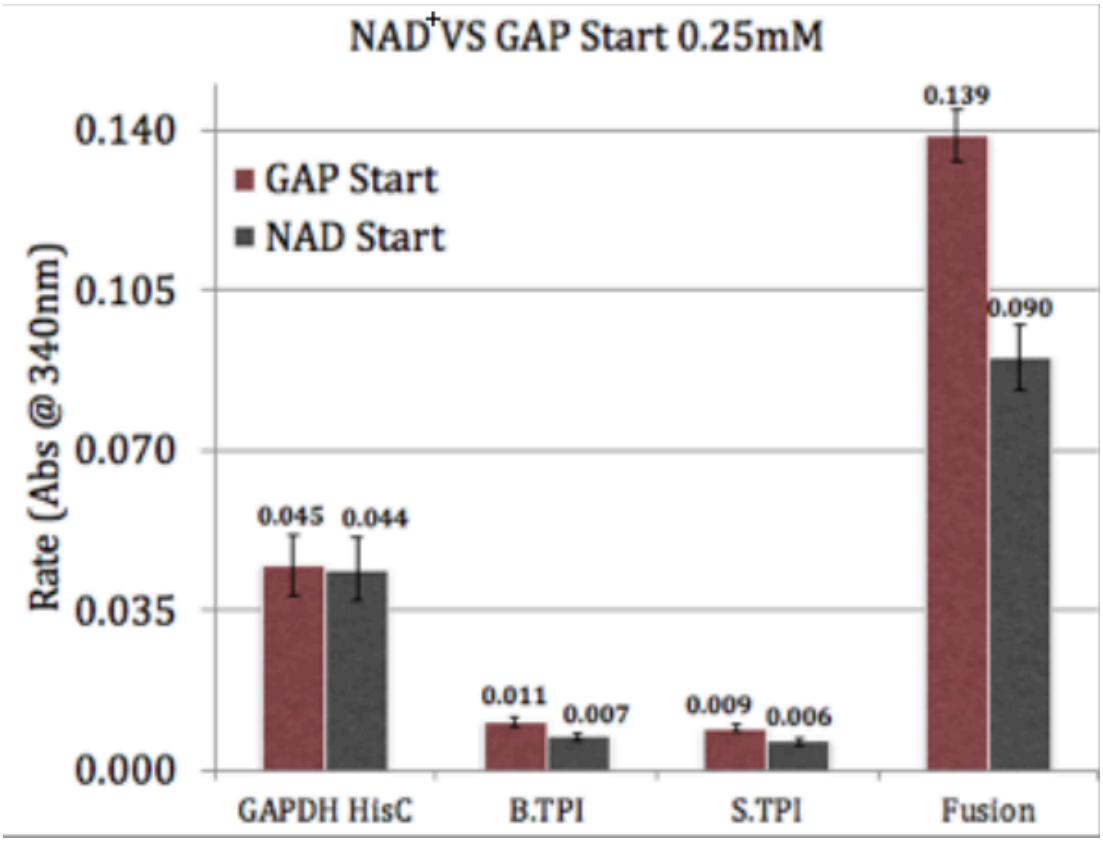


Figure 20: Initial reaction rates of GAPDH His C, GAPDH His C+B.TPI, GAPDH His C+S.TPI, and TPI/GAPDH fusion enzyme for 0.25mM GAP substrate. Reactions were began by the addition of GAP, or allowed to equilibrate and initiated with the addition of NAD^+.

Similar initial rate comparisons were performed using DHAP as substrate. Figure 21 below illustrates the average initial rates for assays either began by the addition of DHAP or allowed to equilibrate and began with NAD^+. Results show an increase in initial rate when DHAP is allowed to equilibrate for all enzymes tested. B.TPI illustrated a 34% increase (P=0.002), S.TPI showed a 21% increase (P=0.015) and the fusion protein showed a 17% increase (P=0.001). These increases in rate are much less than are expected to be seen, and again suggests TPI inefficiencies in equilibrating GAP and DHAP to the canonical ratio of 1:22. Furthermore, the significant increase in initial rate for equilibrated substrate reveals that GAP/DHAP interconversion is not instantaneous.

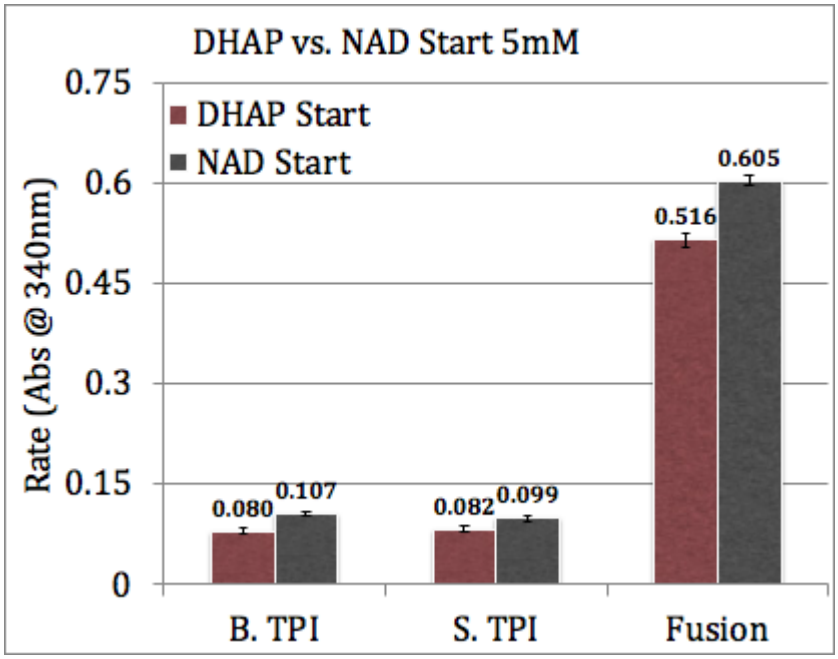


Figure 21: Initial reaction rates of B.TPI + GAPDH, S.TPI + GAPDH, and TPI/GAPDH fusion enzyme for 5mM DHAP substrate. Reactions were begun by the addition of DHAP, or allowed to equilibrate and began with the addition NAD^+. Equilibration of DHAP by TPI proves to increase rate by creating a higher concentration of GAP.

Final equilibration comparisons were performed using equal starting concentrations of GAP and DHAP (figure 22). The enzyme assay mixes were made up with 1mM concentrations of DL-GAP or DHAP and allowed to equilibrate for two minutes prior to the addition of NAD^+. All variants tested showed higher initial rates when DHAP was used as substrate. The data produced by B.TPI (P=0.49) and S.TPI (P=0.19) tests was not statistically relevant and as such was unable to be used for further analysis or conclusion. The fusion enzyme demonstrated a rate 171% faster when DHAP was used instead of GAP (P=0.0002). If the GAP and DHAP availability was truly equilibrated then similar initial rates would be seen. However, an alternative answer may be to call into question the viability of DL-GAP compared to pure D-GAP produced in the conversion of DHAP to GAP by TPI. This information suggests potential GAPDH inhibition by the L-GAP isomer.

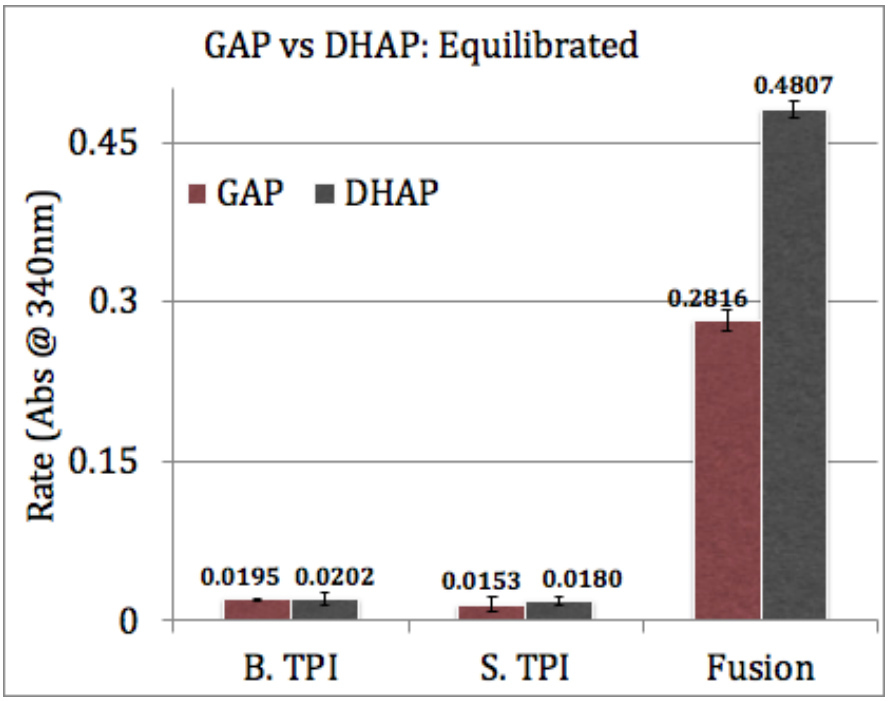


Figure 22: Average initial rates for 1mM GAP or DHAP reactions. By allowing the reaction mix to equilibrate for two minutes with equal concentrations of substrate similar rates are expected. However, The fusion enzyme demonstrated a rate 171% faster when DHAP was used instead of GAP (P=0.0002)

3.2.5 DL-GAP vs D-GAP Point Comparison

In vivo, GAP is produced only in the D isomer form. Current literature cites the D-GAP isomer as a known substrate for GAPDH, and L GAP as a strong competitive inhibitor (Brunner 1998). The results shown in figure 23 below support the findings of Brunner 1998 and confirm that L-GAP does act as an inhibitor for GAPDH. Rate comparisons were performed in triplicates for GAPDH and the GAPDH/TPI fusion protein using DL-GAP or D-GAP as substrate. Higher initial rates were observed for both enzymes at both concentrations when using only D-GAP as a substrate. GAPDH on its own produced an average initial rate 18% faster for both for 2mM D-GAP(P=0.066) and1mM D-GAP(P=0.009). The GAPDH/TPI fusion enzyme produced an average initial rate 21% faster with 2mM D-GAP substrate (P=0.004) and 39% faster with 1mM D-GAP substrate (P=02). Additionally, If the D-GAP and DL-GAP points were along the same curve then the concentration [S] would be off by at least a factor of two as both proteins registered faster rates for 1mM D-GAP than the 2mM DL-GAP rates.

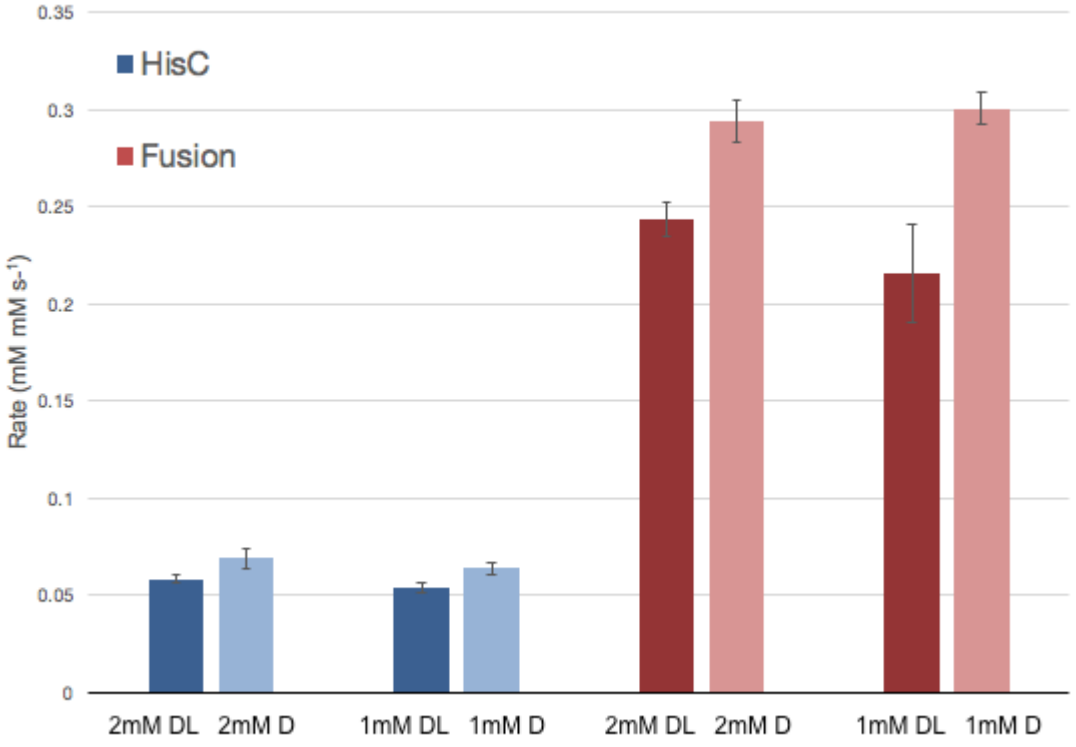


Figure 23: Rate comparisons for GAPDH His C (blues) and TPI/GAPDH fusion (reds) for DL-GAP and D-GAP at 1mM and 2mM concentrations. D-GAP produced higher initial rates across all trials.

Based on the data illustrated in figure 23 above L-GAP is clearly not used as substrate for GAPDH. However, full rate curves would be helpful to better characterise the relationship between L-GAP and the *B. hominis* TPI-GAPDH fusion protein. L-GAP acts as a competitive inhibitor to GAPDH. This does not affect the V_{max} of an the GAPDH enzyme but does dramatically raise the K_m. Testing a range of [S] in order to generate full curves would be needed in order to further explore if this might be the case.

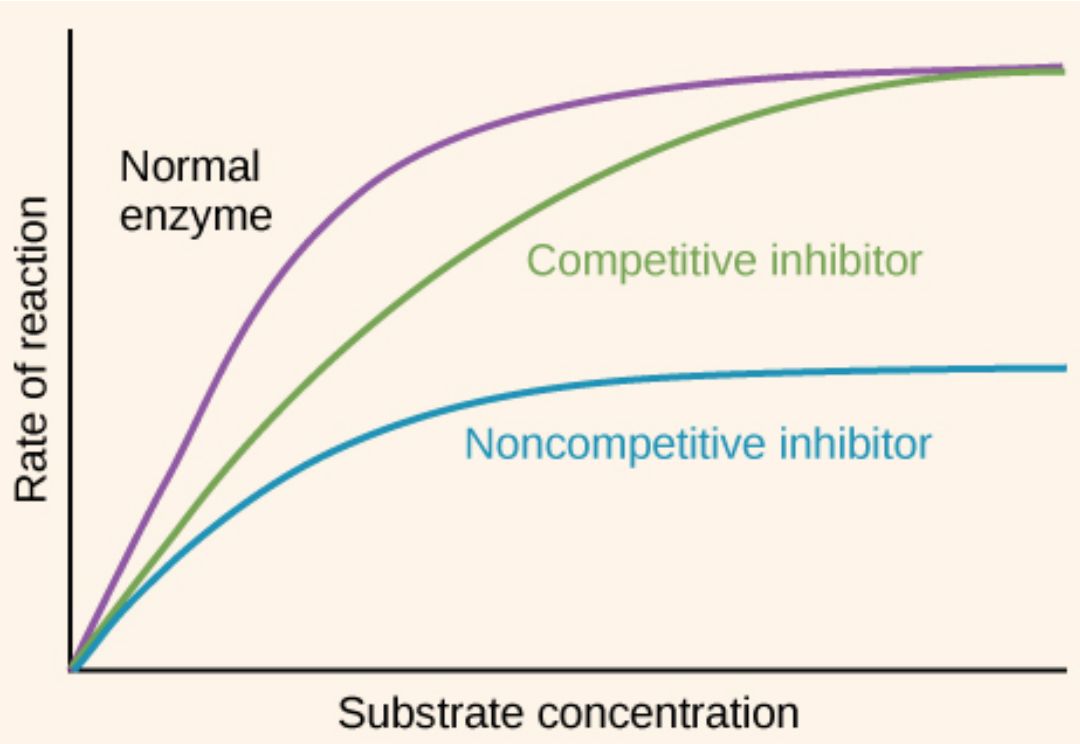


Figure 24: An example of competitive and noncompetitive enzyme inhibition scenarios compared with a normal rate vs [S] relationship (khanacademy). When an enzyme is competitively inhibited V_{max} in unaffected but K_m is increased. Noncompetitive enzyme inhibition lowers V_{max}.

3.3 Thermal Shift Assays

3.3.1 TPI and GAPDH

Thermal shift assays were performed using fluorescence spectroscopy in order to analyse fusion protein stability and to view any potential binding interaction between the B.TPI and GAPDH enzymes. Dissociation curves were produced by denaturing proteins with a temperature ramp from 30°C to 90°C over 2 hours. Fluorescence readings were taken every minute using the FCO3 Quansoft filter cartridge. As temperature increases the proteins denature and hydrophobic residues in the core are exposed. The fluorescent dye binds to the hydrophobic regions and emits a fluorescent beam when charged with light. SYPRO Orange dye was excited by laser at 520nm and emissions measured at 560nm. SYPRO Orange registers an increase in quantum yield when bound to hydrophobic regions in denatured proteins. Therefore, fluorescence can be measured as a reliable indicator for protein dissociation to determine stability.

22µl reactions containing 19µl 15mM protein and 1µl fluorescent dye were setup in PCR tube strips and run in a PrimeQ Real-time Nucleic Acid Detection System as outlined in table 8 below.

Table 8: Thermal shift assay parameters

Step #	Temp°C	Cycles	Hold time	Read Time	Increments
1	20-90	141	30sec	30sec	0.5°C

Table 8: Program parameters for protein stability ramp. Samples were excited at 520nm and emissions measured at 560nm.

Dissociation curves for Blastocystis TPI (red) and Saccharomyces TPI (green) are seen below in figure 25 A and B. Dissociation curves for GAPDH His C (cyan) and GAPDH His N (magenta) are shown in figure 25 C and D. Figure 25 A and C represent fluorescence with respect to temperature, and figure 25 B and D are their first derivative plots showing Δfluorescence / Δtime with respect to temperature. A dissociation difference of about 10°C can be distinguished between the two TPI samples. B.TPI begins to dissociate at a temperature about 10°C lower than S.TPI, and peaks about 10°C lower as well. This shows that the B.TPI is less stable than the S.TPI used in this study. GAPDH His C is much more stable than the His N-tagged protein. A melt peak at approximately 45°C can be seen for His N-tagged GAPDH while a peak at about 77°C can be seen for the His C-tagged variant.

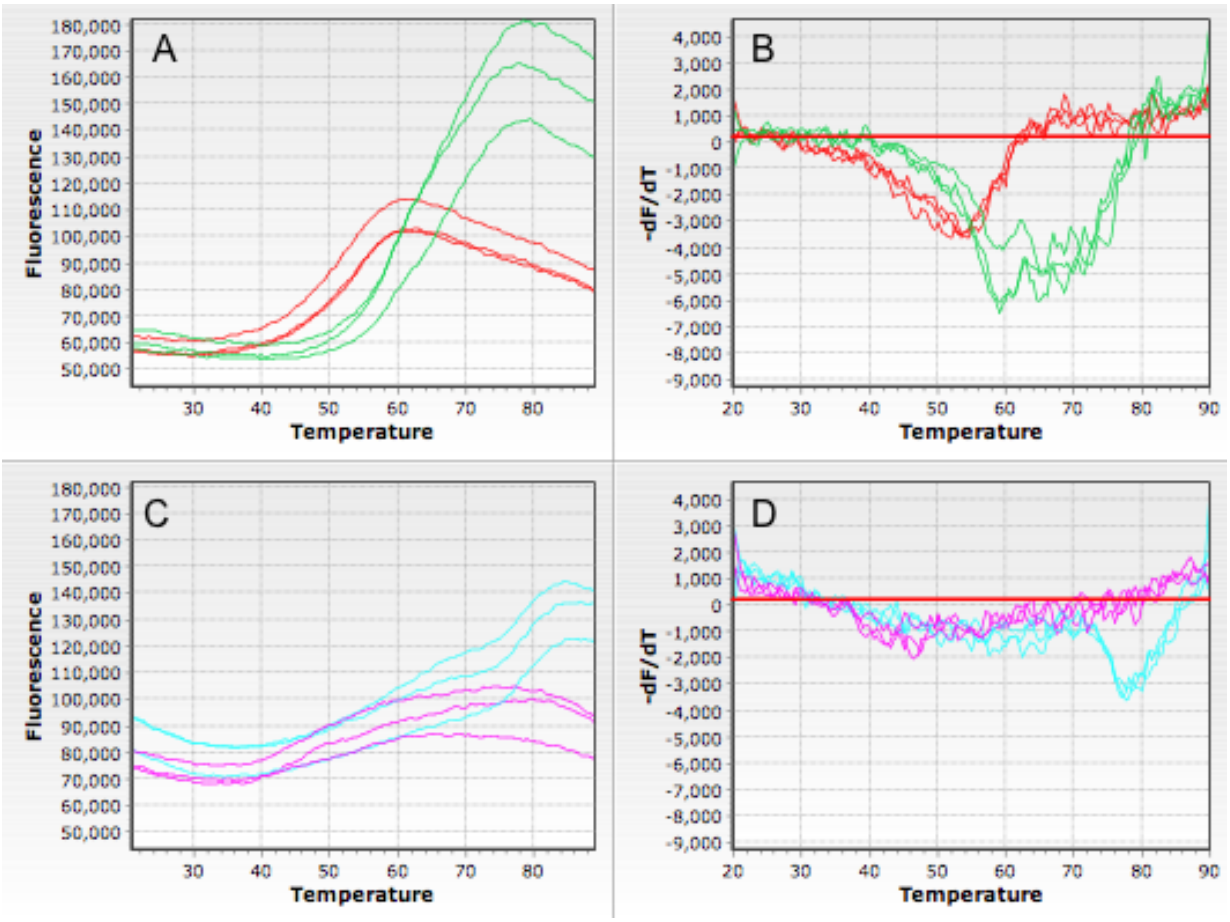


Figure 25: Thermal shift data illustrating dissociation over a range of temperatures. A) B.TPI (red) and S.TPI (green) dissociation in respect to temperature. B) First derivative plot showing Δfluorescence / Δtime. B.TPI begins to denature around 30°C and reaches a peak rate of denaturing around 53C, about 10°C less than the S.TPI. C) His C (cyan) is much more stable than His N (magenta). D) There seems to be an early peak around 45°C for His N, and 77°C for His C.

3.3.2 Fusion Protein and Subunit Combination

The fusion protein (blue) appears much less stable than B.TPI (red) or GAPDH His C (cyan) on their own. The fusion protein is shown here to dissociate in two distinct stages. Transition peaks are seen in the melt data at approximately 43°C and 52°C. The first peak seems to suggest a destabilized TPI moiety, which melts about 10°C lower. The second peak is much sharper and appears to correspond with GAPDH but dissociates approximately 25°C lower than GAPDH His C by itself. Remarkably, TPI appears to significantly destabilise the fusion protein. B.TPI appears to be stabilised by GAPDH (orange). The GAPDH His C + B.TPI combination produces a slow dissociation from 35°C until about 61°C. Transition stages at approximately 69°C and 76°C may be indicative of potential protein-protein binding. If so, this is significant as no known current research has evidenced binding of TPI and GAPDH molecules. More research is needed to determine if this is the case or not.

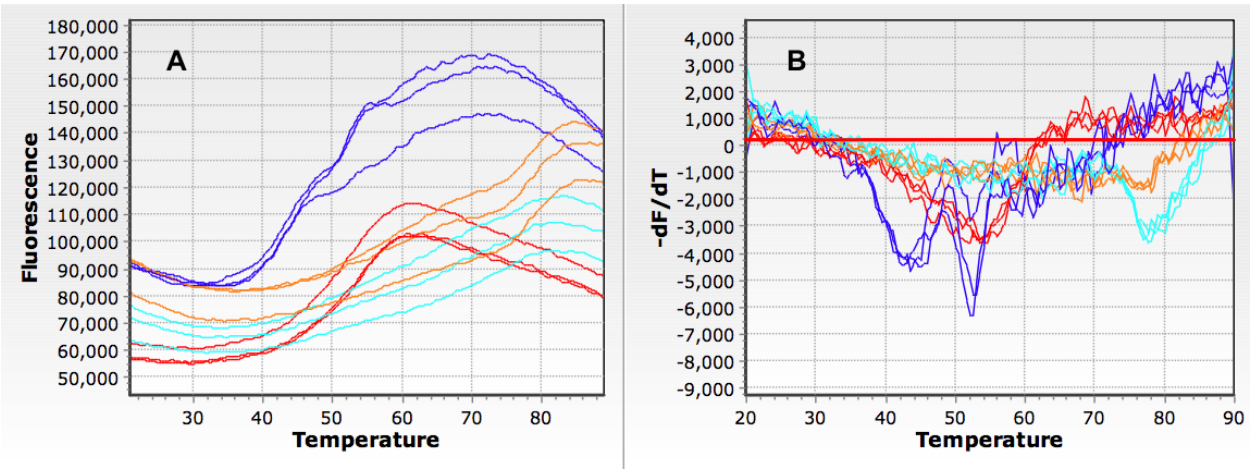


Figure 26: Thermal shift data illustrating dissociation curves for the TPI-GAPDH fusion protein (blue), Blastocystis TPI (red), GAPDH His C (cyan), and a GAPDH His C + B.TPI combination (orange). A) The GAPDH His C + B.TPI combination appears to produce a unique dissociation curve with two transition stages. B) The first derivative plot of -Δfluorescence / Δtime confirms the GAPDH His C + B.TPI double transition peak and suggests B.TPI stabilisation by GAPDH, suggesting that protein-protein binding is occurring.

3.3.3 NAD^+ Binding Stabilises the Fusion Protein

Thermal shift data demonstrates significant NAD^+ stabilisation of the fusion protein, but interestingly has minimal to no effect on the stability of GAPDH His C used in this study. Figure 27 below represents the averages of triplicate melts performed on GAPDH His C and the fusion protein, both with and without the presence of NAD^+.

Samples were prepared as previously stated but with the addition of 2µl of 7.5mM NAD^+ for a final concentration 0.68mM NAD^+. The addition of NAD^+ to the fusion enzyme causes substantial increase in both presumed TPI (43°C) and GAPDH (52°C) stability, along with the formation of a new melt peak at 64°C. The Fusion + NAD^+ second peak occurs at about 64°C,

while the fusion sample without NAD^+ peaks at about 52°C. Interestingly however, the GAPDH His C protein was almost entirely unaffected by the presence of NAD^+, and appears to be stabilised by about 2°C. This data appears anomalous as NAD^+ is a known substrate of GAPDH and is sure to have been bound.

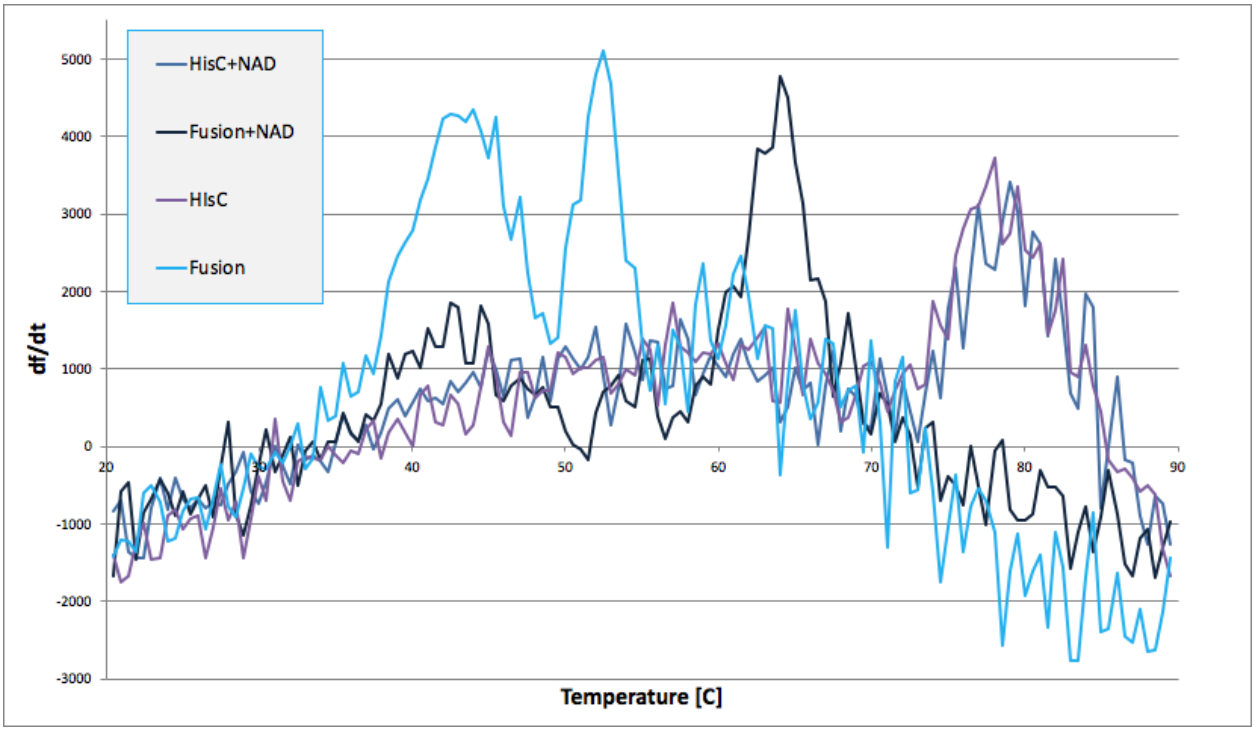


Figure 27: Change in sample fluorescence over change in time. The Fusion + NAD^+ second peaks occurs at about 64°C, while the fusion sample without NAD^+ peaks at about 52°C. The GAPDH His C + NAD^+ appears to have been minimally stabilized by about 2°C.

3.4 Active Site Cysteine Inactivation

3.4.1 MMTS Inactivation Assay

A variety of known parasites contain a TPI cysteine residue at the protein dimer interface. Derivatization with thiol-specific reagents have been shown to induce enzyme inactivation and aggregation, while TPI varieties which are lacking this residue are mostly unaffected (Garza-Ramos et al., 1996). A multiple sequence alignment was produced for protein sequences of B.TPI, S.TPI, and GAPDH/TPI using Unipro UGENE v1.29.0 (figure 28). The results confirm the presence of the Cysteine residue in the Blastocystis TPI and fusion protein, as well as the absence of the residue in the Saccharomyces TPI.

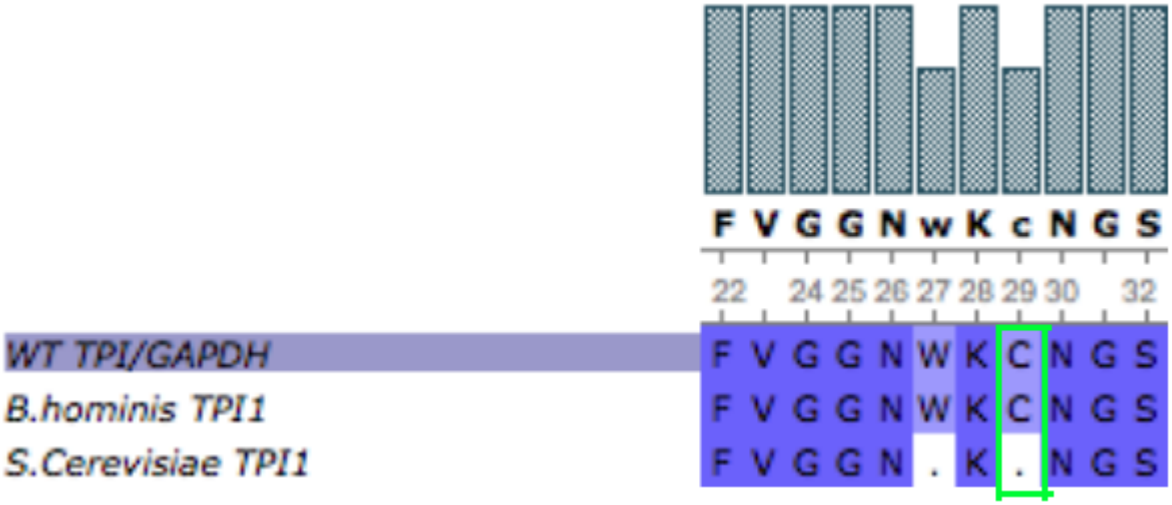


Figure 28: Multiple sequence alignment for B.TPI, S.TPI, and GAPDH/TPI protein sequences evidencing the highly homologous enzyme structure. The wtTPI/GAPDH fusion protein and the B.TPI both contain the target cysteine residue located at the TPI dimer interface (Highlighted in green), but the S.TPI does not.

Methyl Methanethiosulfonate (MMTS) has been previously reported to inactivate TPI proteins by acting on this cysteine residue. Enzyme activity was tested before and after incubating potentially vulnerable B.TPI, S.TPI and fusion proteins in 1mM MMTS. The results can be seen in figure 29 below. Assays were started by the addition of 1mM DHAP and the absorbance was measured at 340nm for each enzyme in triplicates at T=0(blue). 10µl of 15µM GAPDH His C was added to B.TPI and S.TPI samples when preparing each assay mix. Enzymes were then incubated at 4°C for 120 minutes with or without MMTS. Enzymes which had not been exposed to MMTS were used as a control in order to account for natural enzyme degradation (orange). After 120 minutes B.TPI, S.TPI and TPI-GAPDH fusion proteins demonstrated 84%, 89%, and 83% activity respectively. Samples which were incubated in 1mM MMTS after 120 minutes (grey) showed no meaningful decrease in activity compared to their respective control groups (B.PTI P=0.649, S.TPI P=1.000, Fusion P=0.831). This result appears contrary to previously reported findings on the ability of MMTS to cause enzyme inactivation and aggregation (Garza-Ramos et al., 1996).

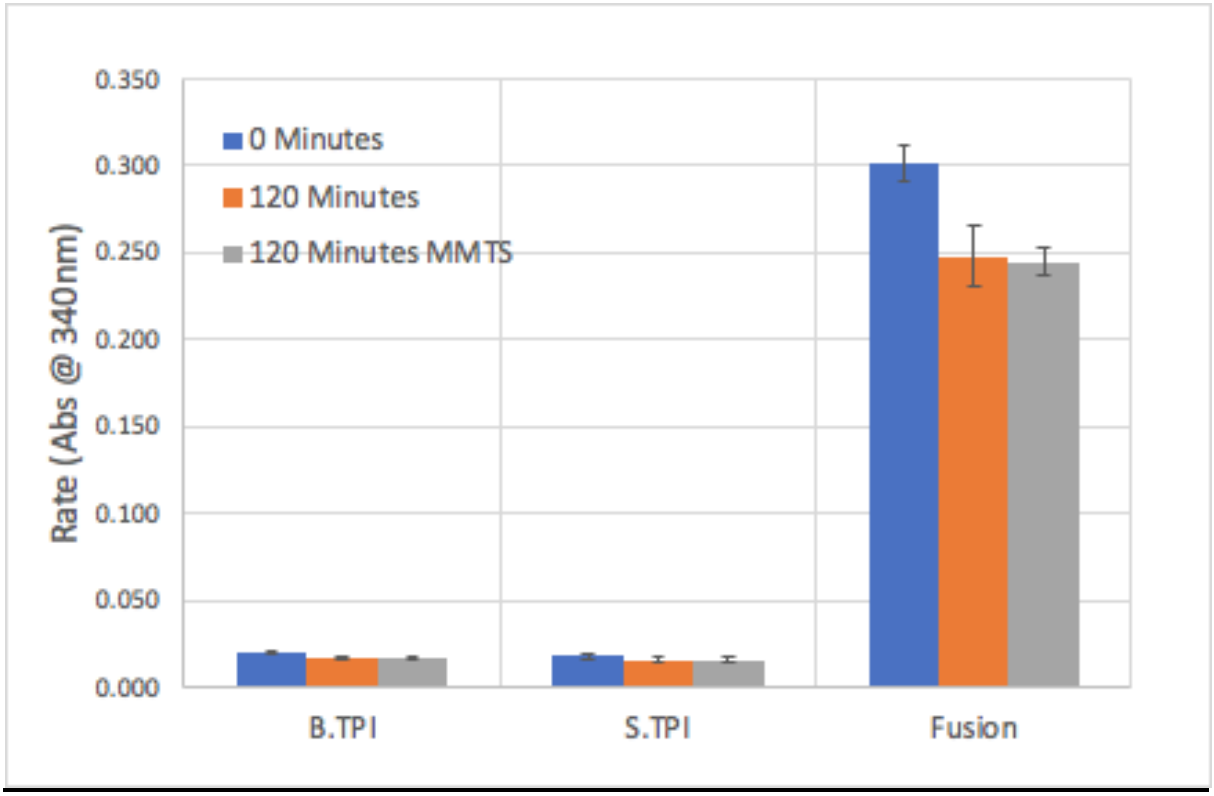


Figure 29: Initial rates for enzymes before and after 120-minute incubation with and without 1mM MMTS. Enzyme activity decreased by 16% (B.TPI), 11% (S.TPI), and 17% (Fusion) over 120 minutes without MMTS (orange). Contrary to expectation no significant difference in enzyme activity was observed between samples incubated in 1mM MMTS (grey).

3.4.2 MMTS Thermal Shift Assays

Dissociation curves of B.TPI and S.TPI were produced with and without MMTS using fluorescence spectroscopy. The samples were incubated at 20°C for 15 minutes, 30°C for 90 minutes and then ramping the temperature from 30°C to 90°C over 2 hours. B.TPI contains the cysteine 29 active site residue and was significantly less stable at temperatures over 20°C when incubated with 1mM MMTS (figure 30 A and B). The first derivative graph evidences dissociation at about 10°C lower (figure 30 B). The S.TPI lacks the cysteine 29 active site residue and was mostly unaffected by MMTS incubation (figure 30 C and D). The S.TPI transition peak suggests marginal change in dissociation temperature of approximately 1-2°C (figure 30 D).

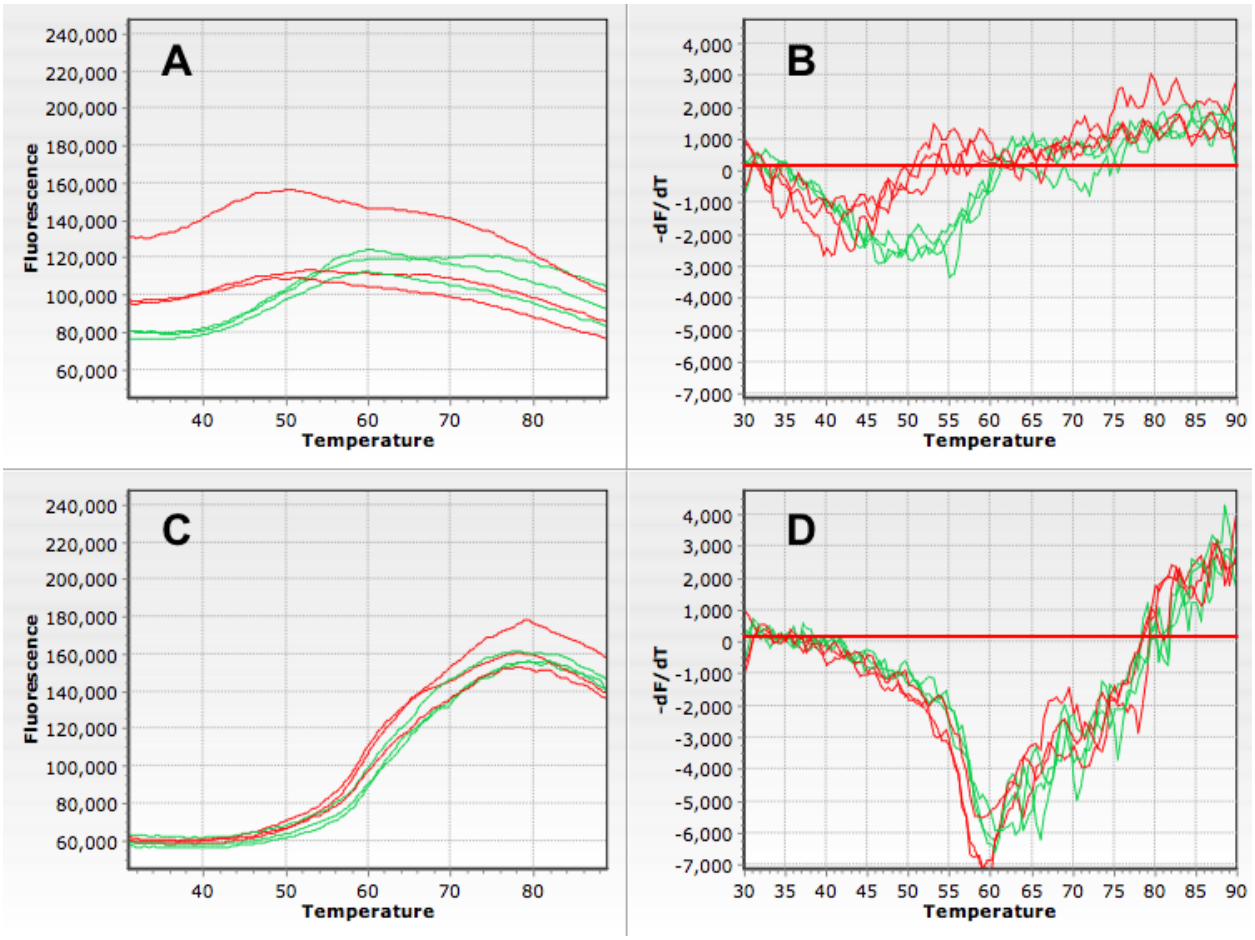


Figure 30: Dissociation curves of B.TPI and S.TPI with MMTS (red) and without MMTS (green). Samples were incubated on a ramp from 30°C-90°C over a 2-hour period. A) Dissociation curve of B.TPI with MMTS shows some prior dissociation at 30°C (data not shown). B) B.TPI with MMTS dissociated about 10°C lower than without. A small second peak is visible at approximately 65°C. C) S.TPI stability was mostly unaffected by MMTS. D) A possible destabilisation of approximately 2°C is illustrated.

Table 9: PrimeQ PCR Fluorescence Spectroscopy MMTS Assay Parameters

Step #	Temp (°C)	Cycles	Hold time	Read Time	Increments
1	20	10	60sec	30sec	N/A
2	30	60	60sec	30sec	N/A
3	30-90	121	30sec	30sec	0.5°C

Table 9: Program parameters for MMTS protein melt. Hold time + read time = total cycle length.

Figure 31 illustrates the fusion enzyme dissociation curves with and without 1mM MMTS during a 20°C (A) and 30°C (B) incubation. The MMTS fusion sample began to denature even prior to the initial incubation. Previous research suggests that this is most likely the result of a cysteine residue in the GAPDH moiety of the fusion enzyme, however replicating this experiment with GAPDH would either confirm or disprove this. The TPI moiety of the fusion protein begins to rapidly dissociate when the incubation temperature was increased to 30°C (figure 31 B).

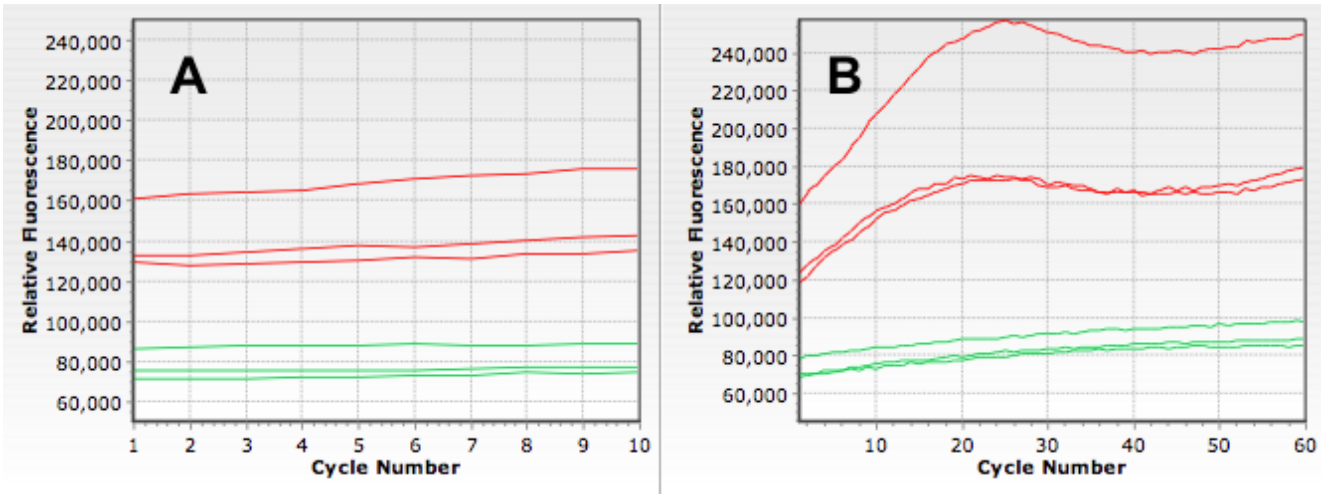


Figure 31: Thermal stability charts representing the TPI/GAPDH with (red) and without (green) incubation in 1mM MMTS. A) MMTS likely facilitated partial dissociation of the GAPDH portion of the enzyme prior to the beginning of this incubation at 20°C. B) The rate of the initial dissociation rapidly increases when incubation temperature increases to 30°C. A second transition period is evidenced during the incubation at 30°C beginning at cycle number 45, or roughly 67 minutes at 30°C.

Chapter 4: Discussion

4.1 The B. hominis TPI/GAPDH fusion enzyme catalyses the same glycolytic function as more common enzyme subunits.

As previously described by current literature the B. hominis TPI-GAPDH fusion enzyme catalyses the same glycolytic reactions performed by the more common TPI and GAPDH protein subunits (Liaud et al., 2000). TPI conversion of DHAP to GAP is an essential step in glycolysis. TPI is understood to maintain a thermal equilibrium of 22:1 DHAP:GAP, and only GAP is used in the next stage of glycolysis by GAPDH. This same function was evidenced by the TPI-GAPDH fusion protein during assays where DHAP was used as substrate. This activity of the fusion enzyme performed in vitro mirrors the essential in vivo TPI glycolytic inter-conversion of the DHAP to usable GAP (figures 19,21,22, table 7). Similarly, the GAPDH portion of the fusion protein moiety performs the same glycolytic functions as it would in cells but is potentially more active. NAD^+ is reduced to form NADH as a result of fusion protein GAPDH catalytic action (figure 18, table 6). The GAPDH portion of the fusion enzyme demonstrated an initial rate 3.08 times that of the separated GAPDH His C when 0.25mM GAP was used as substrate (figure 20). This is surprising as the TPI portion of the enzyme is shown to be rapidly removing available GAP and converting it to DHAP (figure 20) The GAPDH His C enzyme was estimated to have a V_{max} of 0.124 mM mMs^{-1} while the fusion V_{max} was estimated at 0.600 mM mMs^{-1}. This may be an erroneous measurement as the true V_{max} of an enzyme does not change, and the GAPDH portion of the fusion enzyme is structurally identical to the separated GAPDH moiety. However, the fusion enzyme as a whole is structurally different, and the higher V_{max} could be an insight into a far greater enzymatic potential of the fusion enzyme potentially as a biophysical result of being attached the TPI protein in a unique orientation.

4.2 Dividing *B. hominis* GAPDH and TPI moieties results in decreased catalytic activity

Separated GAPDH His C and His N enzyme are less active than the *B. hominis* fusion enzyme as a whole. The protein subunits have significantly decreased catalytic activity across all conditions tested compared to the activity of the intact fusion enzyme. Appendix 7 contains a chart comparing enzyme activity between these three enzymes when GAP is used as substrate. With 0.25mM GAP GAPDH His C in combination with B.TPI demonstrated the lowest activity (0.009mM mMs^{-1}), then GAPDH His N (0.023mM mMs^{-1}), GAPDH His C (0.044mM mMs^{-1}), and the fusion enzyme was the fastest demonstrating a rate of (0.113mM mMs^{-1}). The rapid conversion of GAP to DHAP by TPI justifies the slow rate of GAPDH His C + B.TPI. Removing the TPI enzyme should theoretically cause a potential 22-fold increase in GAPDH activity, but interestingly the GAPDH His C only registered a rate 4.9x faster when the B.TPI was removed. The TPI-GAPDH fusion enzyme demonstrated a rate 12.6x faster than the GAPDH His C + B.TPI combination. This evidences far greater catalytic activity and greater efficiency of the fusion protein compared to the detached TPI and GAPDH subunits.

Interestingly the fusion protein has higher initial rates and a higher estimated V_{max} for both GAP and DHAP as a substrate. The Fusion protein registered a higher Km when GAP was used, and a lower Km when DHAP was used (figure 18,19 Tables 6,7). When using 5mM DHAP as substrate the intact TPI-GAPDH fusion protein demonstrated a rate 5.65x faster than GAPDH His C + B.TPI combination. A possible explanation for this could be an overall increased efficiency of the protein as a result substrate channelling. This would support the hypothesis that these protists have kept the TPI-GAPDH bound together for metabolic efficiency. This suggests

a mitochondrial origin for glycolysis in these organisms when considered with the mitochondrial targeted pre-sequence (Liaud et al., 2000; Nakayama et al., 2012).

The dramatic decrease in GAPDH His N activity compared to GAPDH His C could be a result of the position of the polyhistidine-tag. The N terminus of GAPDH attaches to the C terminus of TPI (Figure 6). The placement of the His-tag at the N terminus of GAPDH may inhibit interaction required between the TPI and GAPDH molecules.

4.3 TPI competes with GAPDH for GAP

Previous research has determined TPI to be a "perfect" diffusion rate-controlled enzyme. As such all known literature discussing glycolytic flux assumed TPI to thermodynamically equilibrate DHAP:GAP at the canonical 22:1 ratio instantly. GAPDH is left absent in the discussion of TPI enzyme kinetics, and vice versa (Albert et al., 2005). The results shown in figures 20,21 and 22 are therefore interesting. Initial rate comparisons showed reduction of 35% and 33% when 0.25mM GAP was equilibrated with B.TPI and S.TPI respectively, compared to when reactions were begun with GAP. This is a significant decrease in activity but is still much higher than expected for a 22-fold decrease in GAP availability. This might suggest that GAPDH is in competition with TPI for GAP. Further evidence to suggest this is the faster initial rates demonstrated when TPI is allowed to equilibrate DHAP prior to the start of the assay. Rate increases of 34%, 21%, 17% are observed for B.TPI, S.TPI, and the TPI-GAPDH fusion protein respectively ($P=0.002$, $P=0.015$, $P=0.001$). If TPI is equilibrating DHAP:GAP at a 22:1 ratio then a much higher initial rate is expected to be seen.

4.4 L-GAP competitively inhibits GAPDH glycolytic functions

Initial rate comparisons for GAPDH His C and the TPI-GAPDH fusion protein both demonstrated higher rates when using D-GAP as opposed to DL-GAP for substrate at 1mM and 2mM concentrations. Higher initial rates were observed for both enzymes at both concentrations when using only D-GAP as a substrate, supporting the findings of Brunner et al, 1998. GAPDH His C produced an average initial rate 18% faster for both 2mM D-GAP(P=0.066) and 1mM D-GAP(P=0.009). The GAPDH/TPI fusion enzyme produced an average initial rate 21% faster with 2mM D-GAP substrate (P=0.004) and 39% faster with 1mM D-GAP substrate (P=0.02). These results suggest possible competitive inhibition of GAPDH glycolytic activity by the L-GAP isomer. This is supported by the evidence that both enzymes tested the 1mM D-GAP registered faster rates than the 2mM DL-GAP. GAPDH His C achieved an average initial rate of 0.064 mM mMs^{-1} for 1mM D-GAP and 0.058 mM mMs^{-1} for 2mM DL-GAP. The fusion protein achieved an average initial rate of 0.301 mM mMs^{-1} for 1mM D-GAP and 0.244 mM mMs^{-1} for 2mM DL-GAP. This means the D-GAP and DL-GAP points were truly along the same curve then the [S] would have to be off by at least a factor of two. The data gathered in this research is limited, but there is some evidence to suggest that L-GAP interacts with and inhibits GAPDH.

4.5 TPI has a destabilising effect on TPI/GAPDH fusion protein

Thermal shift assays suggest that GAPDH is more stable by itself than when bound with TPI. Data from dissociation curves show lower melting temperatures for the TPI-GAPDH fusion protein than GAPDH His C on its own (figure 26). This result suggests that the fusion protein is less stable than GAPDH apart from TPI, and that TPI has a destabilising effect on the TPI-GAPDH fusion protein. The fusion protein appears to dissociate in two distinct stages. Transition peaks can be seen in the melt data at approximately 43°C and 52°C. The first peak seems to suggest a destabilized TPI moiety, which melts 10°C lower than B.TPI by itself. The second peak is much sharper and appears to correspond with GAPDH but dissociates approximately 25°C lower than GAPDH His C by itself.

4.6 GAPDH appears to stabilise TPI and may show evidence of enzyme Interaction

Evidence from dissociation curves suggests that GAPDH has a stabilising effect on TPI. When GAPDH His C and B.TPI were incubated together the TPI protein appeared to be stabilised by about 14°C, with the GAPDH molecule remaining stable (figure 26). The dissociation curve shows a unique melt pattern different to that of the individual GAPDH His C and B.TPI enzymes or the fusion protein. B.TPI appears to be stabilised by GAPDH. The GAPDH His C + B.TPI combination produces a slow melt from 35°C until about 61°C. Transition stages at approximately 69°C and 76°C may be indicative of potential protein-protein binding. If so, this is significant as no known current research has evidenced binding of TPI and GAPDH molecules. More research is needed to determine if this is the case or not.

4.7 NAD^+ stabilises the TPI-GAPDH fusion protein significantly but had no measurable effect on GAPDH His C

The TPI-GAPDH fusion protein was stabilised by binding NAD^+. The addition of NAD^+ to the fusion enzyme caused substantial increase in presumed TPI (43°C) and GAPDH (52°C) moiety stability (figure 27). A new melt peak at 64°C was also visualised. The Fusion + NAD^+ second peak occurs at about 64°C, while the fusion sample without NAD^+ peaks at about 52°C. The GAPDH His C protein was almost entirely unaffected by the presence of NAD^+, and appears to be stabilised by about 2°C. This data appears anomalous as NAD^+ is a known substrate of GAPDH and is sure to have been bound. However, one possible explanation could be that there is a potential for NAD^+ to come bound to GAPDH when purified as it is a co-enzyme. If this were the case, the GAPDH His C molecules may already be nearly saturated with NAD^+ resulting in little effect when exposed to more. This would also result in slightly flawed GAPDH enzyme concentration calculations as the presence of NAD^+ would have to be taken into account.

4.8 MMTS Incubation destabilises but does not inactivate TPI proteins containing active site cysteine.

TPI samples incubated for 120 minutes at 4°C with or without 1mM MMTS showed no comparative decrease in enzymatic activity (figure 29). After 120 minutes all samples tested demonstrated a natural decrease in enzyme activity, but no difference was observed between samples incubated in MMTS. Enzymatic assays performed using 1mM DHAP as substrate with B.TPI, S.TPI, and TPI-GAPDH fusion after 120 minutes without MMTS incubation returned 84%, 89%, and 82% activity compared to the control rate taken at T=0. In contrast samples with 120 minutes 1mM MMTS incubation returned 83%, 89%, and 81% enzyme activity, a difference of -

1%, 0%, and -1%. Contrary to Garza-Ramos, 1996, this data suggests that MMTS alone does not inhibit TPI activity.

The thermal stability results suggest a destabilization of proteins containing a cysteine active site resulting in dissociation at a lower temperature (figure 30). Dissociation curves of B.TPI, S.TPI, and GAPDH/TPI fusion proteins with MMTS show B.TPI and fusion protein destabilisation, both of which contain the cysteine 29 active site residue. The B.TPI enzyme began to dissociate when incubated at 30°C and reached a transition peak approximately 10°C sooner than B.TPI without MMTS. The fusion protein sample began to denature even prior to the initial incubation at 30°C. This is most likely a result of active site cysteine residues in the GAPDH moiety of the fusion enzyme. The TPI moiety of the fusion protein rapidly dissociated when the incubation temperature was increased to 30°C (figure 31). The S.TPI protein does not contain a reactive site cystine 29 residue and was unaffected by MMTS incubation.

4.9 Future Work

Repeating the enzyme assays performed in this study with more trials using multiple batches of protein would provide more reliable results. Producing data over a higher range of [S] would provide more accurate V_{max} and K_m estimates.

Producing data over a range of [S] would be beneficial in order to more accurately determine the relationship between D-GAP the TPI and GAPDH proteins. Furthermore, it would be great to equilibrate D-GAP and DHAP as a proof for potential inhibition. Fitting DL-GAP curves to Michaelis-Menten curves, then fitting to competitive inhibition curves would provide additional insight as to whether competitive inhibition is occurring. If the inhibition curve fits better it would

suggest competitive inhibition. This would be a special case as concentration inhibition (L) and [S] (D) are equal.

All enzyme assays performed in this study are pseudo first-order kinetic measurements performed with an excess of coenzyme NAD^+ substrate. It would be useful to perform similar assays with an excess of GAP or DHAP substrate and vary the amount of NAD^+ available. This would potentially provide new insight into the relationship between the TPI-GAPDH fusion protein and NAD^+.

Size exclusion chromatography would help to determine if the GAPDH and TPI enzymes are showing signs of interaction and binding to one another as the data produced was extremely limited. Unfortunately, time restrictions prevented this test from being performed, but it would be relatively quick and simple to perform. Size exclusion chromatography would also be beneficial to visualise how MMTS is destabilising the TPI-GAPDH fusion protein enzymes, as there has been no similar test performed to date.

Deeper biophysical analysis of the Fusion protein would be beneficial in order to elucidate potential differences in the flow of substrate between the TPI and GAPDH moieties. Several more TPI-GAPDH fusion proteins have recently been discovered in protists. Performing similar enzymatic and structural assays with enzymes from several different organisms would prove insightful. Identifying sequences of fusion proteins with large variations would increase likelihood of exhibiting different properties.

Bibliography

Albert, M. A., Haanstra, J. R., Hannaert, V., Van Roy, J., Opperdoes, F. R., Bakker, B. M. and Michels, P. A. (2005). Experimental and in silico analyses of glycolytic flux control in bloodstream form trypanosoma brucei. *J. Biol. Chem.* **280,** 28306-28315.

Alvarez, A., Martinez-Cadena, G., Silva, M., Avila, E. and Saavedra, E. (2007). Entamoeba histolytica: ADP-ribosylation of secreted glyceraldehyde-3-phosphate dehydrogenase. *Exp. Parasitol.* **117,** 349-356.

Aryan, S. (2015). Glycolysis explained in 10 easy steps. Retreived from microbiologyinfo.com

Banerjee, M., Balaram, H. and Balaram, P. (2009). Structural effects of a dimer interface mutation on catalytic activity of triosephosphate isomerase. *FEBS Journal* **276,** 4169-4183.

Bell, C. A., Cory, M., Fairley, T. A., Hall, J. E. and Tidwell, R. R. (1991). Structure- activity relationships of pentamidine analogs against giardia lamblia and correlation of antigiardial activity with DNA-binding affinity. *Antimicrob. Agents Chemother.* **35,** 1099-1107.

Berg, J. M., Tymoczko, J. L. and Stryer, L. (2007). Glycolysis and cluconeogenesis. In *Biochemistry (6th Edition)*: WH Freeman.

Bringaud, F., Rivière, L. and Coustou, V. (2006). Energy metabolism of trypanosomatids: Adaptation to available carbon sources. *Mol. Biochem. Parasitol.* **149,** 1-9.

Brunner, N., Brinkmann, H., Siebers, B., Hensel, R. (1998) NAD1-dependent Glyceraldehyde-3-phosphate Dehydrogenase from Thermoproteus tenax. Journal of Biological Chemistry 273, 6149-6156.gapdh

Cáceres, A., Michels, P. and Hannaert, V. (2010). Genetic validation of aldolase and glyceraldehyde-3-phosphate dehydrogenase as drug targets in trypanosoma brucei. *Mol. Biochem. Parasitol.* **169,** 50-54.

Cardoso, C. L., de Moraes, M. C., Guido, R. V. C., Oliva, G., Andricopulo, A. D., Wainer, I. W. and Cass, Q. B. (2008). The development of an immobilized enzyme reactor containing glyceraldehyde-3-phosphate dehydrogenase from trypanosoma cruzi: The effect of species' specific differences on the immobilization. *Analyst* **133,** 93-99.

Chuang, D., Hough, C. and Senatorov, V. V. (2005). Glyceraldehyde-3-phosphate dehydrogenase, apoptosis, and neurodegenerative diseases. *Annu. Rev. Pharmacol. Toxicol.* **45,** 269-290.

Clos, J. and Brandau, S. (1994). pJC20 and pJC40-two high-copy-number vectors for T7 RNA polymerase-dependent expression of recombinant genes in< i> escherichia coli. *Protein Expr. Purif.* **5,** 133-137.

Davidson, B., Sajgo, M., Noller, H. and Harris, J. I. (1967). Amino-acid sequence of glyceraldehyde 3-phosphate dehydrogenase from lobster muscle. *Nature* **216,** 1181-1185.

Der Giezen, M. V. (2011). Mitochondria and the rise of eukaryotes. *Bioscience* **61,** 594-601.

Elghareeb, AS et al. (2015) Laboratory diagnosis of Blastocystis spp. In diarrhic patients. Trop parasitol 2015; 5:36-41

Embley, T. M. and Martin, W. (2006). Eukaryotic evolution, changes and challenges. *Nature* **440,** 623-630.

Evans, K. (2014) Characterisation of the novel TPI-GAPDH fusion enzyme in blastocystis hominis. Unpublished

Ginger, M. L. (2006). Niche metabolism in parasitic protozoa. *Philos Trans R Soc Lond B Biol Sci* **361,** 101-118.

Glaser, P. E., Han, X. and Gross, R. W. (2002). Tubulin is the endogenous inhibitor of the glyceraldehyde 3-phosphate dehydrogenase isoform that catalyzes membrane fusion: Implications for the coordinated regulation of glycolysis and membrane fusion. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 14104-14109.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T. et al. (2009). Genome sequence and analysis of the irish potato famine pathogen phytophthora infestans. *Nature* **461,** 393-398.

Hammarton, T. C., Mottram, J. C. and Doerig, C. (2003). The cell cycle of parasitic protozoa: Potential for chemotherapeutic exploitation. *Prog. Cell Cycle Res.* **5,** 91- 101.

Jacobs, R. T., Nare, B. and Phillips, M. A. (2011). State of the art in african trypanosome drug discovery. *Curr Top Med Chem* **11,** 1255-1274.

Jacq, C., Alt-Morbe, J., Andre, B., Arnold, W., Bahr, A., Ballesta, J. P., Bargues, M., Baron, L., Becker, A., Biteau, N. et al. (1997). The nucleotide sequence of saccharomyces cerevisiae chromosome IV. *Nature* **387,** 75-78.

Jones, G. and Harris, J. I. (1972). Glyceraldehyde 3-phosphate dehydrogenase: Amino acid sequence of enzyme from baker's yeast. *FEBS Lett.* **22,** 185-189.

Kim, J. H., Lee, S., Park, J. B., Lee, S. D., Kim, J. H., Ha, S. H., Hasumi, K., Endo, A., Suh, P. and Ryu, S. H. (2003). Hydrogen peroxide induces association between glyceraldehyde 3-phosphate dehydrogenase and phospholipase D2 to facilitate phospholipase D2 activation in PC12 cells. *J. Neurochem.* **85,** 1228-1236.

Knowles, J. R. and Albery, W. J. (1977). Perfection in enzyme catalysis: The energetics of triosephosphate isomerase. *Acc. Chem. Res.* **10,** 105-111.

Kuntz, D. A., Osowski, R., Schudok, M., Wierenga, R. K., Müller, K., Kessler, H. and Opperdoes, F. R. (1992). Inhibition of triosephosphate isomerase from trypanosoma brucei with cyclic hexapeptides. *Eur. J. Biochem.* **207,** 441-447.

Kunz, S., Johner, A., Oberholzer, M., Wentzinger, L., Seebeck, T., Beavo, J. A., Laxman, S., D'Angelo, M.A., Flawia, M. M., Francis, S. H. et al. (2006). Cyclic nucleotide specific phosphodiesterases of the kinetoplastida: A unified nomenclature. *Mol. Biochem. Parasitol.* **145,** 133-135.

Liaud, M., Martin, W., Cerff, R., Lichtlé, C. and Apt, K. (2000). Compartment- specific isoforms of TPI and GAPDH are imported into diatom mitochondria as a fusion protein: Evidence in favor of a mitochondrial origin of the eukaryotic glycolytic pathway. *Mol. Biol. Evol.* **17,** 213-223.

Lõoke, M., Kristjuhan, K. and Kristjuhan, A. (2011). Extraction of genomic DNA from yeasts for PCR-based applications. *BioTechniques* 50, 325-328.

Lu, H. S., Yuan, P. M. and Gracy, R. W. (1984). Primary structure of human triosephosphate isomerase. *Journal of Biological Chemistry* **259,** 11958-11968.

Madigan, M., Martinko, J., Dunlap, P. and Clark, D. (2009). Antimicrobial agents used *in vivo*. In *Brock Biology of Microorganisms. 12th Edition* (ed. L. Berriman and G. Carlson), pp. 791-799. San Fransisco: Pearson Benjamin Cummings.

Mallo, N., Lamas, J. and Leiro, J. M. (2013). Hydrogenosome metabolism is the key target for antiparasitic activity of resveratrol against trichomonas vaginalis. *Antimicrob. Agents Chemother.* **57,** 2476-2484.

Marbach, A., Bettenbrock, K. (2012). Lac operon induction in Escherichia coli: systematic comparison of IPTG and TMG induction and influence of transacetylase LacA. *Journal of biotechnology.* 157(1):82-8

Michels, P., Herman, M., Hannaert, V. and Bringaud, F. (2006). Metabolic functions of glycosomes in trypanosomatids. *Biochimica et Biophysica Acta - Molecular Cell Research* **1763,** 1463-1477.

Moras, D., Olsen, K. W., Sabesan, M. N., Buehner, M., Ford, G. C. and Rossmann, M. G. (1975). Studies of asymmetry in the three-dimensional structure of lobster D-glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **250,** 9137-9162.

Mukherjee, S., Dutta, D., Saha, B. and Das, A. K. (2010). Crystal structure of glyceraldehyde-3-phosphate dehydrogenase 1 from methicillin-resistant staphylococcus aureus MRSA252 provides novel insights into substrate binding and catalytic mechanism. *J. Mol. Biol.* **401,** 949-968.

Müller, J. and Hemphill, A. (2011). Drug target identification in intracellular and extracellular protozoan parasites. *Curr Top Med Chem* **11,** 2029-2038.

Müller, M., Mentel, M., van Hellemond, J.,J., Henze, K., Woehle, C., Gould, S. B., Yu, R., van, d. G., Tielens, A. G. M. and Martin, W. F. (2012). Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev.* **76,** 444-495.

Nakajima, H., Amano, W., Kubo, T., Fukuhara, A., Ihara, H., Azuma, Y. T., Tajima, H., Inui, T., Sawa, A. and Takeuchi, T. (2009). Glyceraldehyde-3- phosphate dehydrogenase aggregate formation participates in oxidative stress- induced cell death. *J. Biol. Chem.* **284,** 34331-34341.

Nakayama, T., Ishida, K. and Archibald, J. M. (2012). Broad distribution of TPI- GAPDH fusion proteins among eukaryotes: Evidence for glycolytic reactions in the mitochondrion? *PLoS ONE* **7,** 1-12.

Nishimura, A., Morita, M., Nishimura, Y. and Sugino, Y. (1990). A rapid and highly efficient method for preparation of competent escherichia coli cells. *Nucleic Acids Res.* **18,** 6169.

Nowak, K., Kuczek, M., Ostropolska, L., Malarska, A., Wolny, M. and Baranowski, T. (1975). The covalent structure of glyceraldehyde-phosphate dehydrogenase from human muscles. isolation and amino acid sequences of peptides from tryptic digest. *Hoppe Seylers Z. Physiol. Chem.* **356,** 1181-1183.

Oesper, P. (1954). The mechanism of action of glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **207,** 421-429.

Omar, A. (2006). The epidemiology of *Blastocystis hominis* in the united states. *Research Journal of Parasitology* **1,** 1-10.

Orosz, F., Oláh, J. and Ovádi, J. (2009). Triosephosphate isomerase deficiency: New insights into an enigmatic disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1792,** 1168-1174.

Garza-Ramos, G., PErez-Montfort, R., Rojo-DomInguez, A., GOmez-Puyou, M. and Gomez-Puyou, A. (1996). Species-Specific Inhibition of Homologous Enzymes by Modification of Nonconserved Amino Acids Residues. The cysteine residues of triosephosphate isomerase. *European Journal of Biochemistry* 241, 114-120.

Roger, A. J. and Hug, L. A. (2006). The origin and diversification of eukaryotes: Problems with molecular phylogenetics and molecular clock estimation. *Philosophical Transactions: Biological Sciences*, 1039.

Ronai, Z. (1993). Glycolytic enzymes as DNA binding proteins. *Int. J. Biochem.* **25,** 1073-1076.

Sirover, M. A. (1997). Role of the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in normal cell function and in cell pathology. *J. Cell. Biochem.* **66,** 133.

Sirover, M. A. (2014). Structural analysis of glyceraldehyde-3-phosphate dehydrogenase functional diversity. *The international journal of biochemistry & cell biology*, *57*, 20–26.

Sleator, R. (2013). A Beginner's Guide to Phylogenetics. *Microbial Ecology* 66, 1-4.

Stanford, N. J., Lubitz, T., Smallbone, K., Klipp, E., Mendes, P. and Liebermeister, W. (2013). Systematic construction of kinetic models from genome-scale metabolic networks. *PloS one* **8,** e79195.

Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. (1990). Use of t7 RNA polymerase to driect expression of cloned genes. *Methods in Enzymology* 185:60-89

Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A. and Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 19333-19338.

Tan, K. S. W. (2008). New insights on classification, identification, and clinical relevance of blastocystis spp. *Clin. Microbiol. Rev.* **21,** 639-665.

Trentham, D. R. (1968). Aspects of the chemistry of D-glyceraldehyde 3- phosphate dehydrogenase. *Biochem. J.* **109,** 603-612.

Wierenga, R. K., Kapetaniou, E. G. and Venkatesan, R. (2010). Triosephosphate isomerase: A highly evolved biocatalyst. *Cell Mol. Life Sci.* **67,** 3961-3982.

Winter, R. W., Kelly, J. X., Smilkstein, M. J., Dodean, R., Hinrichs, D. and Riscoe, M. K. (2008). Antimalarial quinolones: Synthesis, potency, and mechanistic studies. *Exp. Parasitol.* **118,** 487-497.

Wiser, M. F. (2011). *Protozoa and Human Disease.*: Garland Science, Taylor and Francis Group, LLC.

Zhang, Z., Sugio, S., Komives, E. A., Liu, K. D., Knowles, J. R., Petsko, G. A. and Ringe, D. (1994). Crystal structure of recombinant chicken triosephosphate isomerase-phosphoglycolohydroxamate complex at 1.8-. ANG. resolution. *Biochemistry (N. Y.)* **33,** 2830-2837.

Appendices

Appendix 1: TPI Sequence supplied by Mark Van Der Giezen

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCA

TATGTTCGTCGGTGGCAATTGGAAGTGCAACGGTTCTCTCTCCAAGGTCCAGGAGATCG

TGGCTACTCTCAACAACTCCAACCTCAACAACGACGCTGAGGTCGTGATCGCCCCTCCC

ACCGCCTATCTGAGAGACACCGTGGCTACCGTGCGCGCTGATGTGCAGGTGGCTGCTCA

GGACGTCTGGTCCCAGGGCAACGGTGCCTTCACCGGCGAGACCTCTGCTGAAATGCTGA

AGGATCTCAAGGTCGGTTGGGCGATCGTGGGCCACTCCGAGCGCCGTGGCAAGGGCGAG

AGCGATGCCGAGGTTGCGACCAAGGCTGCTTACGCTCAGAAGAACGGTCTGAAGGTGAT

CTGCTGCCTGGGCGAGTCCCTGAAGGAGCGCGAGGCTGGCCGCTTCGCCGAGGTGGTGA

CCCGCCAGCTCAAGGCCTACGCCGACGCCATCAAGAACTGGGATGACGCCGTGATCGCC

TACGAGCCCATCTGGGCCATTGGCACGGGCAAGACCGCCACCCCCGCTCAGGCGGAGGA

GGTGCACGCCGTACTGCGCAAGTGGCTGCGCGACAATGTGAGCGCCGCCGTGGCGGACA

AGGTGCGCATCATCTACGGAGGCTCCGTGAACGCGAAGAACTGCAACGAGCTGGGTAAG

CAGGCCGACATCGACGGCTTCCTGGTGGGCGGCGCTTCGCTGAAGCCCGAGTTCGTGAA

CATCATCAACGCCAACTCCACGGTGCAGAACGCGTAA

Appendix 2: GAPDH Sequence supplied by Mark Van Der Giezen

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCA

TATGCCCGTGAGCGTGGGTATCAACGGCTTCGGCCGCATTGGTCGTCTGGTGTTCCGCG

CCAGCCAGCCGAACCCGCTGGTGAACGTGGTGGCGATCAACGACCCGTTCATCACGCCC

GACTACATGGAGTACATGATCATGCACGACTCCACGCACGGTCCGTTCCAGGGCACCGT

GAAGGCTGAGAAGGACGCGATCATCGTGAACGGCCGCCGCATCGTGGTGTCCAACGAGA

TGGACCCGAAGAAGATCCAGTGGGGCGCTGCGGGTGCGGAGTACATCGTGGAGTCCACG

GGCGTGTTCACGGCGAAGGACAAGGCCGCGCAGCACCTGGAGGGCGGCGCGAAGAAGG

TGGTGATTTCCGCGCCGTCGAAGGACGCTCCGATGTTCGTGATGGGCGTGAACAACACGA

CGTACACGAAGGACCTGAAGGTGGTGTCGAACGCGTCCTGCACGACGAACTGCCTGGCG

CCGCTGGCGAAGATCGTGAATGACAAGTTCGGTCTGAAAGAGGGTCTGATGACGACGG

TGCACTCGGTGACCTCCACGCAAAAGGTGCTGGACGGTCCTTCCAAGAAGGACTGGAGA

GGCGGCCGCTCTGCCTGCTACAACATCATCCCGTCCTCCACGGGCGCCGCCAAGGCCGT

GGGCAAGGTCATTCCCGAGCTGAACGGCAAGCTGACGGGCATGTCCTTCCGCGTGCCGA

CGGAGGACGTCTCCGTGGTGGATCTGACCTGCACGCTGAAGAAGCCCGCCACCTACGAG

C AGATCAAGGCTGCGGTGAAGGAGGCTTCGGAGACGTACATGAAGGGCATTATGGGATA

CGTGGATTACGACGTGGTGTCTCGCGATCTGCTGACCTGCCCGTACTCTTCTGTCTTCG

ACGCGAAGGCAGGAATTGCCCTGAACGACACGTTCGTGAAGCTGGTTTCTTGGTACGAC

AACGAGTGGGGCTACTCCAACCGTATGGTCGACCTCATCCAGTACATGGCGAAGGTGGA

TCGCTCTTAA

Appendix 3: wtTPI-GAPDH Sequence from Mark Van Der Giezen

ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCA
TATGTTCGTCGGTGGCAATTGGAAGTGCAACGGTTCTCTCTCCAAGGTCCAGGAGATCG
TGGCTACTCTCAACAACTCCAACCTCAACAACGACGCTGAGGTCGTGATCGCCCCTCCC
ACCGCCTATCTGAGAGACACCGTGGCTACCGTGCGCGCTGATGTGCAGGTGGCTGCTCA
GGACGTCTGGTCCCAGGGCAACGGTGCCTTCACCGGCGAGACCTCTGCTGAAATGCTGA
AGGATCTCAAGGTCGGTTGGGCGATCGTGGGCCACTCCGAGCGCCGTGGCAAGGGCGAG
AGCGATGCCGAGGTTGCGACCAAGGCTGCTTACGCTCAGAAGAACGGTCTGAAGGTGAT
CTGCTGCCTGGGCGAGTCCCTGAAGGAGCGCGAGGCTGGCCGCTTCGCCGAGGTGGTGA
CCCGCCAGCTCAAGGCCTACGCCGACGCCATCAAGAACTGGGATGACGCCGTGATCGCC
TACGAGCCCATCTGGGCCATTGGCACGGGCAAGACCGCCACCCCCGCTCAGGCGGAGGA
GGTGCACGCCGTACTGCGCAAGTGGCTGCGCGACAATGTGAGCGCCGCCGTGGCGGACA
AGGTGCGCATCATCTACGGAGGCTCCGTGAACGCGAAGAACTGCAACGAGCTGGGTAAG
CAGGCCGACATCGACGGCTTCCTGGTGGGCGGCGCTTCGCTGAAGCCCGAGTTCGTGAA
CATCATCAACGCCAACTCCACGGTGCAGAACGCGGGCCCCGTGAGCGTGGGTATCAACG
GCTTCGGCCGCATTGGTCGTCTGGTGTTCCGCGCCAGCCAGCCGAACCCGCTGGTGAAC
GTGGTGGCGATCAACGACCCGTTCATCACGCCCGACTACATGGAGTACATGATCATGCA
CGACTCCACGCACGGTCCGTTCCAGGGCACCGTGAAGGCTGAGAAGGACGCGATCATCG
TGAACGGCCGCCGCATCGTGGTGTCCAACGAGATGGACCCGAAGAAGATCCAGTGGGGC
GCTGCGGGTGCGGAGTACATCGTGGAGTCCACGGGCGTGTTCACGGCGAAGGACAAGGC
CGCGCAGCACCTGGAGGGCGGCGCGAAGAAGGTGGTGATTTCCGCGCCGTCGAAGGAC
GCTCCGATGTTCGTGATGGGCGTGAACAACACGACGTACACGAAGGACCTGAAGGTGGTG
TCGAACGCGTCCTGCACGACGAACTGCCTGGCGCCGCTGGCGAAGATCGTGAATGACAA
GTTCGGTCTGAAAGAGGGTCTGATGACGACGGTGCACTCGGTGACCTCCACGCAAAAGG
TGCTGGACGGTCCTTCCAAGAAGGACTGGAGAGGCGGCCGCTCTGCCTGCTACAACATC
ATCCCGTCCTCCACGGGCGCCGCCAAGGCCGTGGGCAAGGTCATTCCCGAGCTGAACGG
CAAGCTGACGGGCATGTCCTTCCGCGTGCCGACGGAGGACGTCTCCGTGGTGGATCTGA
CCTGCACGCTGAAGAAGCCCGCCACCTACGAGCAGATCAAGGCTGCGGTGAAGGAGGCT
TCGGAGACGTACATGAAGGGCATTATGGGATACGTGGATTACGACGTGGTGTCTCGCG
ATCTGCTGACCTGCCCGTACTCTTCTGTCTTCGACGCGAAGGCAGGAATTGCCCTGAAC
GACACGTTCGTGAAGCTGGTTTCTTGGTACGACAACGAGTGGGGCTACTCCAACCGTAT
GGTCGACCTCATCCAGTACATGGCGAAGGTGGATCGCTCTTAA

Appendix 4: ProtParam data for TPI

Number of amino acids: 267
Molecular weight: 28556.10
Theoretical pI: 6.76
Amino acid composition:

Ala (A)	36	13.5%
Arg (R)	11	4.1%
Asn (N)	18	6.7%
Asp (D)	13	4.9%
Cys (C)	4	1.5%
Gln (Q)	9	3.4%
Glu (E)	16	6.0%
Gly (G)	24	9.0%
His (H)	9	3.4%
Ile (I)	13	4.9%
Leu (L)	16	6.0%
Lys (K)	17	6.4%
Met (M)	3	1.1%
Phe (F)	5	1.9%
Pro (P)	6	2.2%
Ser (S)	17	6.4%
Thr (T)	12	4.5%
Trp (W)	6	2.2%
Tyr (Y)	5	1.9%
Val (V)	27	10.1%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%

(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 29
Total number of positively charged residues (Arg + Lys): 28

Atomic composition:

Carbon	C	1248
Hydrogen	H	1980
Nitrogen	N	368
Oxygen	O	387
Sulfur	S	7

Formula: $C_{1248}H_{1980}N_{368}O_{387}S_7$
Total number of atoms: 3990

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 40700
Abs 0.1% (=1 g/l) 1.425, assuming all pairs of Cys residues form cystines

Ext. coefficient 40450
Abs 0.1% (=1 g/l) 1.417, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 16.30
This classifies the protein as stable.

Aliphatic index: 85.17

Grand average of hydropathicity (GRAVY): -0.249

Appendix 5: ProtParam Data for GAPDH

Number of amino acids: 356
Molecular weight: 38711.40
Theoretical pI: 8.88
Amino acid composition:

Ala (A)	28	7.9%
Arg (R)	12	3.4%
Asn (N)	16	4.5%
Asp (D)	21	5.9%
Cys (C)	5	1.4%
Gln (Q)	7	2.0%
Glu (E)	13	3.7%
Gly (G)	29	8.1%
His (H)	11	3.1%
Ile (I)	19	5.3%
Leu (L)	19	5.3%
Lys (K)	28	7.9%
Met (M)	14	3.9%
Phe (F)	10	2.8%
Pro (P)	17	4.8%
Ser (S)	29	8.1%
Thr (T)	23	6.5%
Trp (W)	4	1.1%
Tyr (Y)	13	3.7%
Val (V)	38	10.7%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%

(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 34
Total number of positively charged residues (Arg + Lys): 40

Atomic composition:

Carbon	C	1714
Hydrogen	H	2717
Nitrogen	N	469
Oxygen	O	513
Sulfur	S	19

Formula: $C_{1714}H_{2717}N_{469}O_{513}S_{19}$
Total number of atoms: 5432

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 41620
Abs 0.1% (=1 g/l) 1.075, assuming all pairs of Cys residues form cystines

Ext. coefficient 41370
Abs 0.1% (=1 g/l) 1.069, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (*Escherichia coli*, in vivo).

Instability index:

The instability index (II) is computed to be 28.99
This classifies the protein as stable.

Aliphatic index: 80.45

Grand average of hydropathicity (GRAVY): -0.173

Appendix 6: ProtParam data for TPI-GAPDH

Number of amino acids: 603 **Molecular weight:** 65012.0 **Theoretical pI:** 8.36

Amino acid composition:

Ala (A) 64 10.6% Arg (R) 22 3.6% Asn (N) 34 5.6% Asp (D) 34 5.6% Cys (C) 9 1.5% Gln (Q) 16 2.7% Glu (E) 29 4.8% Gly (G) 51 8.5% His (H) 13 2.2% Ile (I) 32 5.3% Leu (L) 34 5.6% Lys (K) 45 7.5% Met (M) 15 2.5% Phe (F) 15 2.5% Pro (P) 22 3.6% Ser (S) 41 6.8% Thr (T) 35 5.8% Trp (W) 10 1.7% Tyr (Y) 18 3.0% Val (V) 64 10.6% Pyl(O) 0 0.0% Sec(U) 0 0.0%

(B) 0 0.0% (Z) 0 0.0% (X) 0 0.0%

Total number of negatively charged residues (Asp + Glu): 63 **Total number of positively charged residues (Arg + Lys):** 67

Atomic composition:

Carbon C Hydrogen
Nitrogen N 800 Oxygen O 874

Sulfur S 24

2869 H 4558

Formula: $C_{2869}H_{4558}N_{800}O_{874}S_{24}$ **Total number of atoms:** 9125

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 82320
Abs 0.1% (=1 g/l) 1.266, assuming all pairs of Cys residues form cystines

Ext. coefficient 81820
Abs 0.1% (=1 g/l) 1.259, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro). >20 hours (yeast, in vivo).

>10 hours (*Escherichia coli*, in vivo).

Instability index:

The instability index (II) is computed to be 22.59 This classifies the protein as stable.

Aliphatic index: 84.08
Grand average of hydropathicity (GRAVY): -0.177

Appendix 7: Data

Table of Averages GAP Substrate:

GAP Conc.	HisC	HisN	Fusion	HisC B.TPI
1	0.085507	0.034833	0.28265	0.019935
0.5	0.070327	0.028307	0.18645	0.014625
0.25	0.044	0.022733	0.112717	0.008955
0.125	0.026987	0.012613	0.0568	0.00702
0.0625	0.016133	0.0077	0.029067	0.00396
0.03125	0.009093	0.005427	0.015133	0.003105

Table of Averages DHAP Substrate:

DHAP Conc (mM)	STPI	BTPI	Fusion
5	0.07976	0.08232	0.8605
2.5	0.04552	0.0456	0.574
1.25	0.02864	0.02928	0.3723
0.625	0.01576	0.01448	0.1741
0.3125	0.00784	0.00736	0.0893
0.1563	0.00408	0.00312	0.0503

Table of Averages Equilibraion Assay

	1mM GAP			Average
B.TPI	0.021	0.019	0.019	0.020
S.TPI	0.014	0.018	0.015	0.015
Fusion	0.285	0.278	0.282	0.282
	1mM DHAP			Average
B.TPI	0.020	0.021	0.020	0.020
S.TPI	0.020	0.017	0.018	0.018
Fusion	0.471	0.489	0.482	0.481

Table of averages 0.25mM GAP Comparison

NAD^+ Start				
	1min			Average
HisC	0.051	0.038	0.042	0.044
B.TPI+HisC	0.007	0.008	0.007	0.007
S.TPI+HisC	0.007	0.006	0.006	0.006
Fusion	0.089	0.098	0.084	0.090
GAP Start				
HisC	0.053	0.040	0.045	0.045
B.TPI+HisC	0.012	0.010	0.010	0.011
S.TPI+HisC	0.010	0.009	0.009	0.009
Fusion	0.145	0.136	0.136	0.139

Table of Averages 5mM DHAP Comparisons

DHAP Start	1min			Average
B.TPI +HisC	0.08208	0.08304	0.07416	0.080
S.TPI + HisC	0.08784	0.07776	0.08136	0.082
Fusion	0.502	0.526	0.519	0.516
NAD^+ Start				
	1min			Average
B.TPI +HisC	0.110	0.104	0.106	0.107
S.TPI + HisC	0.096	0.097	0.105	0.099
Fusion	0.600	0.616	0.598	0.605

Table of Averages Isomer

	D GAP HisC	DL GAP HisC	D GAP Fusion	DL GAP Fusion
2mM	0.069	0.05843333	0.2944	0.244
1mM	0.06373333	0.054	0.3007	0.21593333

Table of Averages MMTS

	0 Min	120min Control	120min MMTS
B.TPI	0.020	0.017	0.017
S.TPI	0.018	0.016	0.016
Fusion	0.302	0.248	0.245

Appendix 8: DHAP Certificate of Analysis:

SIGMA-ALDRICH

3050 Spruce Street, Saint Louis, MO 63103 USA
Email USA: techserv@sial.com Outside USA: eurtechserv@sial.com

Certificate of Analysis

Product Name:	DIHYDROXYACETONE PHOSPHATE HEMIMAGNESIUM SALT HYDRATE >= 95 % TLC
Product Number:	51269
Batch Number:	BCBT9318
Brand:	Sigma
CAS Number:	57-04-5
Formula:	$C_3H_5Mg_{0.5}O_6P * xH_2O$
Formula Weight:	180.19
Storage Temperature:	
Quality Release Date:	24 APR 2017
Recommended Retest Date:	AUG 2019

TEST	SPECIFICATION	RESULT
APPEARANCE (COLOR)	COLORLESS OR WHITE	WHITE
APPEARANCE (FORM)	SOLID	SOLID
PURITY (TLC AREA %)	≥ 95 %	96 %
CARBON CONTENT	20.00 % (THEORY, ANHYDROUS BASIS)	15.82 %
HYDROGEN CONTENT	2.80 % (THEORY, ANHYDROUS BASIS)	4.52 %
REM. ON ELEMENTAL ANALYSIS	CORRESPONDS TO WATER XX %	20.89 %
PROTON NMR SPECTRUM	CONFORMS TO STRUCTURE	CONFORMS
PHOSPHORUS 31 NMR SPECTRUM	CONFORMS TO STRUCTURE	CONFORMS

Dr. Claudia Geitner
Manager Quality Control
Buchs, Switzerland

Sigma-Aldrich warrants that at the time of the quality release or subsequent retest date this product conformed to the information contained in this publication. The current specification sheet may be available at Sigma-Aldrich.com. For further inquiries, please contact Technical Service. Purchaser must determine the suitability of the product for its particular use. See reverse side of invoice or packing slip for additional terms and conditions of sale.