X-RAY CRYSTALLOGRAPHIC ANALYSIS OF THE PYROPHOSPHATE-DEPENDENT PHOSPHOFRUCTOKINASE OF *SPIROCHAETA THERMOPHILUM*

A thesis presented in partial fulfillment of the requirements for the degree of Master of Science in Biochemistry at Massey University, New Zealand.

Andrew James Welham

2002

"I think, therefore I am"

René Descartes

ABSTRACT

The structure of a homodimeric, non-allosteric, PP_i-dependent phosphofructokinase from the thermophilic bacterium *Spirochaeta thermophilum* has been resolved by X-ray crystallography in two distinct conformations at 2.2 (R = 0.1991 [R_{free} = 0.2288]) and 1.85 Å (R = 0.1923 [R_{free} = 0.2035]) resolution. The 554 residue (M_r 61080 $g.mol^{-1}$) subunit, a homologue of the plant PP_i-PFK β-subunit exhibits an asymmetrical quaternary structure and shares both sequence and tertiary structure with the N- and C-terminal Rossmann-like domains of prokaryotic ATP-PFKs. *Spirochaeta thermophilum* PP_i-PFK exhibits three major inserts relative to the prokaryotic ATP-PFK of *E. coli*, an N-terminal insert, a C-terminal insert, and an insert within the PFK C-terminal domain which forms an autonomous α-helical domain. The active site is formed at the interface of the N and C domains. The 'open' and 'closed' subunit asymmetry of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model mirrors that of the *B. burgdorferi* PP_i-PFK (1KZH [Moore et al.2002]) with the exception that the two unique β–hairpins (380-390 [α16-α17] and 485-495 [β14-β15]) of subunit A are not displaced into the active site. Both subunits of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model adopt an 'open', apparently inactive conformation. The conformational change involves concomitant closure of the active site of both subunits via a rigid-body displacement of the C and α-helical domains, relative to the N domain. The N domain of one subunit and the C domain of the opposing subunit can be thought of as a rigid body, therefore closure of one active site dictates closure of the other. Rotation of the small domain forces Met251 of the MGR motif to adopt an active conformation and displacement of the α-helical domain, specifically the 380-390 β-hairpin into the active site 'folds' Arg253 (MG**R**) into an active conformation. Closure of the active site, which prevents wasteful hydrolysis, involves movement of the β14-β15 β-hairpin into the active site and simultaneous rearrangement of the PP_i-binding GGDD motif. The conformational change of the *S. thermophilum* PP_i-PFK is surprisingly complex and unique relative to prokaryotic ATP-PFKs and involves displacement of novel structural elements. These movements change the conformation of conserved motifs at the active site and therefore function to modulate PP_i-dependent activity.

ACKNOWLEDGEMENTS

This thesis is representative of the efforts of a number of people to whom I am indebted. Primarily I would like to thank my supervisor Dr. Stan Moore for his unbounded patience, encouragement, and support. Unforeseeably, giving me the opportunity to undertake this research also ceded a passion for science and research, a life long pursuit.

I would like to thank Dr. Ron Ronimus and Dr. Hugh Morgan from the Thermophile Research Unit, University of Waikato for the purified *Spirochaeta thermophilum* pyrophosphate-dependent phosphofructokinase.

I would like to thank Dr. Kathryn Stowell for her immense patience in listening to my ideas and the proof reading of this thesis.

I would like to thank past and present members of the Institute of Molecular Biosciences and the Institute of Fundamental Sciences for there tutorage, particularly Assoc. Prof. John Tweedie, Dr. Mark Patchet, Prof. Pat Sullivan, Dr. Gill Norris, Dr. Max Scott, Assoc. Prof. Geoff Jameson, Dr. Bryan Anderson, Dr. Emily Parker, and last but not least Dr. Simon Brown (an inspirational teacher).

I gratefully acknowledge financial support from the award of a Massey University Masterate Scholarship.

Thank you to my family: Mum, Uncle Howard, Aunt Beth, Gaylyn, Sylvia, Shaun, Martyn, Pieter, Jeannie and Michelle for your perpetual love, support, patience and faith in me. I will always be indebted.

TABLE OF CONTENTS

List of Figures

Chapter 1 Introduction and Literature Review

Chapter 3 Results: Data Quality and Atomic Model Statistics

Chapter 4 *Spirochaeta thermophilum* PP_i-PFK: Structure and Function

List of Tables

Abbreviations

$\|Fc\|$	Structure factor (calculated)
$\|F_o\|$	Structure factor (observed)
1,3-BPG	1,3-bisphosphoglycerate
2-PG	2-Phosphoglycerate
3-PG	3-Phosphoglycerate
Å	Angstrom (10^{-10}m)
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
APS	Ammonium peroxidisulphate
ATP	Adenosine triphosphate
CC	Correlation Coefficient
Da	Dalton
DHAP	Dihydroxyacetone phosphate
ED	Entner Doudoroff Pathway
F16bP	Fructose 1,6-bisphosphate
F26bP	Fructose 2,6-bisphosphate
F6P	Fructose 6-phosphate
FBPase	Fructose bisphosphatase
FOM	Figure of merit
G6P	Glucose 6-phosphate
GAP	Glyceraldehyde-3-phosphate
GAP:FdOR	Glyceraldehyde-3-phosphate: ferrodoxin oxidoreductase
GLK	Glucokinase
Hepes	N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
M_r	Molecular mass (g. mol^{-1})
NAD	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
PEG	Polyethylene Glycol
PEP	Phosphoenolpyruvate
PFK	Phosphofructokinase
PGI	Phosphoglucose isomerase
PP_i	Pyrophosphate
PPP	Pentose-Phosphate Pathway
RMS	Root mean squared
TCA	Tricarboxylic acid cycle
TEMED	N,N,N',N'-tetramethylethylenediamine
TIM	Triose phosphate isomerase
Tris	Tris(hydroxymetyl)aminomethane
$\Delta G^{o'}$	Free energy change

Amino Acids

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic acid
Cys	C	Cystine
Gln	Q	Glutamine
Glu	E	Glutamic acid
Gly	G	Glycine
His	H	Histidine
Ile	I	Iso-leucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine

Chapter 1

Introduction and Literature Review

1.1 Metabolism

Maintenance of life requires the coupling of free energy (Eqn. 1) gained through exergonic processes with endergonic processes, for instance the coupling of catabolic reactions to anabolic reactions, mechanical work, or active transport.

$$\Delta G = \Delta H - T\Delta S$$

(Eqn.1.1)

Eqn.1.1: The free energy change (ΔG) of a reaction. ΔH is the change in enthalpy term, ΔS is the change in entropy term, and T the absolute temperature. If ΔG is negative the process is spontaneous.

Heterotrophic and photoautotrophic organisms oxidize organic metabolic fuels to release free energy (photoautotrophs are autotrophic in that they synthesize their own carbohydrates from CO_2 via photosynthesis). The free energy released is often not coupled directly to an endergonic process, rather the exergonic process is coupled to the immediate synthesis of a ATP which has a high energy of hydrolysis, NADH, or NADPH, which have a high redox potential. Endergonic processes can subsequently be driven via the coupled cleavage of a phosphoanhydride bond in ATP or the oxidation of NADH or NADPH. In biological systems endergonic and exergonic processes are predominantly catalyzed via multistep enzymatic pathways, such as glycolysis.

1.2 Glycolysis

Glycolysis (Greek: *glykus*, sweet + *lysis*, loosening), alternatively known as the Embden-Meyerhof pathway is a sequence of enzyme-catalyzed reactions responsible for the initial stages of glucose catabolism, which occurs in the cell cytosol.

A. **B.**

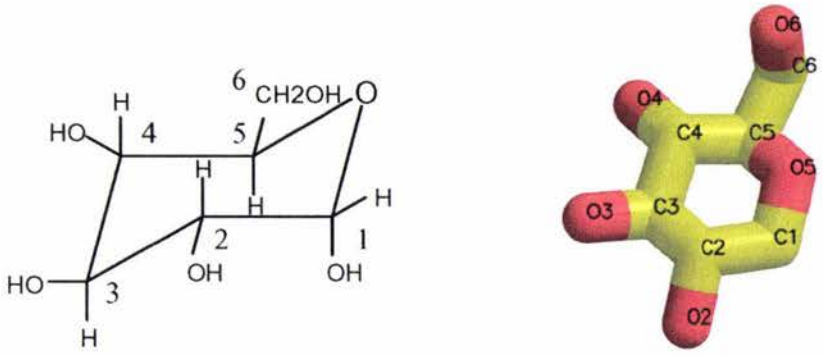


Figure 1.1: **A**.Chair conformation of the hexose glucose. **B**. Ball and stick diagram of beta-D-glucose. Figure prepared using MOLSCRIPT (Kraulis 1991) and Raster3D (Merrit and Bacon 1997).

The glycolytic/gluconeogenic pathway, along with the oxidative tricarboxylic acid cycle (TCA), the reductive pentose phosphate pathway (PPP), and the Entner Doudoroff pathway (ED) are found in all three domains of the phylogenetic tree and are considered central to the origins of metabolism (Selig *et al.* 1997; Ronimus and Morgan 2002)[1]. The overall reaction of the glycolytic pathway involves the conversion of a glucose molecule into two pyruvate molecules with the concomitant generation of two NADH molecules and two ATP molecules via a series of 10 enzymatic reactions (see Eqn. 1.2; see Figure 1.2).

$$\text{Glucose} + 2\ NAD^+ + 2\ ADP + 2\ P_i \longrightarrow$$
$$2\ \text{Pyruvate} + 2\ NADH + 2\ ATP + 2\ H_2O + 4\ H^+$$

(Eqn.1.2)

The glycolytic pathway can be split into two phases. The primary phase involves the consumption of two ATP molecules and the conversion of a hexose into two triose molecules. The secondary phase is responsible for substrate-level phosphorylation; the generation of ATP from ADP via phosphoryl transfer from a high-energy metabolite such as 1,3-bisphosphoglycerate (13BPG) (see Figure 1.2). For an exhaustive discussion of the glycolytic reactions and the pathway's regulation see Voet *et al.* (1999). Phosphofructokinase (PFK, EC 2.7.1.11) catalyzes the third reaction in the glycolytic pathway, the irreversible phosphoryl-transfer (see section 1.4) from ATP to the C-1 of fructose 6-phosphate (F6P) to form β-D-fructose 1,6-bisphosphate (F16bP) ($\Delta G^{\circ\prime}$= -18.45

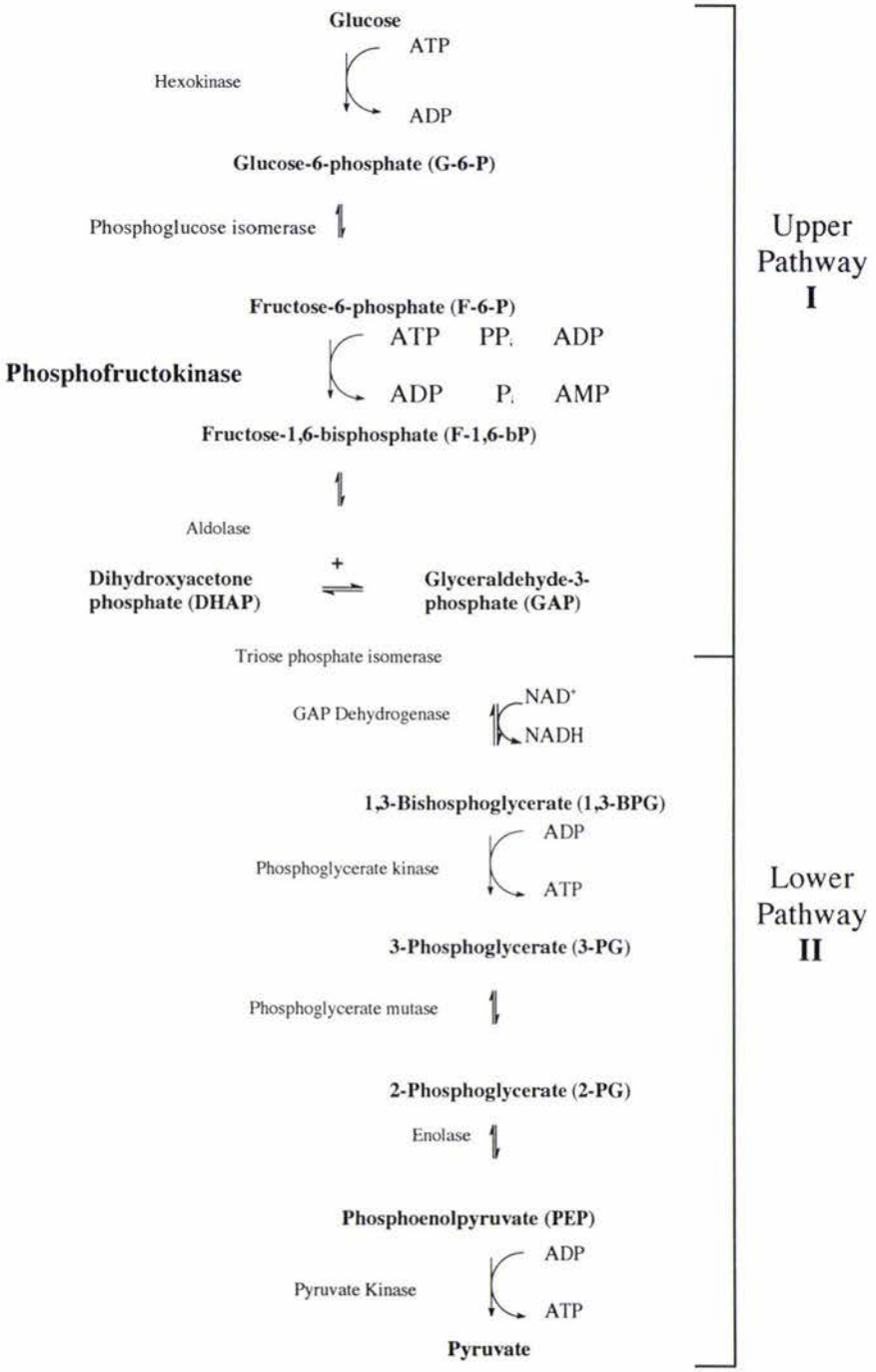


Figure 1.2: Embden-Meyerhof pathway. Reactions of the glycolytic pathway can be split into phase I (upper pathway) and phase II (lower pathway). Note PFKs exist which utilize alternative phosphoryl donors, including PP_i and ADP.

$kJmol^{-1}$). Fructose-1,6-bisphosphatase (FBPase) catalyses the reverse, gluconeogenic reaction. PFK catalyses the first glycolytic step in the metabolism of hexoses other than glucose, which enter the pathway by conversion to F6P. PFK therefore catalyses the first committed step of the glycolytic pathway. PFKs are generally subject to allosteric regulation, as such they play a key role in the regulation of glycolytic flux (Voet et al. 1999). Prokaryotic tetrameric ATP-dependent PFKs (ATP-PFKs) are generally allosterically activated by ADP and inhibited by phosphoenolpyruvate (PEP) and bind F6P cooperatively (Shirakihara and Evans 1988). Mammalian PFKs are also tetramers, although their molecular weight is higher and their regulation more complex (Kemp and Gunasekera 2002).

The existence of multiple gene sequence families encoding glucokinases (GLK), phosphoglucose isomerases (PGI), and phosphorfructokinases (PFK) at the beginning of the glycolytic pathway compared to the highly conserved and universal distribution of the trunk pathway (the last five reactions of glycolysis) enzymes throughout the phylogenetic tree is interpreted as evidence for a anabolic, gluconeogenic, 'bottom-up' evolution of the glycolytic pathway (Ronimus and Morgan 2002)[1].

1.3 Modified Embden-Meyerhof Pathways

The conservation of the glycolytic pathway throughout the phylogenetic tree is not strictly mirrored by the pathway's constitutive enzymes. The domain archaea, credited with containing the most ancient organisms displays the greatest diversity with respect to glucose catabolism, exhibiting traditional and modified Embden-Meyerhof pathways (Ronimus and Morgan 2002)[1]. Diversity of the nearly ubiquitous pathway involves novel reactions, enzymes and methods of control, including PP_i-PFKs, ADP-PFKs, ADP-GLKs, and a novel glyceraldehyde-3-phosphate ferrodoxin oxidoreductase (GAP:*Fd*OR) (Ronimus and Morgan 2001; Verhees *et al.* 2002; Verhees *et al.* 2001; Mukund and Adams 1995). PFK exemplifies this diversity. Evolutionary divergence of PFKs is likely to be reflective of the varying complexities of glycolytic regulation within eukarya, bacteria, and archaea.

1.3.1 Pyrophosphate-dependent Glycolysis

As mentioned earlier phosphofructokinase catalyses the committed step of the glycolytic pathway, the phosphorylation of F6P to form F16bP. ATP is the near-universal phosphoryl donor, however homologues that utilize PP_i, initially identified in the amitochondriate protist *Entamoeba histolytica* (Lipman 1965) have been identified in several bacterial species, the archaea *Thermoproteus tenax,* and plants (see Eqn 1.3). A PP_i-PFK has significant metabolic consequences. Firstly, the reaction catalyzed by PP_i-PFKs is freely reversible ($\Delta G^{\circ\prime} = -8.70$ kJmol^{-1}) and is therefore capable of a glycolytic and gluconeogenic role, abrogating the requirement for a FBPase (FBPase is absent from the PP_i-PFK-exhibiting genomes of *Thermotoga maritima*, *Borrelia burgdorferi* and *Treponema pallidum* [Fraser *et al.* 1997; Subramanian *et al.* 2000]). PP_i-PFKs are not allosterically regulated (with the exception of plant PP_i-PFKs, which has fructose-2,6-bisphosphate (F26bP) as an allosteric effector) therefore the major control point of the glycolytic pathway has been lost.

$$PP_i + F6P \underset{PP_i\text{-PFK}}{\rightleftharpoons} F16bP + P_i \qquad \text{(Eqn.1.3)}$$

$$ATP + F6P \xrightarrow[ATP\text{-PFK}]{} F16bP + ADP \qquad \text{(Eqn. 1.4)}$$

$$\text{sedoheptulose-1,7-bisphosphate} + P_i \rightleftharpoons \text{sedoheptulose-7-phosphate} + PP_i \qquad \text{(Eqn.1.5)}$$

Enzyme kinetics and substrate concentrations become determining factors in activity. Non-allosteric PP_i-PFKs are predominantly found in pathogenic anaerobic bacteria and primitive eukaryotes, which do not exhibit the TCA cycle or oxidative phosphorylation, therefore they rely upon glycolysis for ATP production. A PP_i-PFK utilizing PP_i as a high-energy phosphoryl donor conserves ATP, effectively increasing the ATP yield of glycolysis by 50%. This however is of negligible consequence if oxidative phosphorylation is functional. Why then do plants exhibit a PP_i-PFK? The

amitochondriate protist, *E. histolytica* lacks the first two enzymes in the oxidative pentose phosphate pathway and transaldolase. In their stead pentoses are formed by a C2 + C3 condensation reaction, which is catalyzed by a transketolase, aldolase, and PP_i-PFK, where PP_i-PFK catalyses the interconversion of sedoheptulose-1,7-bisphosphate and sedoheptulose-7-phosphate (Mertens 1993 [Eqn.1.5]). It is therefore hypothesized that PP_i-PFK function in plants may not be glycolytic, rather functioning in gluconeogenesis and/or the pentose phosphate pathway (Mertens, E. 1993).

1.4 Kinases

1.4.1 Phosphoryl Transfer

Phosphofructokinase is a kinase, a ubiquitous group of enzymes that catalyze phosphoryl transfer from a *phosphoryl donor* to an *acceptor.* The kinase group is enormously diverse in sequence, structure, and specificity indicative of the multiple roles these enzymes play in cellular pathways. However, the overall theme of the reaction is the same. The reaction involves binding and orientation of a phosphoryl donor; binding and orientation of a phosphoryl acceptor, a magnesium divalent cation (Mg^{2+}) for catalytic function (often a secondary divalent or monovalent cation is also required), and a catalytic mechanism of phosphoryl transfer, which generally involves a conformational change preventing wasteful hydrolysis (Cheek *et al.* 2002; Voet *et al.* 1999). In general three mechanisms exist: transition state stabilization, acid-base catalysis, and ping-pong. A common component of all mechanisms is nucleophilic attack upon the γ-phosphoryl group of ATP complexed to the divalent cation, Mg^{2+}, by the phosphoryl acceptor substrate or an intermediate (see 1.4.4 Ping-Pong). Formation of a Mg^{2+}·ATP complex is essential as the Mg^{2+} coordinates the γ- and β-phosphates, serving to correctly orient the γ-phosphoryl group and to partially ionize the phosphoanhydride bond, yielding the γ-phosphate more susceptible to nucleophilic attack. Phosphoryl transfer proceeds via an associative (S_{N2}), dissociative (S_{N1}), or a hybrid mechanism. The identity of a specific mechanism can be deduced from the transition state structure, which differs in both charge distribution and bond length for purely associative and dissociative mechanisms (Schlichting, I., Reinstein, J. 1997).

1.4.2 Dissociative-S_{N1}

A purely dissociative (S_{N1}) mechanism of phosphoryl transfer involves breakage of the P_γ–O bond before any significant covalent bonding contribution from the nucleophile. It therefore proceeds via a trigonal planar metaphosphate intermediate. Non-enzymatic phosphoryl transfer reactions are typically of a dissociative nature due to the unfavorable electrostatic interactions of an associative mechanism, between the nucleophile and the γ-phosphoryl non-bridging oxygen atoms. Within an enzymatic environment an associative mechanism is possible due to positively charged groups and/or coordinated metal ions stabilizing the negatively charged pentavalent phosphorane intermediate (Cook, A. *et al.* 2002).

Based on structural evidence (Cook, A. *et al.* 2002) the lone pair of electrons from the nucleophile is oriented in-line with the γ–phosphate and γ–β bridging oxygen atom. A consequence and defining characteristic of an S_{N1} mechanism is charge development on the γ–β bridging oxygen upon bond breakage (see Figure1.3a), which is stabilized by coordination to the Mg^{2+} ion, and a decrease in charge on the metaphosphate transition state relative to its ground state. Subsequent to bond breakage the van der Waals interaction between the nucleophile and the trigonal planar metaphosphate intermediate progressively gains in covalent nature. Phosphoryl transfer is rarely a purely associative or dissociative mechanism, rather a fractional combination of both. The associative/dissociative component of a catalyzed phosphoryl transfer reaction can be calculated using the Pauling formula (Mildvan 1997) (see Eqn. 1.6).

$$D(n) = D(1) - 0.60\log(n) \qquad \text{(Eqn. 1.6)}$$

Equation 1.6: Pauling formula, where D(n) is the bond distance between the nucleophile and the Pγ, D(1) is the Pγ -O bond distance (1.73Å), and n is the fractional bond number.

1.4.3 Associative - S_{N2}

A purely associative (S_{N2}) mechanism involves partial bond formation, between the γ-phosphoryl group and the nucleophile concomitant with partial bond cleavage of the γ-β phosphoanhydride bond (see Figure 1.3b). Phosphoryl transfer proceeds via a

pentacovalent phosphorane intermediate. Distinct from a dissociative mechanism, charge development is localized to the equatorial oxygen atoms of the pentacovalent intermediate in an S_{N2} mechanism (see Figure 1.3b).

A number of kinase structures have been resolved complexed with their products and a metaphosphate analogue, which putatively represents the transition-state, allowing structural definition of associative and dissociative mechanisms (see Phospho-CDK2/Cyclin A (1GY3), Cook *et al.* 2002 [dissociative]; UMP kinase (3UKD), Schlichting and Reinstein 1997[associative]).

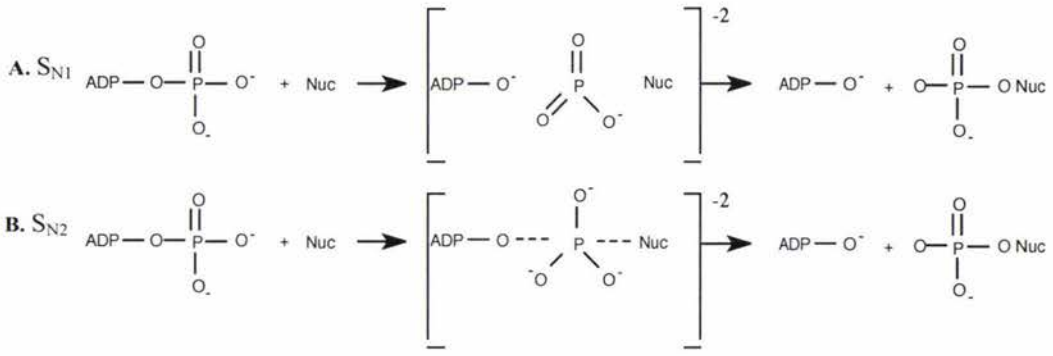


Figure 1.3: Schematic representation of dissociative and associative mechanisms of phosphoryl transfer. **A.** An dissociative mechanism proceeds via an trigonal planar metaphosphate intermediate (S_{N1}). **B.** An associative mechanism proceeds via a pentavalent intermediate (S_{N2}).

$$\text{Enz} + \text{Q-A} \rightleftharpoons \text{Enz-A.Q} \rightleftharpoons \text{Enz-A} + \text{Q}$$

$$\text{Enz-A} + \text{P} \rightleftharpoons \text{Enz.P-A} \rightleftharpoons \text{Enz.P-A}$$

(Eqn.1.7)

Equation 1.7: A ping-pong mechanism involves the release of a product (Q) before all substrates (Q-A and P) are bound.

1.4.4 Ping-Pong mechanism

A ping-pong mechanism (see Eqn. 1.7) involves the release of a product before all substrates are bound. Nucleoside-diphosphate kinase (NDP kinase) exhibits a bi-substrate

ping-pong mechanism (Xu *et al.* 1997). The γ-phosphoryl group of ATP is transferred to an active site histidine (His122) to form a phosphoryl-histidine intermediate. Once ADP is released nucleoside diphosphate is bound and the phosphoryl group transferred. Nucleoside-diphosphate kinase is returned to its original state with the release of the nucleoside triphosphate.

1.4.5 General acid-base catalysis

General acid-base catalysis involves the catalytic transfer of a proton. General acid catalysis involves lowering the free energy of a reaction's transition state via the transfer of a proton from an acid. General base catalysis involves the deprotonation of a functional group rendering it a stronger nucleophile. The mechanism may primarily be general acid or general base catalysis or a concerted mechanism (Voet et al. 1999). For example a general base accepts the proton from a hydroxyl oxygen atom concomitant with its nucleophilic attack of the γ-phosphoryl group of ATP. A general acid subsequently protonates the leaving group. There are several variations on this theme, including a proton shuttle such as the imidazole group of a histidine residue concomitantly accepting a proton from the nucleophilic hydroxyl and donating a proton to the leaving group (Cho and Cook 1988).

1.4.6 Transition state stabilization

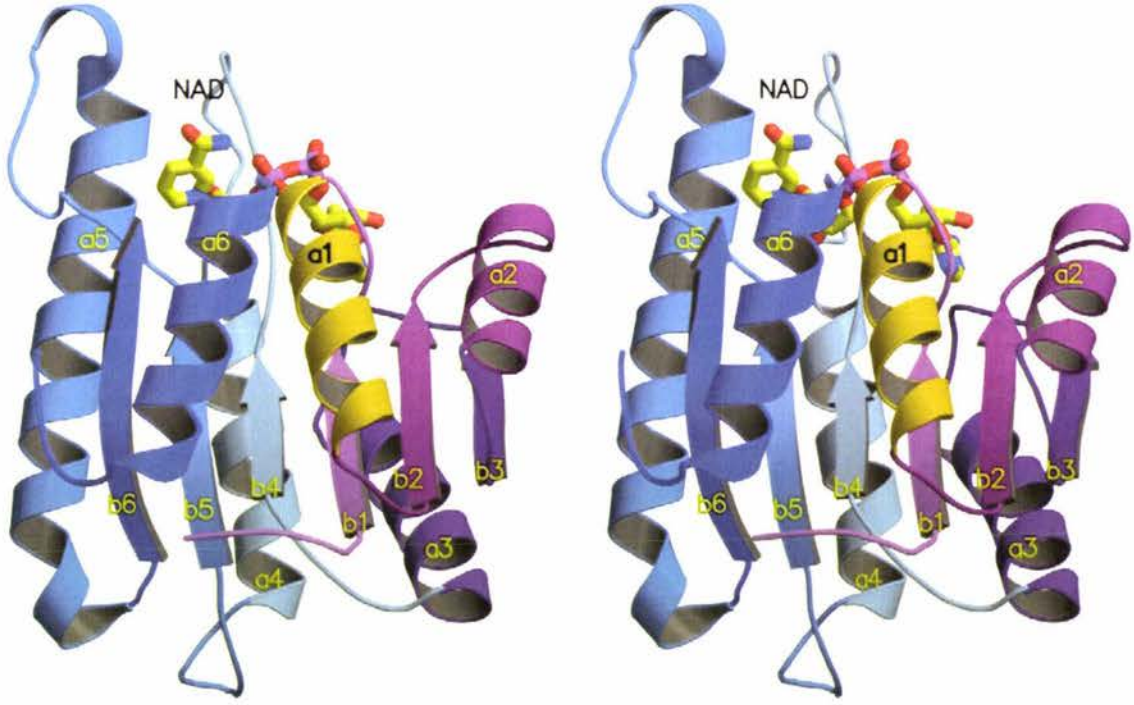
An enzyme with a transition state stabilization mechanism binds the transition state with a greater affinity than either the substrate(s) or product(s). Substrates are bound with a geometry similar to that of the transition state, which induces steric strain. Correct orientation of the substrate/s and steric strain promote the reaction.

1.4.7 Kinase Structural Folds

Cheek *et al* (2002) recently carried out a survey of over 17000 kinase amino acid sequences, classifying them into 30 families based upon sequence similarity. Correlating this information with known three-dimensional structures and secondary structure predictions the authors were able to further classify 19 of these families (comprising 98% of the sequences) into 7 groups, which share a general structural-fold. Group 2 is

comprised of eight families, which share a Rossmann-like fold structure; a core three layered α/β/α domain of 8 βα repeats. The orientation of the bound nucleotide is conserved, bound at the C-terminus of the β-sheet with the triphosphate tail located under the N-terminus of a α-helix (see Figure 1.4). Families are differentiated based upon variation within the β-sheet (e.g. strand order/anti-parallel/parallel), large insertions, or extra domains generally associated with phosphoryl-acceptor substrate binding (Cheek *et al.* 2002). Alcohol dehydrogenase (E.C 1.1.1.1) exhibits an archetypal Rossmann fold (see Figure 1.4). ATP-PFK's belong to the phosphofructokinase-like family of Group 2. Two domains with a Rossman-like fold define the family with the active site situated at the domain interface (see Figure 1.4 and Figure 1.5) (Cheek et al. 2002; Shirakihara and Evans 1988).

A.



B.

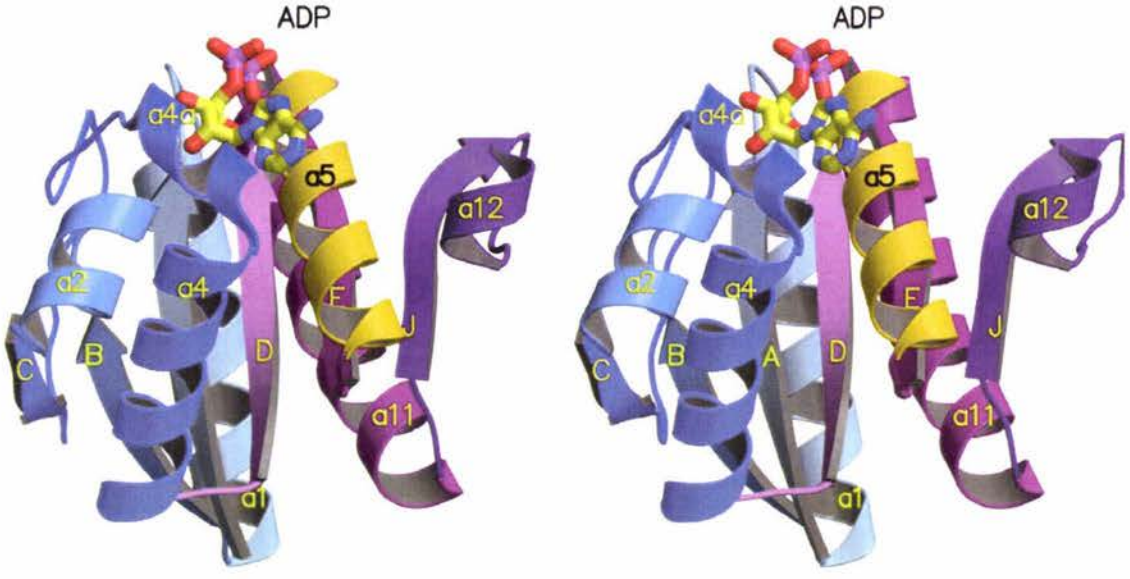


Figure 1.4: Rossmann-like folds exhibited by Group II family members (Cheek *et al.* 2002).

A. Ribbon diagram of the archetypal Rossmann fold (βα repeat) exhibited by alcohol dehydrogenase from *Drosophila lebanonensis* (Benach *et al.* 1998). NAD (carbon = yellow, nitrogen = blue, oxygen = red, phosphorous = purple) is bound at the C-terminus of the β-sheet, the phosphates bound at the N-terminus of α1 (yellow). Only secondary structure elements comprising the Rossmann fold are shown for clarity. **B**. Ribbon diagram of the Rossmann-like core of the *E. coli* ATP-PFK (1PFK) N-domain (Shirakihara and Evans (1988). Note the βα repeat. ADP is bound at the C-terminus of the β-sheet, the phosphate moieties bound at the N-terminus of α5 (yellow). Note the similarity to **A**. Figures prepared with MOLSCRIPT (Kraulis1991) and Raster3D (Merritt and Bacon 1997).

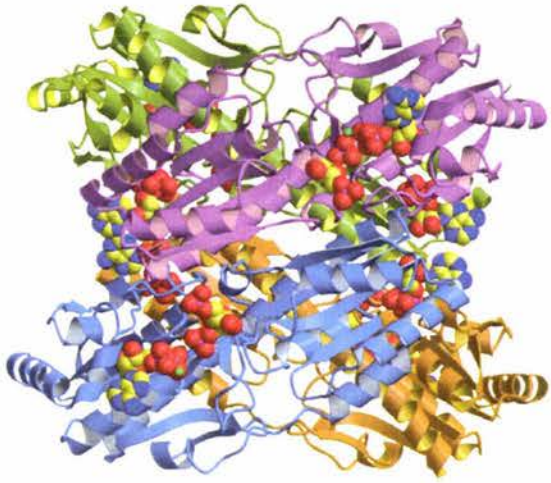
1.5 Structure: Prokaryotic ATP-PFKs

The homotetrameric quaternary structures of *E. coli* and *B. stearothermophilus* ATP-PFK exist as a dimer of rigid dimers (A:B/C:D) (see Figure 1.5a). The subunit of each bacterial homotetramer consists of 319 and 320 residues with molecular masses of 35,000 and 33,900 Da, respectively. Their subunit tertiary structure is comprised of a large, N-terminal and small, C-terminal domain. The main chain traverses the two domains three times with the C-domain situated mid-sequence (see Figure 1.5 B and C and Figure 1.11). This is somewhat unusual, as domains are typically sequential (Shirakihara and Evans 1988; Evans and Hudson 1979). The N-domain of *E. coli* ATP-PFK is comprised of residues 1-131, which form strands A-E and helices 1-5 and residues 251-301, which form strands J and K and helices 10-12. The strands C and K of the N-terminal domain β-sheet are antiparallel with respect to the other five. The C-domain is comprised of residues 132-250, which form four parallel strands F-I and helices 7-9 and residues 302-319, which comprise helix 13 (Shirakihara and Evans 1988) (see Figure 1.5b and c).

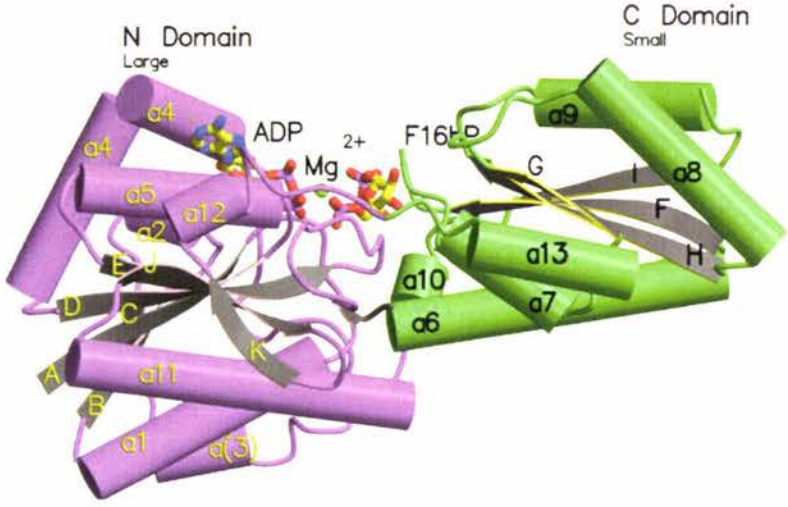
The four subunits of the homotetramer are approximately related by three orthogonal dyad axes (222); p, q, and r as defined by Shirakihara and Evans (1988) (see Figure 1.5a). Their relationship is approximate because the A and B subunits exhibit different conformations. Each subunit contacts two other subunits. Non-interacting subunits are separated by a solvent-filled cavity formed parallel to the q-dyad axis, spanning the length of the macromolecule, approximately 7 Å in diameter (Shirakihara and Evans 1988). Crystals of the *E. coli* ATP-PFK R-state structure (1PFK) complexed with F16bP and ADP belonged to the primitive orthorhombic $P2_12_12$ space group with two subunits (A and B) in the asymmetric unit. The molecular dyad relating the A:B dimer to the C:D dimer of the intact tetramer is coincident with the crystallographic two-fold axis parallel to q.

The subunit A active site is 'closed' relative to the 'open' active site of subunit B . Closure of the subunit A active site is partly a result of a shearing displacement of the upper helical layer (helices 2, 4a, 4, and 5 [see Figure 1.5d]) of the N domain, relative to the β-sheet, towards the active site. Note these helices comprise most of the ATP binding site.

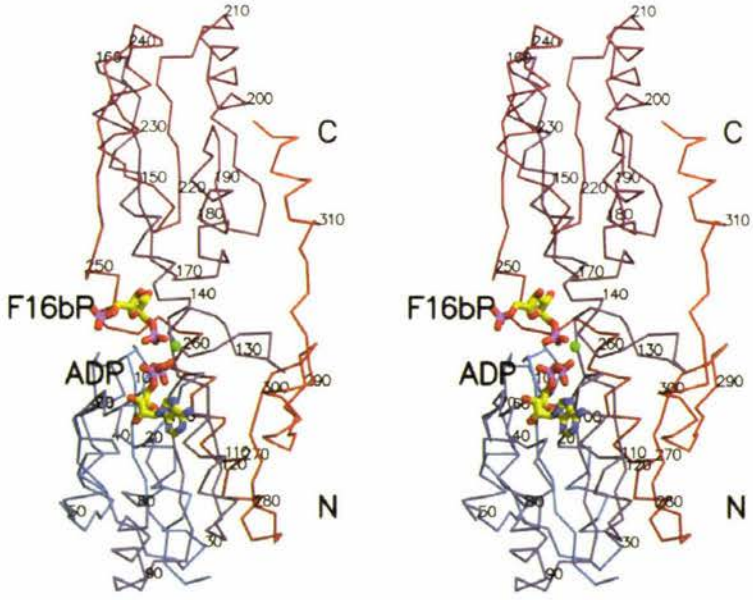
A.



B.


N Domain
Large
C Domain
Small
ADP
Mg
2+
F16bP
α4
α5
α12
α2
α11
α1
α(3)
α9
α8
α13
α10
α7
α6
G
I
F
H
A
B
C
D
K

C.


F16bP
ADP
C
N
F16bP
ADP
C
N

D.

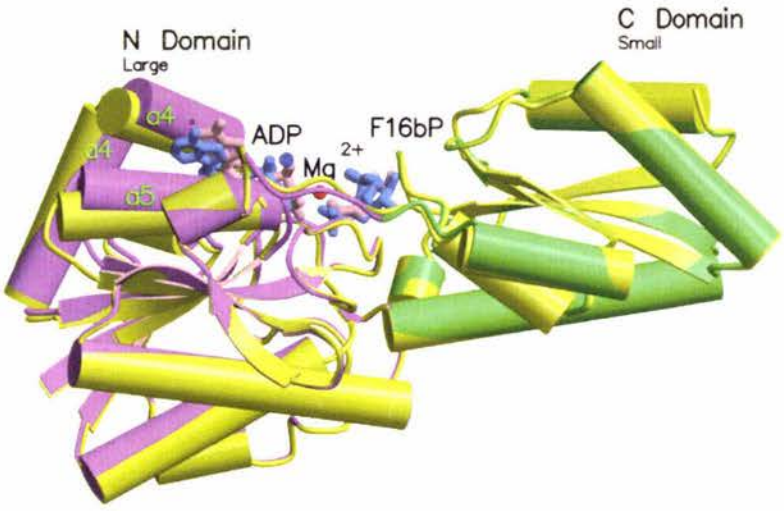


Figure 1.5: Structure of the tetrameric prokaryotic ATP-PFK of *E. coli* (Shirakihara and Evans 1988).

A. Ribbon diagram illustrating the tetrameric quaternary structure of *E. coli* ATP-PFK (subunit A = purple, subunit B = green, subunit C = orange, subunit D = blue). F16bP and ADP are bound at the active site and ADP at the allosteric site (carbon = yellow, oxygen = red, nitrogen = blue, phosphorous = purple, magnesium = green). The orthogonal dyad axis, p, is vertical, parallel to the page, through the center of the tetramer. The orthogonal dyad axis, q, is horizontal, parallel to the page, through the center of the tetramer. The orthogonal dyad axis, r is perpendicular to the plane of the page through the center of the molecule. **B**. Tertiary structure of *E. coli* ATP-PFK (1PFK [Shirakihara, Y. and Evans, P. 1988]) complexed with ADP (bound by the N domain), F16BP (bound by the C domain), and Mg^{2+} (ligands coloured as in **A**.). The N-domain is shown in purple, the C-domain in green. Secondary structural elements are labeled. **C**. Cα trace of *E. coli* ATP-PFK illustrating the domain structure. Note the main chain traverses between domains three times. **D.** Superposition of subunits A (purple, green, and pink) and D (yellow and blue) of *E. coli* ATP-PFK, illustrating their asymmetry. Helices 4a, 4, and 5, which form the ATP-binding site, of subunit A are displaced relative to the β-sheet, closing the active site and concomitantly altering the orientation of ADP to F16bP. Figures prepared with MOLSCRIPT (Kraulis1991) and Raster3D (Merritt and Bacon 1997).

1.6 Active site: Substrate binding and Catalysis

Referring to the *E. coli* ATP-PFK structure resolved by Shirakihara and Evans (1988) the A:D dimer forms two autonomous active sites (see Figure 1.6a). ADP is bound predominantly by the large domain and involves two conserved motifs, GGD**G** and P**G**TIDND comprising residues 101-104 and 123-129, respectively (see Figure 1.6b). The consensus motifs are GGDx and PxTIDzD, where x is a highly conserved residue specific to the phosphoryl donor and z is weakly conserved. Gly104 and Gly124 are synonymous with ATP phosphoryl donor specificity. PP_i-PFKs exhibit highly conserved GGD**D** and P**K**TIDzD motifs (Siebers *et al.* 1998; Xu *et al.* 1994; Chi *et al.* 2000). F16bP is bound by the conserved motifs MGR and LG**H**IQ**R** comprising residues 169-171 and 247-252, respectively and two conserved Arg residues; Arg162 and Arg243, both from the opposing subunit (see Figure 1.6d). The MGR motif is exclusively conserved in all PFKs (Wang *et al.* 1998). The LG**H**IQ**R** motif is not highly conserved and its identity is dependent upon phosphoryl donor specificity (Deng *et al.* 2000). His249 of *E. coli* ATP-PFK is predominantly conserved in ATP-PFKs, whereas PP_i-PFKs exhibit a conserved Tyr residue. Arg162 and Arg243 of *E. coli* ATP-PFK are not conserved in PP_i-PFKs.

The phosphoryl transfer reaction of *E. coli* ATP-PFK proceeds via an associative mechanism (S_{N2} [see Figure 1.3a]). Asp127 acts as a general base, removing the proton from the F6P O1 hydroxyl group, rendering it a stronger nucleophile (Asp127 also functions in destabilization of the product complex [Hellinga and Evans 1987]). Arg72 and Arg171 putatively function in stabilization of the pentavalent transition state (Hellinga and Evans 1987; Berger and Evans 1990). The proton from the protonated general base (Asp127) is subsequently transferred indirectly to the bridging oxygen atom of the β-phosphate group via a passive mechanism.

1.6.1 F16bP binding site

The F6P-binding site involves two conserved motifs, MGR and LG**H**IQ**R**. The MGR motif is situated upon the F-α7 loop and the LG**H**IQ**R** motif upon α10 of the C-domain. His249, Arg252, and Arg162 (βF) and Arg243 (βI) from the opposing subunit bind the F16bP 6-phosphate. The MGR motif is involved in binding the F16bP fructose moiety. Met169 Sδ forms a hydrogen bond with the O2 hydroxyl group and the peptide nitrogen

atom of Gly170 forms a hydrogen bond with O3 hydroxyl group. The fructose moiety is effectively 'sandwiched' between Met169 and Arg171 (see Figure 1.6d). Arg171 functions in transition state stabilization (Hellinga and Evans 1987). Mutation of Met169 to an Ala (Met169Ala) results in a decrease in k_{cat}, leading Wang *et al.* (1998) to erroneously conclude that it directly functions in transition state stabilization. Met169 likely functions in the correct coordination of the MGR motif and of F6P into the active site. Based upon analysis of prokaryotic ATP-PFK structures (Shirakihara, Y. and Evans, P. 1988; Schirmer, T and Evans, P. 1990) the observed k_{cat} effect of the Met169Ala mutation is resultant from disordering of the F-α7 loop, causing misorientation of Arg171 and of F6P bound at the active site.

1.6.2 ADP-binding site

The shearing of helices 4a, 4, and 5, which comprise the Mg^{2+}ADP binding site, relative to the β-sheet in subunit A positions the ADP β-phosphate 1.58 Å closer to the 1-phosphate of F16bP, relative to subunit B (see Figure 1.5d) (Shirakihara, Y., Evans, P. 1988). Binding of the adenine and ribose moieties at the active site of both subunits is relatively equivalent. Major deviations between the active sites of subunits A and B include coordination of the Mg^{2+} cation and α- and β-phosphates due to the closer proximity of F16bP in subunit A. Arg77A and Phe76A of the 3_{10} helix 4a, Met107A of helix 5 and Tyr41A of helix 2 form a hydrophobic cleft into which the adenine group binds (see Figure 1.6b). The adenine group forms indirect hydrogen bonds with Arg111A of helix 5 and Arg77A of 3_{10} helix 4a. In subunit B Arg77B forms the only, indirect interaction. The ribose hydroxyl groups form hydrogen bonds with the peptide oxygen and amino nitrogen of Phe73. In the 'closed' active site of subunit A the α-phosphate forms a salt-bridge with Arg72A and indirectly interacts with Arg171A and the Mg^{2+} cation. The α-phosphates bridging oxygen forms a hydrogen bond with R72A. The β-phosphate is bound at the base of the GGDG motif, at the N-terminus of helix 5, forming hydrogen bonds with Ser105A, D103A, and the peptide amino-groups of Gly102A, Asp103A, Gly104A, and Ser105A (see Figure 1.6c). The β-phosphates negative charge is stabilized via coordination to the Mg^{2+} cation and the helical dipole of α5 (Shirakihara and Evans 1988).

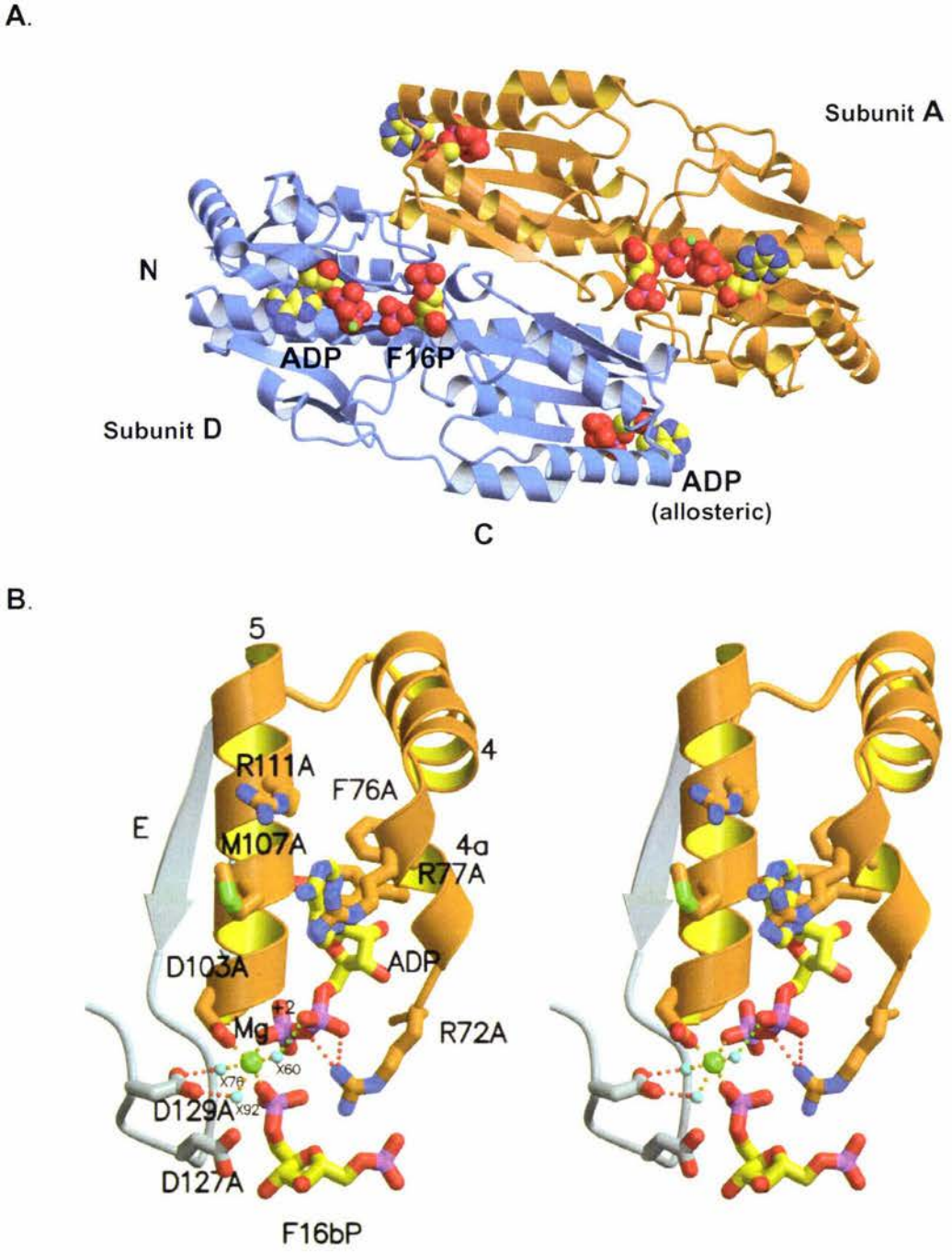


Figure 1.6: Active site of the *E. coli* ATP-PFK (1PFK [Shirakihara and Evans 1988]).

A. Ribbon diagram of subunits A (orange) and D (blue) of the *E. coli* ATP-PFK. The A:D dimer forms two functional active sites. ADP and F6P are shown bound at the active site (carbon = yellow, nitrogen = blue, phosphorous = purple, oxygen = red, magnesium = green). **B.** Stereo view of ADP bound at the ATP-binding site of subunit A of the *E. coli* ATP-PFK (1PFK) structure. Bonds contributing to the octahedral geometry of the Mg^{+2} are shown in orange, the hydrogen bond between the coordinated water (X60) and the α-phosphate is shown in green.

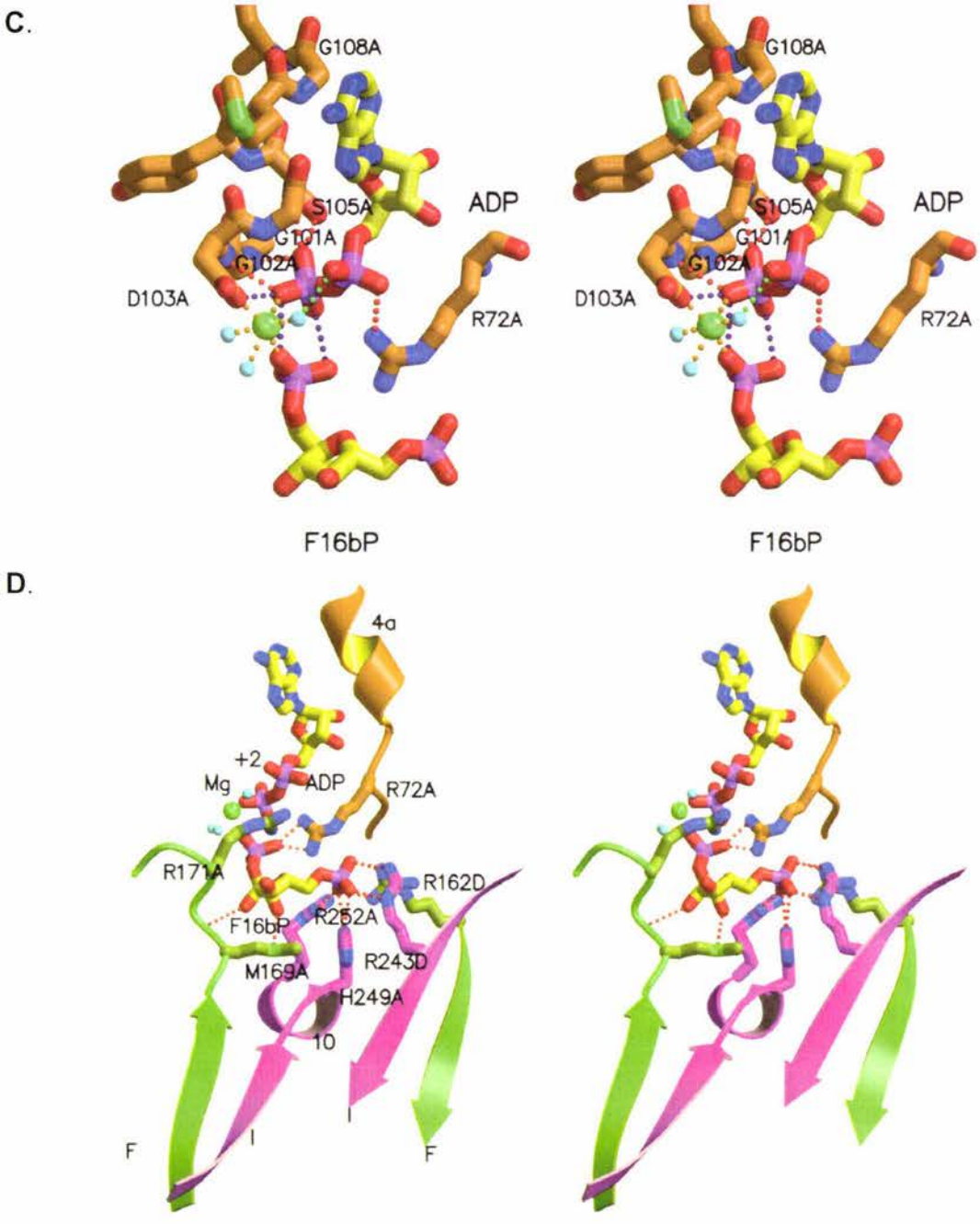


Figure 1.6 continued: Active site of the *E. coli* ATP-PFK (1PFK [Shirakihara and Evans 1988]).

C. Ball-and-stick diagram of the contacts formed by the ADP α- and β-phosphates bound at the N-terminus of helix α5. Note that bonds shown in purple are indicative of protonated ADP α-phosphate oxygen atoms.

D. Stereo view of the F16bP bound at the F6P binding site of subunit A of the *E. coli* ATP-PFK (1PFK) structure. The MGR motif is shown in green and the LGHIQR motif in purple. R162D and R243D from the opposing subunit are shown. Figure prepared with MOLSCRIPT (Kraulis 1991) and Raster3D (Merritt and Bacon 1997).

In the 'open' active site of subunit B the ATP α-phosphate forms a salt-bridge with Arg72B and is directly coordinated to the Mg^{2+} cation. The β-phosphate forms equivalent interactions with the exception of interaction with the 1-phosphate of F16bP and the peptide nitrogen of Asp103 (Shirakihara and Evans 1988).

1.6.3 Mg^{2+}-binding

In the 'closed' active site of subunit A the Mg^{2+} cation bridges the β- and 1-phosphates of ADP and F16bP, respectively. Its octahedral geometry is satisfied by the 1-phosphate of F16bP, the β-phosphate of ADP, Asp103 (GG**D**G), and three water molecules, which are coordinated by Asp129 of the PGTIDN**D** motif and the α-phosphate of ADP (see Figure 1.6b) (Shirakihara and Evans 1988). Coordination of the Mg^{2+} cation in the 'open' active site of subunit B is distinct, coordinated to the α- and β-phosphates. Its octahedral coordination is unclear, coordination to the α- and β-phosphates of ADP and a single water molecule are the only highly ordered interactions observed (Shirakihara and Evans 1988).

1.7 Allosteric site

The allosteric regulation of prokaryotic ATP-PFKs has traditionally been explained in terms of the *Monod-Wyman-Changeux* concerted two-state model for the allosteric regulation of an enzyme (Schirmer and Evans 1990). This model postulates that an enzyme exists in two forms in equilibrium, an active *relaxed* (R) state and an inactive *tense* (T) state. Implicit in the two-state model is cooperative binding of substrate (F6P) and resultant sigmoidal enzyme kinetics. Binding of an allosteric activator induces quaternary and/or tertiary changes resulting in a transition from the T-state to the R-state. An allosteric inhibitor elicits the opposite, R-state to T-state transition (Monod et al. 1965). *E. coli* and *B. stearothermophilus* ATP-PFKs exhibit sigmoidal kinetics for F6P but not ATP (F6P binding of *B. stearothermophilus* ATP-PFK is only cooperative in the presence of PEP). The reversible transition between a high- (R) and low-affinity (T) state for F6P, has been proposed to account for the allosteric and cooperative behavior of

bacterial ATP-PFKs (Schirmer and Evans 1990). Allosteric activators (ADP) and inhibitors (PEP) supposedly elicit their affect by inducing transition between the R- and T-states, respectively. ATP has erroneously been characterised as an allosteric inhibitor (Lau and Fersht 1989). ATP inhibition is not allosteric rather inhibition results from substrate antagonism. Binding of the first substrate lowers affinity for the second (see Figure 1.7). The random, non-rapid equilibrium mechanism prevents substrate equilibration. Wang and Kemp (2001) propose the following scheme in which substrate binding is rate limiting.

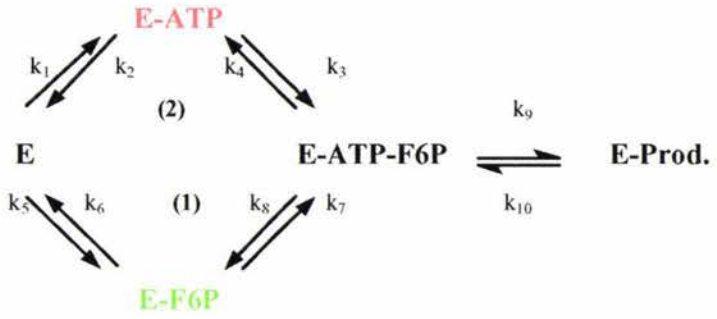


Figure 1.7: Proposed kinetic scheme for substrate binding and catalysis for *E. coli* ATP-PFK. The reaction pathway through (1) is proposed to be faster than via (2) (Wang and Kemp 2001).

Inhibition is a result of the product of the rate constants for route **(1)** and **(2)** being disparate and the high catalytic turnover (k_9) preventing substrate equilibration (Zheng and Kemp 1992).

Arg162 and Arg243 along with Arg72 have been implicated by site directed mutagenesis studies of *E. coli* ATP-PFK in cooperativity (Berger and Evans 1990). Mutation of either Arg162, Arg243, or Arg72 to serine increased $K_{M[F6P]}$ by 165-, 53-, and 3-fold and reduced the degree of cooperativity from 4 to 2.1, 2.7, and 2.2, respectively with minimal effect on k_{cat} or $K_{M[ATP]}$ (Evans and Hudson 1979). Bacterial ATP-PFKs are allosterically activated by Mg^{2+}·ADP, which is observed, bound at the allosteric site of the R-state crystal structures of *E. coli* (1PFK) and *B. stearothermophilus* (4PFK) ATP-PFK. The deep cleft constituting the allosteric site is formed at the A:B subunit interface, between

opposing N- and C-terminal domains (Evans and Hudson 1979; Shirakihara and Evans 1988) (see Figure 1.8a). The allosteric site is formed by loop 7-G and loop 8-H of the C domain and pseudo-helix 3, the C-terminus, and side chains from helices 1 and 6 from the N-terminal domain (see Figure 1.8a) (Shirakihara and Evans 1988). Four allosteric sites are formed by the prokaryotic ATP-PFK tetramer (see Figure 1.5a). The adenine group is bound at the surface of a deep cleft, sandwiched between and forming van der Waals contacts with the planar peptide group of Gly212A of loop 8-H and the aromatic ring of Tyr55B of helix (3), respectively. The only hydrogen-bonding interaction is indirect, via a water molecule to Agr54B. The ribose group forms hydrogen bonds with Lys214A and Asp59B. The α- and β-phosphates are bound at the base of the cleft by four Arg residues (21B, 25B, 54B, and 154A) and are coordinated to a Mg^{2+} cation. The Mg^{2+} cation exhibits octahedral geometry, coordinated to the α- and β-phosphates, Gly185B O, Glu187B O, and two highly ordered water molecules (see Figure 1.8b) (Shirakihara and Evans 1988). All protein atoms involved in direct coordination of the Mg^{2+} cation are contributed by the same subunit (Shirakihara and Evans 1988). The allosteric site of the *B. stearothermophilus* ATP-PFK R-state structure (4PFK) is highly similar to that of *E. coli* ATP-PFK (1PFK). Equivalent residues form equivalent bonding interactions with Mg^{2+}·ADP, albeit for a few minor exceptions. Arg54 is not conserved in *B. stearothermophilus* ATP-PFK, rather Arg211 fulfills its function; forming direct bonding interactions with the adenine group, the C-terminus, and α-phosphate. Arg211, as a functional homologue of Arg54 accommodates the one-residue C-terminal extension of *B. stearothermophilus* (Evans, P., Hudson, P. 1979).

The T-state structure of *B. stearothermophilus* ATP-PFK (6PFK) complexed with the non-physiological PEP analogue, 2-phosphoglycolate (2-PGA) has been resolved to 2.5 Å resolution (Schirmer and Evans 1990). The 2-phosphoglycolate-phosphate group is bound at the base of the deep allosteric cleft by Arg154B of helix 6, Arg21A and Arg25A of helix 1, and the peptide amino nitrogen of Asp59A of loop (3)-4a. Position and binding interactions are homologous to binding of the β-phosphate of ADP in the R-state structures of both *E. coli* and *B. stearothermophilus* ATP-PFKs (see Figure 1.8c). The 2-phosphoglycolate-carboxyl group is bound by Arg211B of helix 8, Arg25A of helix 1,

and the peptide amino nitrogen of Lys214B of loop 8-H at the surface of the cleft (see Figure 1.8c) (Schirmer and Evans 1990). Structural superposition of the allosteric site with the corresponding *B. stearothermophilus* ATP-PFK R-state structure (4PFK) reveals that the overall structure of the allosteric site is homologous albeit for loops 8-H and 7-G, which narrow the allosteric site. It seems apparent that the smaller ligand induces and facilitates the narrowed cleft. Glu187B of loop 7-G, involved in the direct ligation and coordination of two water molecules to the Mg^{2+} cation alters its conformation significantly with 2-phosphoglycolate bound. Rotated around χ_1 by approximately 140° away from the active site it no longer forms a hydrogen bond with I217B, though it retains its salt-bridge with Lys213B (see Figure 1.8b and c) (Schirmer and Evans 1990).

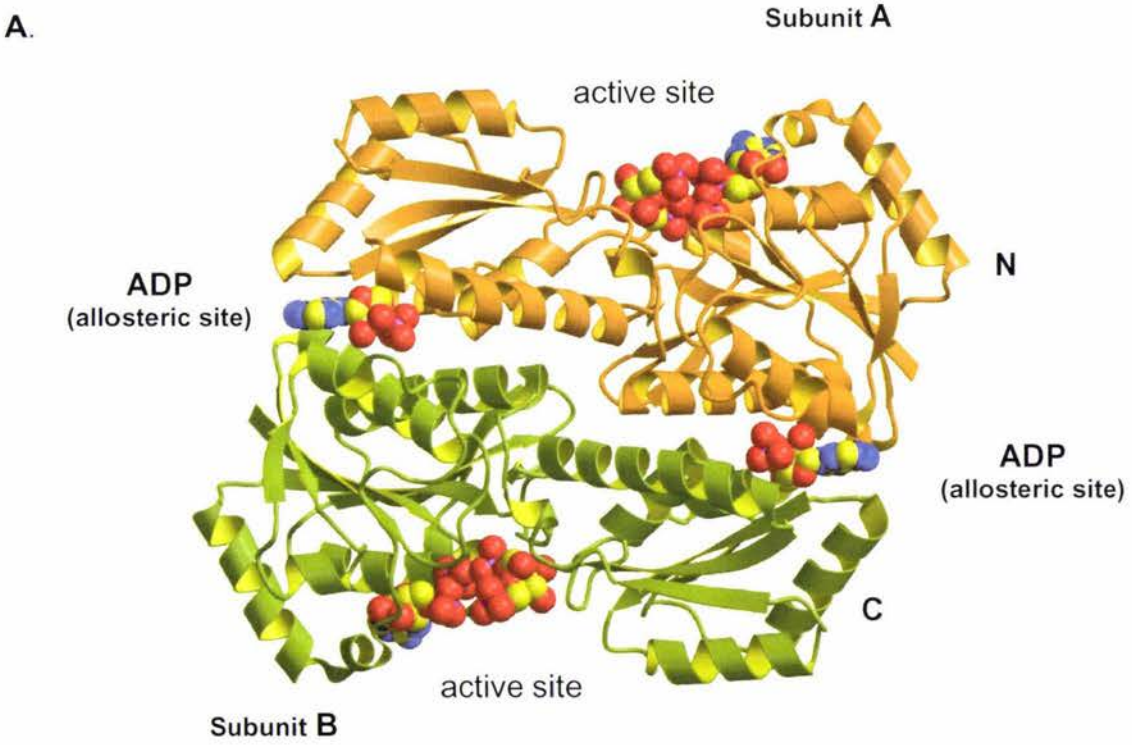


Figure 4.8: The allosteric binding site of prokaryotic ATP-PFKs.

A. Ribbon diagram of the *E. coli* ATP-PFK A:B dimer. Two allosteric sites are formed at the interface of subunits A (orange) and B (green). ADP is shown bound at the allosteric sites along with ADP and F16bP at the active sites cation (carbon = yellow, nitrogen = blue, oxygen = red, phosphorous = purple, magnesium = green).

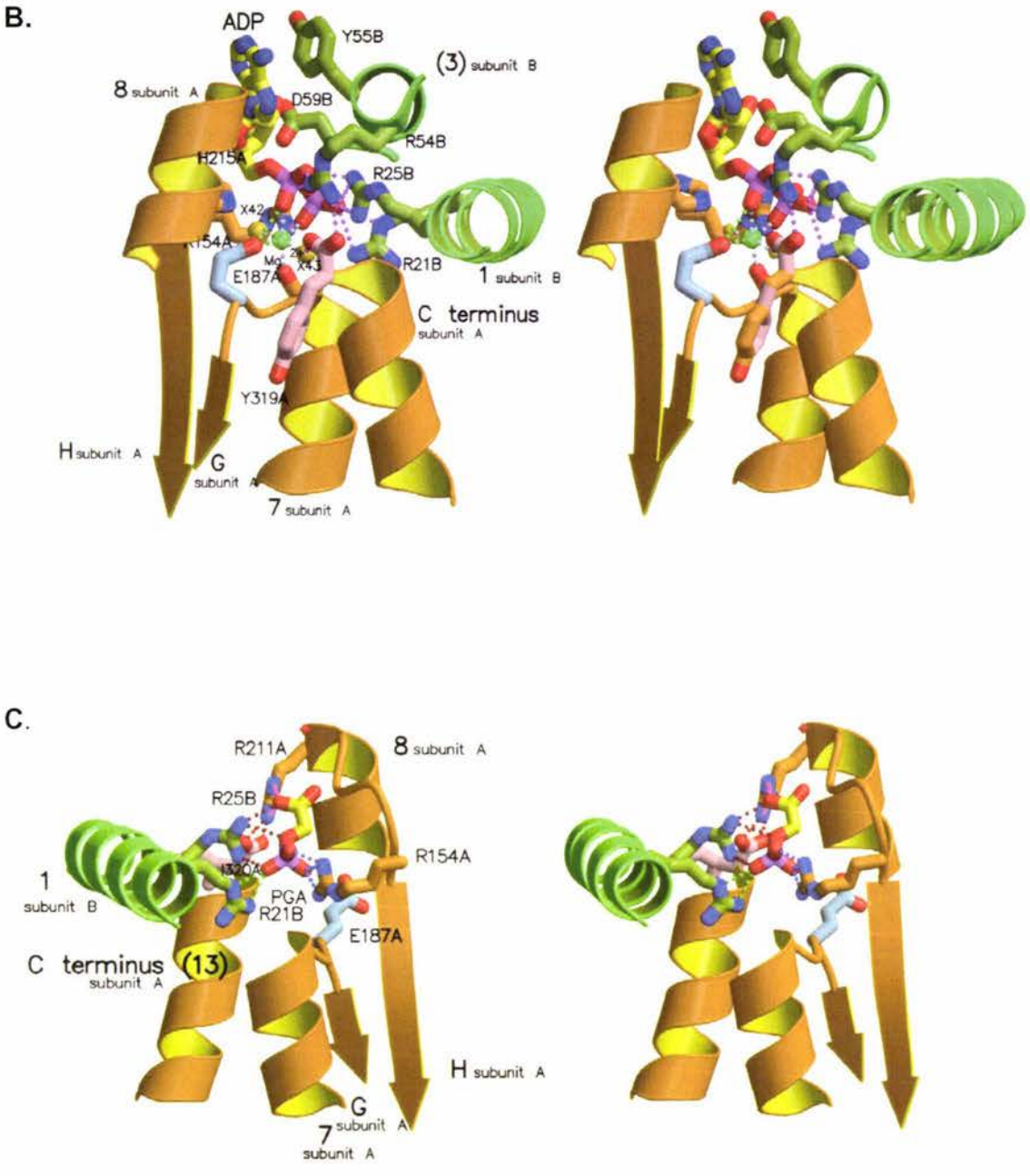


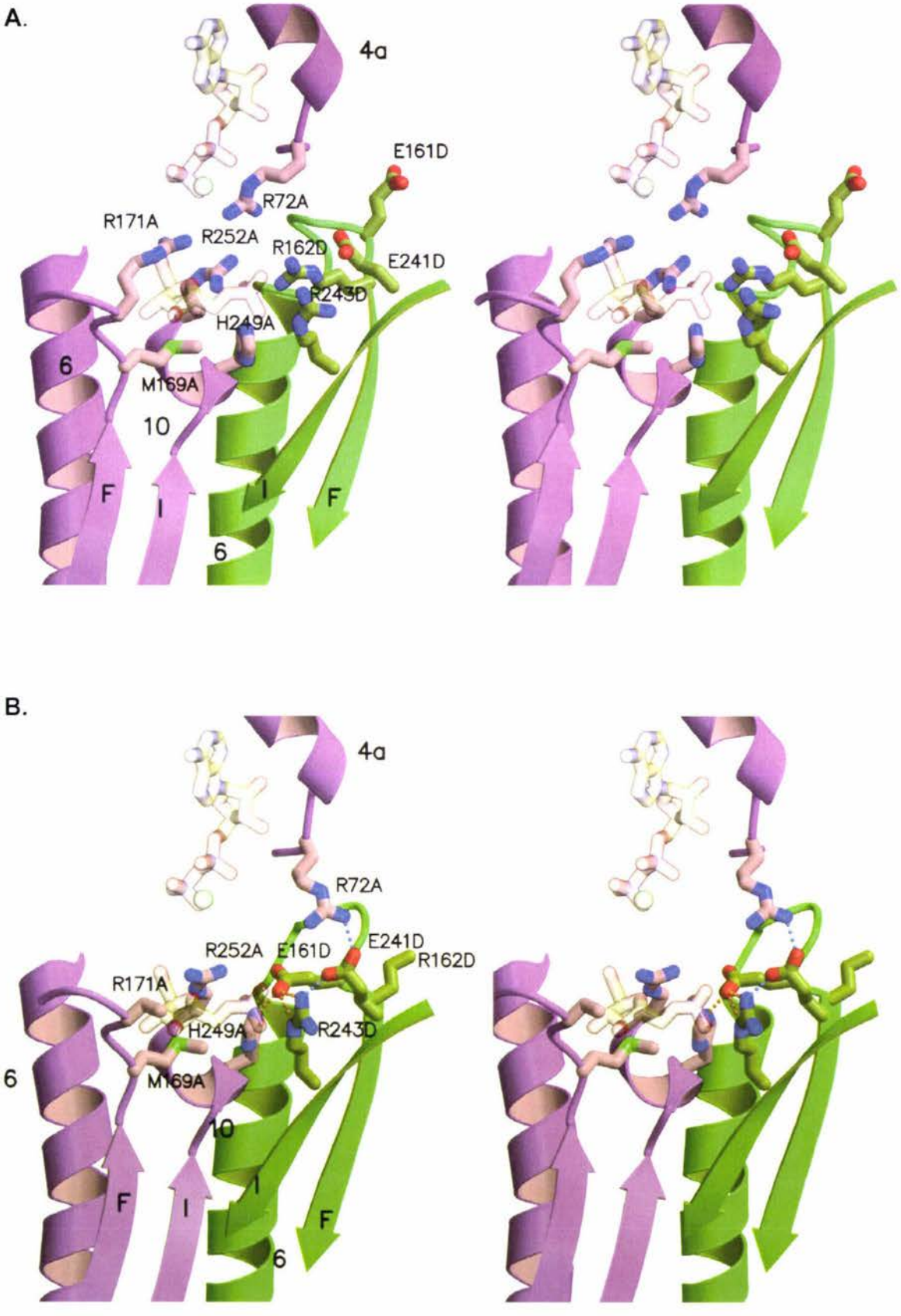
Figure 1.8 continued: The allosteric binding site of prokaryotic ATP-PFKs.

B.Stereo view of ADP bound at the allosteric site (A:B) of *E. coli* ATP-PFK (1PFK). Note that Glu187A is rotated into the active site, away from β-strand H, and directly coordinates the Mg^{2+}.

C. Stereo view of PGA (a PEP analogue) bound at the allosteric site (A:B) of *B. stearothermophilus* (6PFK). The view is rotated 180° relative to **A.** and coloured identically for clarity. Note Glu187A is rotated out of the allosteric site, towards β-strand H. Figures prepared with MOLSCRIPT (Kraulis1991) and Raster3D (Merritt and Bacon 1997).

1.7.1 R-state ⇌ T-State transition

Transition from the R- to T-state in *B. stearothermophilus* involves a conformational change and quaternary structure rearrangement; a rotation of the A:B dimer relative to the C:D dimer by approximately 7° about the p-dyad axis, across the small interface, resulting in a conformational change at the active site (formed at the A:D dimer interface). Conformational change at the A:B dimer interface involves expulsion of the water layer between symmetry-related β-strands I, resulting in symmetry-related direct antiparallel hydrogen bonding between Val246A and Val246D and hydrogen bonding between the side chain hydroxyl group of Thr245 and the peptide carbonyl oxygen of Leu247. The resultant coupled quaternary and tertiary change reorients Arg252A and Glu222A within the individual subunit and Arg243D from the opposing subunit of the A:D dimer. In concert these residues act to narrow the F6P binding site, presenting a steric barrier to F6P binding (Schirmer and Evans 1990). Concomitant with expulsion of the water layer is the melting of the N-terminus of helix 6 and 6-F loop and reorganization of Arg162D and Glu161D (see Figure 1.9b and c). In the R-state structure helix 6 ends with a 3_{10} hydrogen bond between T158D and E161D. E161D goes on to make a left-handed helical turn (dihedral angles of $\phi = 45°$, $\psi = 63°$) which orients the side chain Arg162D into the active site of subunit A (see Figure 1.9a). The peptide carbonyl oxygen of Arg162D forms the first hydrogen bond between the parallel β-sheets F and H and its peptide amino group forms a C-cap hydrogen bond with Ala157D of helix 6. The guanidinium group of Arg162D makes two contacts with F6P, and contacts with Arg243D and Glu241D. Glu161D is pointing away from the active site, towards the allosteric site of subunit B. Upon transition to the T-sate the N-terminus of helix 6 melts, terminating with a 3_{10} hydrogen bond between the peptide carbonyl oxygen of Lys152D and the peptide amino group of Asp155D. Glu161D of the now extended 6-F loop adopts β-sheet-like dihedral angles ($\phi = -104°$, $\psi = 85°$), which concomitantly reorients Arg162D and Glu161D (see Figure 1.9b). Arg162D is rotated away from the F6P binding site, towards helix 8. It is highly disordered in the T-state structure (6PFK), only visible in subunit C. Arg162C forms contacts with the peptide carbonyl oxygen's of Lys213C and Lys214C and the imidazole nitrogen of His209C. Glu161D is rotated into the active site of subunit A, adopting the former position of Arg162D, subsequently forming a salt-


A.
4a
E161D
R72A
R171A
R252A
R162D
E241D
R243D
H249A
6
M169A
10
F
I
1
F
6
B.
4a
R72A
R252A
E161D
E241D
R162D
R171A
H249A
R243D
6
M169A
10
F
I
1
F
6

C.

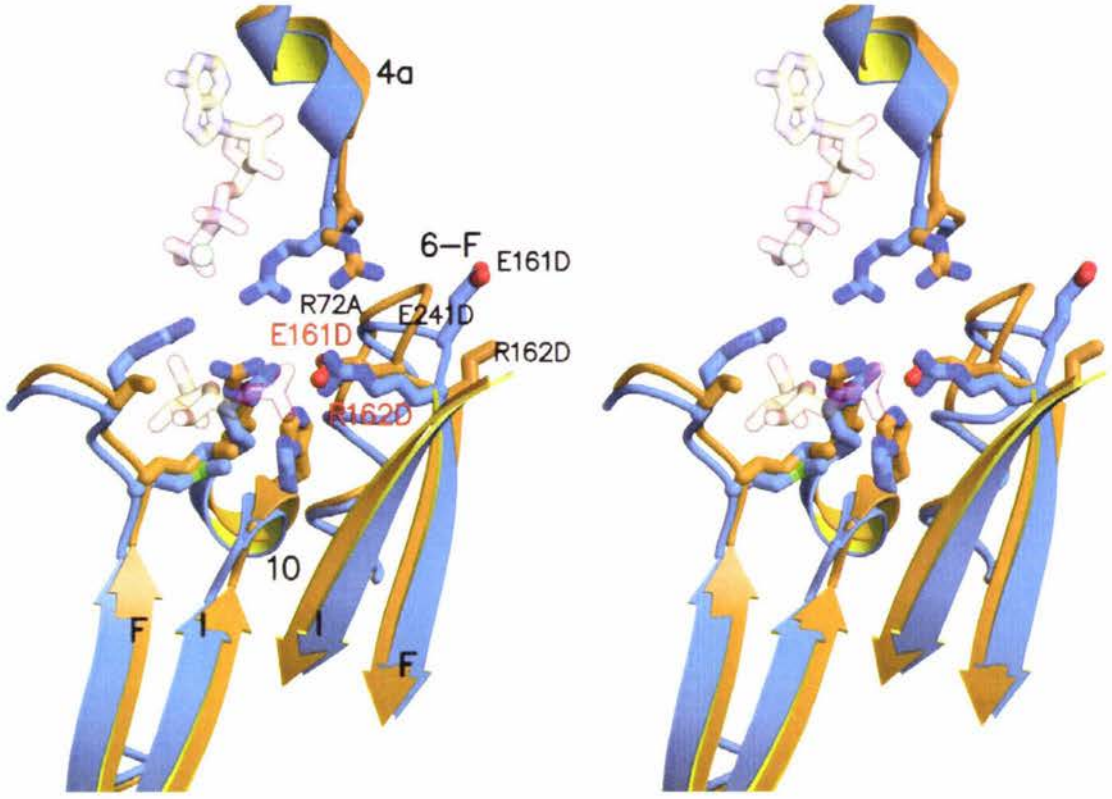


Figure 1.9: The putative R- to T-state transition of *B. stearothermophilus* ATP-PFK.

A. Stereo view of the R-state F6P binding site at the A:D dimer interface of *B. stearothermophilus* ATP-PFK (4PFK) (subunit A = purple, subunit D = green). ADP, Mg^{+2}, and F6P (transparent) are shown bound at the active site (carbon = yellow, nitrogen = blue, oxygen = red, phosphorous = purple, magnesium = green) B. Stereo view of the F6P-binding site of the *B. stearothermophilus* ATP-PFK (6PFK) T-state structure (coloring consistent with A.). ADP, Mg^{+2}, and F6P (transparent) bound at the active site of the *B. stearothermophilus* (4PFK) R-state are superimposed. Note that E241D is now forming salt bridging contacts with R72A and R243D and that E161D is reorientated, into the F6P 6-phosphate binding site. In the T-state structure (6PFK) the side chain of R171A and R162D is disordered. **C.** Stereo view of superposition of the F6P-binding site of the R-state (4PFK [blue]) and T-state (6PFK [orange]) of *B. stearopthermophilus* ATP-PFK structures illustrating the conformational change of E161D and R162D involved in the R- to T-state transition. Note the conformational change of the 6-F loop and the concomitant reorientation of E161D and R162D. Note also that in the T-state the carboxylate group of E161D superimposes with the R-state R162D guanidinium group. Arg243D and E241D have been omitted for clarity.

bridge with Arg243A (see Figure 1.9b). The orientation of Arg72A is also altered in the R- to T-state transition. In the R-state it bridges the two substrates. Upon transition to the T-state it forms a salt bridge with Glu241D, the conformation of which is slightly altered (see Figure 1.9b) (Schirmer and Evans 1990). Based upon structural comparisons of R- and T-state structures it seems evident that the reorientation of Arg162D, Arg243D and Arg72A diminishes their ability to bind F6P, which concomitant with the introduction of a negatively charged residue, Glu161D, which would in theory elicit an electrostatic repulsion with an incoming F6P molecule at the active site, accounts for the decreased affinity of the T-state for F6P (Evans and Hudson 1979) (Schirmer and Evans 1990).

1.8 Caveat: Failure of the Concerted Two-state Model

Explanation of the allosteric regulation of prokaryotic ATP-PFKs within the bounds of the *Monod-Wyman-Changeux* concerted two-state model and the Schirmer and Evans (1990) *B. stearothermophilus* ATP-PFK model has been comprehensively dismantled based upon lack of structural information, induced hypercooperativity and site directed mutagenesis studies (Kimmel and Reinhart 1999; Deville-Bonne *et al.* 1994; Auzat. *et al.* 1995; Auzat. *et al.* 1994; Johnson and Reinhart 1994; Zheng and Kemp 1994; Wang and Kemp 2001; Pham *et al.* 2001).

1.9 Evolution of the Phosphofructokinase A Family

Phosphofructokinases can be differentiated into two monophyletic groups based upon sequence similarity, the PFK A and PFK B families. The ADP-dependent phosphofructokinases were thought to comprise a third, PFK C family until structural and sequence similarity to the PFK B family was established (Ito *et al.* 2001). The PFK A family is distributed almost exclusively within bacteria and eukaryotes. This family is comprised of the higher eukaryotic ATP-dependent PFKs, bacterial ATP- and PP_i-dependent PFKs, the eukaryotic PP_i-dependent PFKs, and the archaeal PP_i-PFK from *T. tenax*. The PFK B family is more widespread, comprised of a broad range of kinases with varying substrate specificities, which include the ADP-dependent PFKs, the non-allosteric ATP-PFK from *A. pernix*, the highly characterized PFK2 from *E. coli*,

adenosine kinases, ribokinases, fructokinases, and tagatokinases (Ronimus and Morgan 2001). The generally accepted view is that the two families evolved independently (Kotlarz and Buc 1982). This is supported by the fact that the PFK A and PFK B families are structurally unrelated.

1.9.1 PFK A Family

The PFK A family forms a monophyletic cluster. The derived clades can be designated into three groups, group I, II, and III based upon sequence similarity (Siebers *et al.*1998) (see Figure 1.10). The common ancestor of the PFK A family is considered to have arisen in a thermophilic, anaerobic environment (Ronimus and Morgan 2001; Yamagata *et al.* 1991). The extinct ancestor is proposed to have arisen prior to the divergence of eukarya, bacteria, and archaea due to all lineage's harboring a PFK A PFK with a Rossmann-like architecture and conserved sequence motifs involved in substrate specificity and catalysis. The identities of underlined positions within these motifs (see Table 1.1) are a fingerprint of a PFKs phosphoryl specificity and a generalized indication of the phylogenetic group (based upon alignments constructed by this author), exceptions exist however.

A.

Motif	Residue	Phosphoryl Specificty	Group(s)
P(**K/G**)TID(N/G)D	**K**	PP_i/ATP	II and III
	G	ATP	I
GG(D/E)(**G/D**)	**D**	PP_i	all
	G	ATP	all
(**Y/H**)XXRX_7D	**Y**	PP_i/ATP	II
	H	PP_i/ATP	I and III

B.

Motif Combination	Residues	Phosphoryl Specificty	Group(s)
P**K**TID(N/G/A)D	**K**	PP_i	II
YXXRX_7D	**Y**		
P**K**TID(N/G/A)D	**K**	PP_i/ATP	III
HXXRX7D	**H**		
P**G**TID(N/G/A)D	**G**	ATP	I
HXXRX7D	**H**		

Table 1.1. Sequence motifs common to all PFK A PFKs. A primary structure fingerprint of phosphoryl donor specificity (Based upon a sequence alignment constructed by this author).

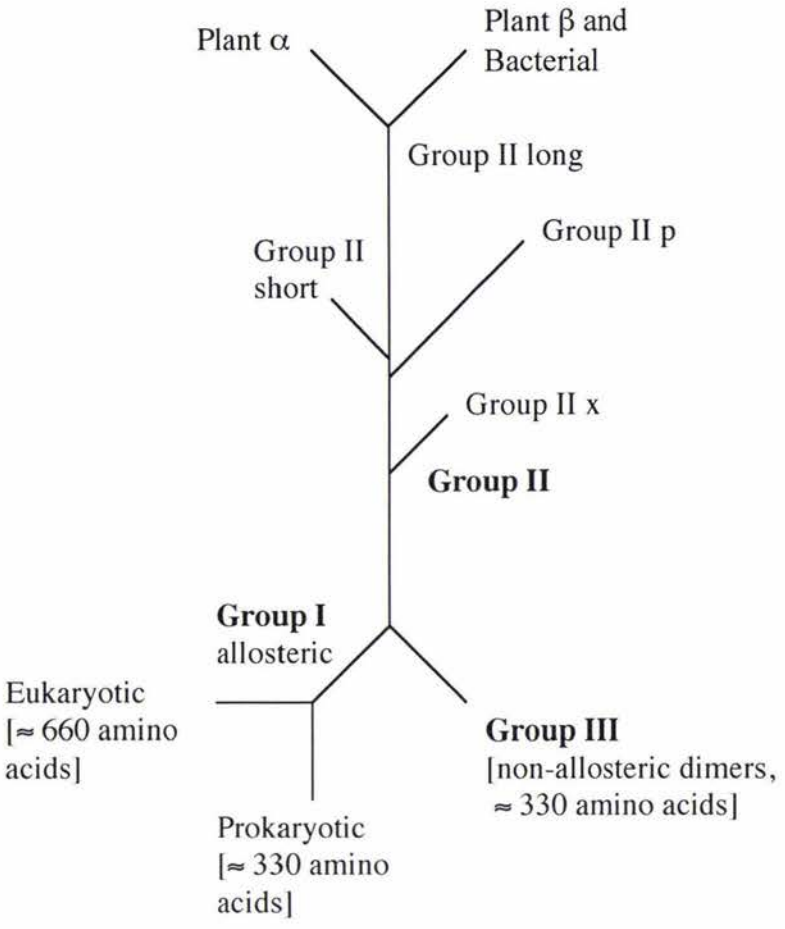


Figure 1. 10: Schematic unrooted phylogenetic tree of the PFK A family depicting the three groups and their putative relationship. The diagram is based upon phylogenetic relationships suggested by Muller *et al.* 2001, Mertens *et al.*1998, Siebers *et al.* 1988, and Poorman *et al.* 1984. The diagram is purely schematic and is intended to demonstrate the relative relationship of the three groups only: no inferences can be drawn from the branch lengths.

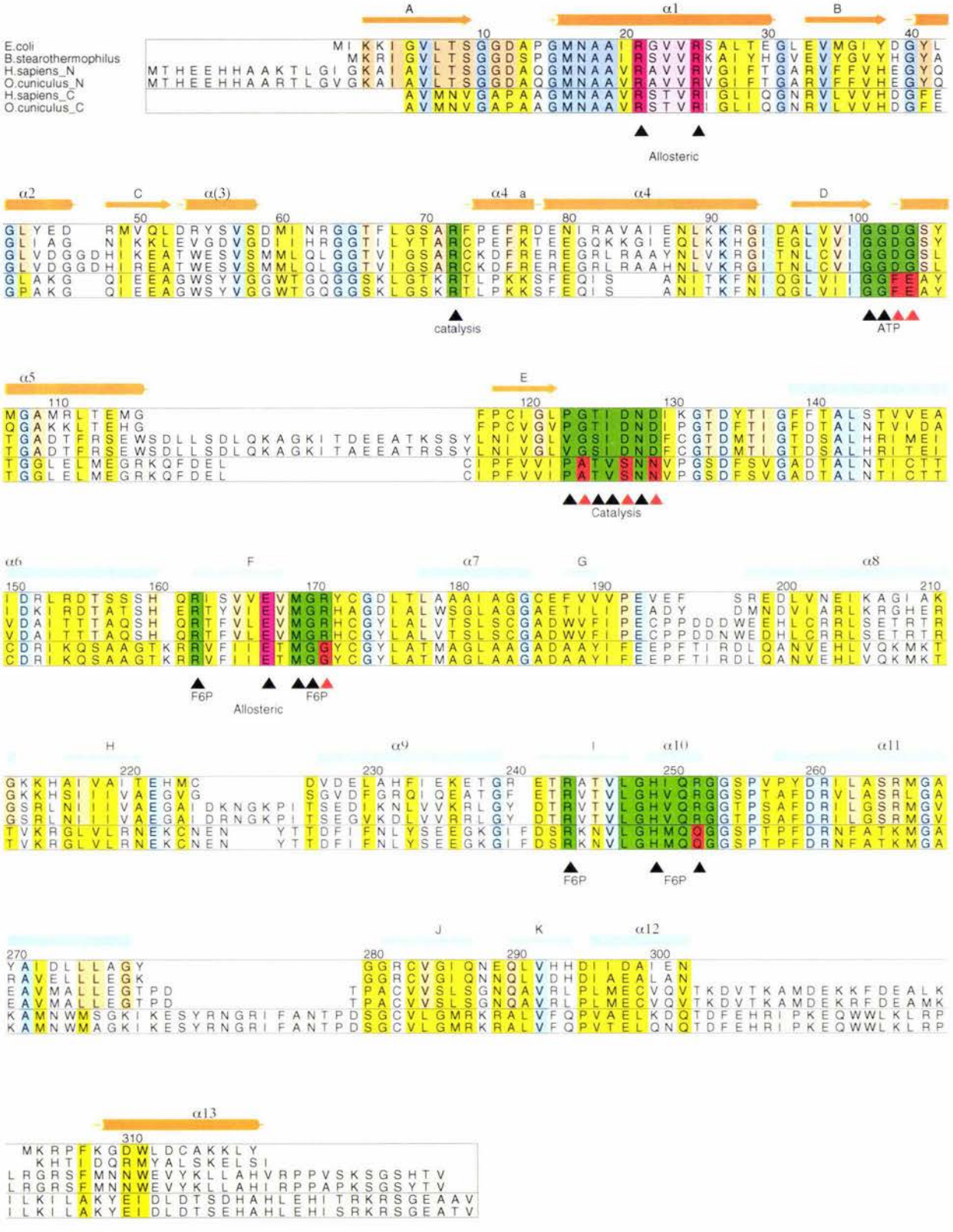
The ancestral phosphoryl donor specificity is open to debate, however the evolutionary history of the PFK A family involves multiple gene duplications, lateral gene transfers, and phosphoryl donor differentiation events explaining the discord between apparent PFK evolution and organismic relationships as suggested by 16S RNA phylogenetic trees (Woese *et al.* 1990).

1.9.2 Group I ATP-PFKs

Group I family members are exclusively ATP-dependent. The group is split into two groups comprised of prokaryotic and eukaryotic ATP-PFKs, respectively. Prokaryotic and eukaryotic ATP-PFKs are generally allosterically regulated homotetramers with subunit M_r of 33-38 kDa and approximately 82k Da, respectively. Poorman *et al.* (1984) hypothesized and demonstrated that higher eukaryotic ATP-PFKs evolved via gene duplication, tandem fusion, and divergence from the prokaryotic gene. The higher-eukaryotic ATP-PFK subunit represents a conserved, covalently linked prokaryotic ATP-PFK dimer (A:B) where the second PFK motif has evolved to adopt a regulatory function (see Figure 1.11) (Poorman *et al.* 1984). The N-terminus of *Oryctolagus. cuniculus* ATP-PFK shares higher sequence identity with *B. stearothermophilus* ATP-PFK than the C-terminus (42% and 34%, respectively). A closer inspection of the sequence alignment reveals that every sequence motif characteristic to PFK A family members involved in substrate binding and catalysis has a crucial mutation in the duplicated C-terminus, which would render it catalytically inactive (see Figure 1.11). Equilibrium binding studies indicate four active sites within the rabbit muscle PFK homotetramer (Kemp and Krebs 1967). It was therefore concluded that the N-terminus contains the active site, whilst the C-terminal site has diverged to adopt an allosteric or regulatory role. For a detailed discussion of eukaryotic ATP-PFK allosteric regulation see Kemp and Foe (1983), Lau and Fersht (1987), Li *et al.* (1999), Zheng and Kemp (1994), Kemp and Gunasekera (2002), and Heinisch *et al.* (1989).

The regulatory α- and catalytically active β-subunits of the novel yeast $\alpha_4\beta_4$ octameric ATP-PFK (110-120kDa) are predicted to have arisen via a subsequent duplication of the higher-eukaryotic ATP-PFK gene (Heinisch *et al.* 1989).

Figure 1.11: (Over page) Alignment of prokaryotic ATP-PFKs with the N- and C-termini of eukaryotic ATP-PFKs based upon the atomic structures of *E. coli* and *B. stearothermophilus* ATP-PFK demonstrating the evolution of the eukaryotic gene via gene duplication, tandem fusion and divergence. Conserved motifs are shown in green. Mutations of these motifs within the C-terminus are shown in red. Exclusively conserved residues are shown in blue, moderately conserved in yellow. Diagram drawn with ALSCRIPT (Barton 1993).


A
α1
B
E.coli
B.stearothermophilus
H.sapiens_N
O.cuniculus_N
H.sapiens_C
O.cuniculus_C
Allosteric
α2
C
α(3)
α4 a
α4
D
catalysis
ATP
α5
E
Catalysis
α6
F
α7
G
α8
F6P
Allosteric
F6P
H
α9
I
α10
α11
F6P
F6P
J
K
α12
α13

1.9.3 Group II PP_i-PFKs

Group II exhibits PP_i- and ATP-dependent family members. The monophyletic group is comprised of four clades supported by strong bootstrap values (Mertens *et al.* 1998). The first clade, dubbed 'short' (~30-45 kDa) includes the dimeric PP_i-PFK of *Trichomonas vaginalis* (Mertens *et al.* 1989) and the tetrameric PP_i-PFK of *Naegleria fowleri* (Mertens *et al.* 1993). The second clade, dubbed 'long' contains the large-subunit PP_i-PFKs (~60-68 kDa), including the allosterically regulated plant PP_i-PFK ($\alpha_2\beta_2$) α- and β-subunits and the homologous β-subunit PP_i-PFKs exhibited in some amitochondriate protists; *Hexamita inflata*, *Entamoeba histolytica*, and *Giardia lamblia* and the two bacterial groups, the Spirochetes and *Chlamydia* (Moore *et al.* 2002) (see Figure 1.12). The relatively distant taxonomic relationship between plants, the protists, and the two bacterial groups suggests that their exhibition of a homologous plant-like β subunit (β_2) is indicative of lateral gene transfer on multiple independent occasions (Fraser *et al.* 1998). The third clade, dubbed 'x' (~40 kDa) is comprised of ATP-PFKs with the exception of the predicted PP_i-PFK of *Oryza sativa*, which interestingly lacks the signatory PP_i-PFK GGD**D** motif. The group is enigmatic in that it contains secondary PFKs from the amitochondriate protist *E. histolytica* (*pfk1*), and the Spirochetes *Treponema pallidum (pfk1)* and *B. burgdorferi* (*pfk1*), which also exhibit a group II non-allosteric, β_2 PP_i-PFK. Whilst the *pfk1* gene product from *B. burgdorferi* and *T. pallidum* remain uncharacterized the *pfk1* gene product from *E. histolytica*, previously thought to be inactive has been characterized as an ATP-PFK (Chi *et al.* 2001). The fourth clade, dubbed 'p' (~40kDa) includes the PP_i-PFK of *Propionibacterium freudenreichii*.

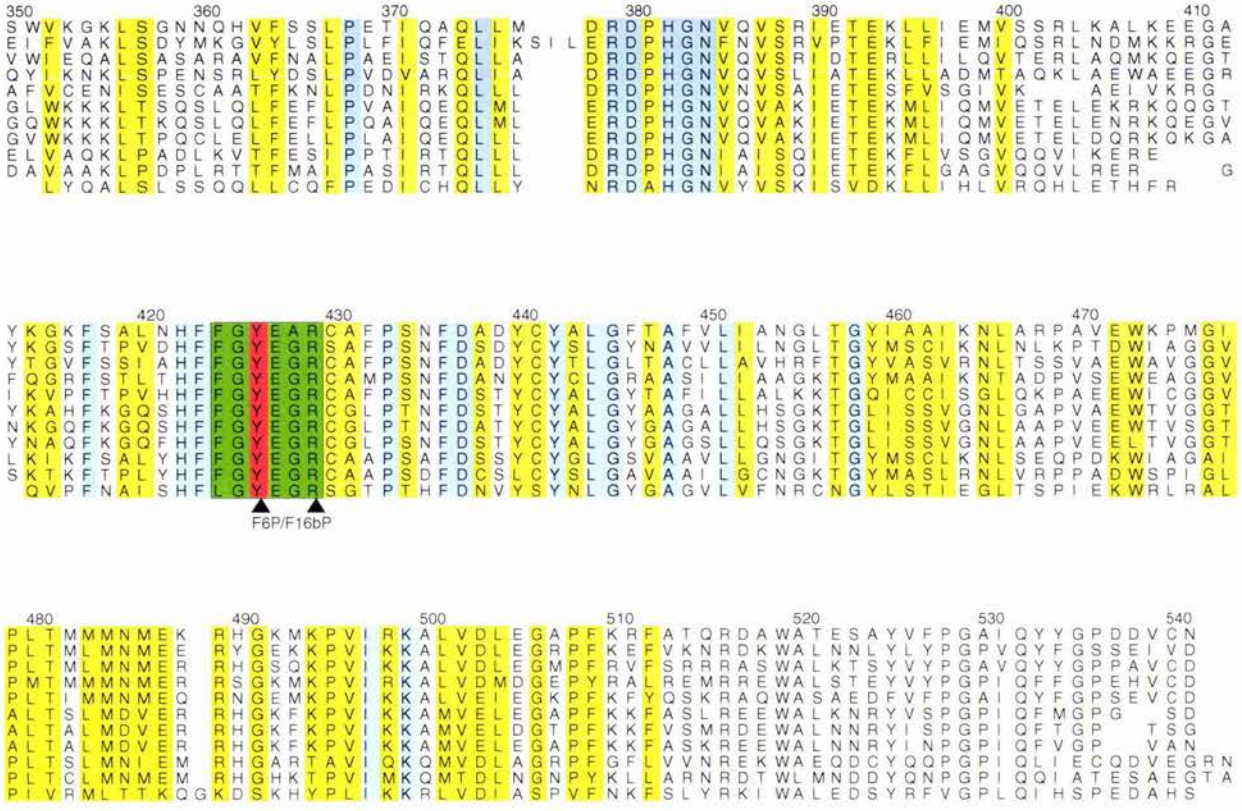
Group II and III are proposed to have diverged via a change in phosphoryl *donor* specificity subsequent to divergence of clade 'x'. Following the divergence of clade 'x' the other three group II clades ('long', 'short', and 'p') diverged without further change in phosphoryl *donor* specificity. A subsequent gene duplication within the 'long' clade is thought to have given rise to the plant regulatory (α) and catalytic (β) subunits, the smaller α-subunit presumably binding the allosteric effector, F-2,6-bP (Muller *et al.* 2001). Though group II and III family members are likely to be non-allosteric sequence alignment reveals hypothetical residual allosteric sites within certain members of group III and the group II's 'x' clade PP_i- and ATP-PFKs, resembling the allosteric sites of

group I prokaryotic and eukaryotic ATP-PFKs (see Figure 1.13). This corroborates the proposed early divergence of groups I, III, and clade 'x' relative to the 'long', 'short', and 'P' clades of group II (Muller *et al.* 2001; Mertens *et al.*1998; Siebers *et al.* 1988). It is also suggestive of the early evolution of allosteric regulation within the PFK A family. The uncharacterized *pfk1* gene products from the Spirochetes *T. pallidum* and *B. burgdorferi* , which share high sequence similarity with the characterized 48 kDa ATP-PFK of *E. histolytica* (Chi *et al.* 2001) are among sequences exhibiting remnants of an allosteric site.

Figure 1.12: (page 34 - 35) Sequence alignment of group II, long clade family members illustrating that lateral gene transfer is putatively involved in the evolution of the clade. Exclusively conserved residues are shown in blue, broadly conserved residues in yellow and motifs involved in catalysis and/or substrate binding in green. Sequences were aligned using ClustalX and the figure produced with ALSCRIPT (Barton 1993).

Figure 1.13: (page 36) Sequence alignment of group I (*E. coli* ATP-PFK, *H. sapiens* ATP-PFK [N-terminus]), group III (*A. mediterannei* PP_i-PFK, *A. methalonica* PP_i-PFK, *C. perfringens* ATP-PFK, and *M. tuberculosis* ATP-PFK) and group II, x clade (*B. burgdorferi* ATP-PFK? and *T. pallidum* ATP-PFK?) family members illustrating that the group III and group II, x clade family members exhibit at least remnants of an allosteric site. The absence of Arg162 and Glu168 in the *B. burgdorferi* and *T. pallidum* putative ATP-PFKs suggests that the active site is inactive, an artifact of its ancestral PFK. Exclusively conserved residues are shown in blue, broadly conserved residues in yellow, and motifs involved in catalysis and/or substrate binding in green. The allosteric site is shown in purple. Sequences were aligned using Clustalx (Jeanmougin *et al.* 1998) and the figure produced with ALSCRIPT (Barton 1993).


S.thermophilum
B.burgdorferi
T.pallidum_beta_PPi
P.gingivalis_PPi
E.histolytica_PPi
R.communis_beta_PPi
C.paradisi_beta_PPi
S.tuberosum_beta_PPi
H.inflata_PPi
G.lamblia_PPi
C.trachomatis_PPi
Catalysis
PPi
Catalysis
F6P/F16bP


F6P/F16bP


E.coli
A.mediterranei_PPi
A.methanolica_PPi
M.tuberculosis_ATP
C.perfringens_ATP
H.sapiens_N
B.burgdorferi_ATP?
T.pallidum_ATP?
Allosteric
catalysis
ATP
Catalysis
F6P
Allosteric
F6P
F6P
F6P

1.9.4 Group III PFKs

Group III family members have an approximate subunit M_r of 37 kDa and exhibit varying phosphoryl donor specificities. The group is comprised of bacterial PFK sequences with the exception of the archaeal PP_i-PFKs from *Thermoproteus tenax* (Siebers *et al.*1998). The scarcity of archaeal PP_i-PFKs suggest that the *T. tenax* PFK A PFK evolved from a possible lateral gene transfer event. Group III sequences represent a hybrid of higher-eukaryotic PP_i- and ATP-motifs (see Table 1.1). The group III clade exhibits short branch lengths, indicative of early differentiation. Group III PP_i-PFKs exhibit higher sequence similarity with group I and III ATP-PFKs than their functionally related group II counterparts, suggesting differentiation in phosphoryl donor specificity subsequent to divergence of the group.

1.9.5 Evidence for a Latent Nucleotide-Binding Site?

The high conservation of sequence motifs between ATP- and PP_i-PFKs suggest that only subtle differences differentiate phosphoryl donor specificity. Mutation of Gly 105 (GGD**G**) and 124 (P**G**TIDND) to Asp (GGD**D**) and Lys (P**K**TIDND), respectively in *E. coli* ATP-PFK effectively abolishes ATP-dependent activity, no subsequent PP_i-dependent activity is observed (Chi and Kemp 2000). When reciprocal mutations were carried out in the Group II, long clade family member *E. histolytica* PP_i-PFK, Asp 175 (GGD**D**) and Lys 201 (P**K**TIDND) to Gly (GGD**G** and P**G**TIDND) a global shift in phosphoryl donor specificity was observed (Chi and Kemp 2000). Mutations individually resulted in an approximate 10,000-fold reduction in PP_i-dependent k_{cat}, whilst the double mutation totally abrogated PP_i-dependent activity. Concurrently the single D175G mutation yielded a 200-fold increase in k_{cat} for ATP-dependent activity, while the K201G mutation yielded a modest 4-fold increase. Interestingly the double mutant yielded a 100-fold increase in k_{cat}, half that observed for D175G. Defining the enzyme substrate preference as the ratio of k_{cat}/K_M for substrate one to k_{cat}/K_M for substrate two (Chen *et al.* 1995), the D175G mutation changed substrate preference by 10^7!

1.9.6 Summary: Evolution of PFK A PFKs

The PFK A family forms a monophyletic cluster. The derived clades can be designated into three groups, group I, II, and III based upon sequence homology (Siebers, B. *et al.*

1998). ATP-PFKs near universal distribution, in contrast to the sparse distribution of PP_i-PFK throughout the phylogenetic tree is suggestive of an ancestral PFK with ATP specificity. Distribution of PP_i-PFKs can thus be explained by secondary adaptations from an ATP-PFK on multiple independent occasions. Group I family members are exclusively ATP-dependent. The group is split into two clades comprised of prokaryotic and eukaryotic ATP-PFKs. Eukaryotic ATP-PFKs evolved via duplication, tandem fusion, and divergence from the prokaryotic gene (Poorman et al. 1984). Group III exhibits both PP_i- and ATP-PFKs with short branch lengths, indicative of early differentiation of the clade. Group III PP_i-PFKs exhibit higher sequence similarity with group III ATP-PFKs than their functionally related group II counterparts, suggesting differentiation in phosphoryl donor specificity subsequent to divergence of the group. Group II exhibit both PP_i- and ATP-PFKs. The group is split into four clades; short, long, x, and p. The regulatory α-subunit of plant PP_i-PFKs ($\alpha_2\beta_2$) is likely to have evolved via gene duplication and divergence. The exhibition of plant β-like homodimers in the *Spirochetes*, *Chlamydia*, and amitochondriate protists is putatively a result of multiple lateral gene transfer events (Fraser *et al.* 1998).

Evolution of the PFK A family is very intricate and is likely to have involved multiple phosphoryl donor differentiation events, gene duplications, and lateral gene transfer events. Evolutionary divergence of PFKs is likely to be reflective of the varying complexities of glycolytic regulation within eukarya, bacteria, and archaea.

1.10 Aims of this Project

At the commencement of this project one PP_i-PFK structure had been completed. The aim of the proposed research was to determine the three dimensional structure of *S. thermophilum* PP_i-PFK by X-ray crystallography, advancing our understanding of this unique family of enzymes. Solution of the structure with and without substrate bound will give an insight into which residues are actually involved in substrate binding and catalysis, which up unitl now could only be suggested by sequence alignments and site directed mutagenesis studies.

- The role of the Lys (PKTIDZD) and the Asp (GGDD) residues of the conserved motifs has been inferred based upon site directed mutagenesis experiments. Determination of the structure will allow a more definitive explanation.
- Structural determination will allow direct investigation of the latent nucleotide-binding site in PP_i-PFKs (only demonstrated in *E. histolytica* PP_i-PFK) proposed by Chi *et al.* (2000). An Asp/Gly mutation of the GGDD motif in *E. histolytica* PP_i-PFK resulted in a global shift in the enzymes phosphoryl donor specificity, from PP_i to ATP. The reciprocal Gly/Asp mutation of the GGDG motif in *E. coli* ATP-PFK abrogated ATP-dependent activity without a concomitant increase in PP_i-dependent activity. The hypothesis proposed to explain this observation is that secondary structure changes modulate the effects of the primary sequence change hence point mutations are not informative. *E histolytica* contains a Group II, long clade, β_2 PFK whilst *E.coli* exhibits a Group I small subunit PFK, therefore an insertion may be involved in substrate binding and/or determination of phosphoryl donor specificity.
- Structural analysis will allow verification of the catalytic and substrate-binding roles of the conserved, PP_i-dependent sequence motifs involved in F6P/F16bP binding. For example the function of Tyr420 in *E. histolytica* PP_i-PFK (Tyr426 in *S. thermophilum,*) is putatively orthologous to His249 in *E.coli* ATP-PFK (Deng et al. 2000). Structural analysis of His249 in *E.coli* has shown it to form a salt bridge with the 6-phosphate group of Fru-6-P. The positive charge of His249 stabilizes the negatively charged phosphate. Obviously Tyr420 cannot fulfill this function. In

addition the Tyr420Phe mutant of *E. histolytica* PP_i-PFK exhibited higher affinity for F6P than the Tyr420His mutant, which would not be expected if His and Tyr were functionally homologous. From the results for the Tyr420His mutant it can be hypothesized that the aromatic property of Tyr420 is more important than its hydrogen bonding ability. Structural analysis of *S. thermophilum* PP_i-PFK will assign the conserved Tyr residues function in PP_i-PFK.

- Structural determination of the *S. thermophilum* PP_i-PFK will allow a more accurate identification of residues at the catalytic site, distinct from ATP-PFKs to be mutated by site directed mutagenesis to determine how essential they are to substrate binding and/or catalysis. This will fulfill the prerequisite for designing an anti-microbial agent for pathogenic spirochetes such as *B. burgdorferi* (Lyme's disease) and *T. pallidum* (syphilis).

The structural determination of a PP_i-PFK is essential for the advancement of our understanding of their catalytic function, conformational flexibility, enzymatic mechanism, evolutionary relationship, and the development of anti-microbial agents for pathogens that utilize PP_i-PFKs in their glycolytic pathways.

Chapter 2

Materials & Methods

2.1 Introduction

Recombinant *S. thermophilum* PP_i-PFK was cloned, expressed in *E. coli*, and purified by Dr. Ron Ronimus (Thermophile Research unit, Department of Biological Sciences, University of Waikato), as described earlier (Ronimus *et al.* 2001). Initial crystallization screens of *S. thermophilum* PP_i-PFK, which yielded crystals used in the initial diffraction experiment were carried out by Dr. Stanley Moore formerly of Massey University (see section 2.3.1) (unpublished data). *S. thermophilum* PP_i-PFK was transported and stored in a buffered solution of 30 mM Tris HCl, pH 7.5, 1 mM $MgCl_2$, 30% glycerol, 7 mM β-mercaptoethanol, and 50 mM NaCl at -18°C. The protein concentration was approximately 1.25 $mg.ml^{-1}$.

2.1.1 Experimental strategy

The aim of this research was to solve the structure of the *S. thermophilum* PP_i-PFK in its unliganded state and complexed with its reactants (F6P and PP_i) or products (F16bP and P_i). The structures of a number of enzymes complexed with F16bP, bound at the active or allosteric site have been resolved. These include the bifunctional inositol monophosphatase/fructose 1,6-bisphosphatase from the hyperthermophile *Archaeoglobus fulgidus* (Stieglitz *et al.* 2002 [1LBZ]), rabbit muscle fructose 1,6-bisphosphate aldolase (Choi *et al.* 1999 [6ALD]), *E. coli* glycerol kinase (Ormo *et al.* 1998 [1BO5]), human erythrocyte pyruvate kinase (Valentini *et al.* 2002 [1LIU]), *E. coli* ATP-PFK (Shirakihara and Evans 1988 [1PFK]), and pig kidney fructose 1,6-bisphosphatase (Villeret *et al.* 1995 [1FPI]). Crystallization of all the complexes, with the exception of *E. coli* ATP-PFK was achieved either by co-crystallization or soaking the crystal in mother liquor containing F16bP. Preliminary soaking experiments by Dr. Stanley Moore of *B. burgdorferi* PP_i-PFK crystals (unpublished information) resulted in the failure to observe bound substrate. Therefore co-crystallization experiments were employed.

2.2 Materials

The chemicals listed were used throughout the course of this research and were obtained from the following companies:

Hampton Research (California, U.S.A)
Crystal Screen Macromolecular Crystallization Kit, Crystal Screen 2 Macromolecular Crystallization Kit, Polyethylene glycol (PEG) 1000, PEG 4000, PEG 6000, PEG 8000, PEG 10000.

Sigma Chemical Company (St. Louis, U.S.A)
Bis-Tris-Propane, Bis-Tris, D-Fructose 1,6-bisphosphate, D-Fructose 6-phosphate, Ammonium chloride, N,N,N',N'-tetramethylethylenediamine,

United States Biochemical Company (Ohio, U.S.A)
Hepes, Trisw

Ajax Chemicals (Sydney, Austrailia)
Potassium chloride, Ammonium acetate, Sodium chloride, Potassium chloride, Ammonium chloride, Nickel chloride, Ammonium sulfate, Magnesium chloride

Bio-Rad Laboratories (California, U.S.A)
30 % Acrylamide/Bis solution,

BDH Laboratory Supplies (Poole, England)
Ammonium peroxidisulfate, Sodium acetate, *tetra*-Sodium pyrophosphate

Riedel-De Haen Ag (Hannover, Germany)
Manganese chloride,

2.3 Methods

2.3.1 Crystallization of *S. thermophilum* PP_i-PFK

An aliquot of *S. thermophilum* PP_i-PFK was dialyzed from the storage buffer into a buffered solution containing 1 mM $Na_4P_2O_7 10H_20$ (tetra-sodium pyrophosphate), 1 mM $MgCl_2$, 1 mM dithiothreitol and 10 mM Tris HCl, pH 7.5 and concentrated to 10 $mg.ml^{-1}$. The solution was centrifuged at 3000 rpm (Sorvall GS3 rotor) for 10 minutes, prior to laying down crystallization screens, to sediment any particulate matter that may have formed in solution. The solution was stored on ice at all times when not stored at 4°C. Vapor diffusion by hanging drops was the crystallization method employed, entailing the suspension of a 4 μL hanging drop of equal volumes of ~10 $mg.ml^{-1}$ buffered protein solution and reservoir solution on plastic cover-slips over individual wells of a 24 well tissue culture plate containing the respective reservoir solution. Each well was sealed using petroleum jelly. Crystals with a needle-like morphology, which diffracted to 2.2 Å were grown from a reservoir solution of 2 M $(NH_4)_2SO_4$, 0.1 M NaCl, and 0.1 M Hepes, pH 7.5 (this work was carried out by Dr. Stan Moore [unpublished information]).

2.3.2 Determination of *S. thermophilum* PP_i-PFK Concentration

$$\text{Beer's law} = A = \varepsilon cl \qquad \text{(Eqn. 2.1)}$$

The precise concentration of the buffered *S. thermophilum* PP_i-PFK stock solution was determined by UV spectroscopy to be 0.9627 $mg.ml^{-1}$. The concentration was calculated from the absorbance using Beer's law (see Eqn. 2.1). The peak absorbance, measured at 278 nm for a twenty-fold diluted solution was 0.147. The extinction coefficient was calculated to be 50781 $M^{-1}.cm^{-1}$, as determined from the primary sequence based upon the absorbance of Trp, Tyr, and Phe residues at 280 nm using the ProtParam tool (http://kr.expasy.org/tools/protparam.html). The concentrated protein was stored at 4 °C.

2.3.3 SDS-PAGE Gel Electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was employed to verify purity of the enzyme after each equilibration of the *S. thermophilum* PP_i-PFK from the stock buffer. Stock solutions, 12% resolving and 4% stacking gels were freshly prepared. The 12% resolving gel contained 3.2 ml acrylamide stock (Bio-Rad laboratories), 2.0 ml resolving buffer (1.5 M Tris HCl, pH 8.8), 2.68 ml H_2O, 80 µL 10% SDS, 40 µL 10% ammonium peroxodisulphate (APS), and 4 µL N,N,N',N'-tetramethylethylenediamine (TEMED). The acrylamide to bisacrylamide ratio of the acrylamide stock was 37.5:1. The 4% stacking gel contained 0.399 ml acrylamide stock (Bio-Rad laboratories), 0.75 ml stacking buffer (0.5 M Tris HCl, pH 6.8), 1.815 ml H_2O, 30 µL 10 % SDS, 15 µL 10% ammonium peroxodisulphate (APS), and 3 µL N,N,N',N'-tetramethylethylenediamine (TEMED).

40 µg, 10 µg, and 4 µg of protein was mixed with an equivalent volume of loading buffer and heated at 100°C for 10 minutes upon a Multi-Block heater (Lab-Line Instruments Inc.). The solutions were allowed to cool and centrifuged before loading onto the gel. A constant voltage of 200 volts was applied across the electrodes at 50mA until the dye front reached the bottom of the gel. Gels were stained with a solution of 0.1% Coomassie Brilliant Blue R_{250}, 40% methanol, and 10% acetic acid. Gels were then destained in an equivalent solution minus the Coomassie Brilliant Blue R_{250} until bands were clearly distinguishable.

2.3.4 Equilibration

A 5 ml aliquot of *S. thermophilum* PP_i-PFK in storage buffer was transferred into a buffered solution containing 10 mM Hepes, pH 6.8, 3.5 mM $MgCl_2$, 100 mM NaCl, 0.1 mM Fructose-1,6-bisphosphate (F16bP buffer). Equilibration was achieved by the dropwise addition of 5 ml of the F16bP buffer to the 5 ml aliquot, in 200 µL aliquots. Particular care was taken to observe possible precipitation, indicative of the protein coming out of solution. No precipitate was observed during or after the addition of 5ml F16bP buffer. The volume of the solution was then approximately halved by centrifugation at 3000 rpm (Sorvall GS3 rotor) in a 30 K cut-off centricon microfilter (Millipore) at 4°C. The process was repeated until approximately 99% of the storage

buffer had been replaced with F16bP buffer. The final *S. thermophilum* PP_i-PFK concentration was 8.8 mg.ml $^{-1}$.

2.3.5 Commercial Crystallization Screens

Initial crystallization conditions were identified by vapor diffusion experiments using commercial screens: Crystal Screen and Crystal Screen 2: Macromolecular Crystallization Kits (Hampton Research). Each screen contained 50 individual reservoir solutions, the first 48 of which were used. 0.5 ml of each of the 96 reservoir solutions was added to an individual well of a 24 well tissue culture plate. 1.5 μL of the buffered 8.8 mg.ml^{-1} *S thermophilum* PP_i-PFK solution was mixed with 1.5 μL of the reservoir solution upon a plastic cover-slip to form a small spherical drop of 3 μL, which was suspended over the reservoir solution. Each well was sealed using petroleum jelly. A large, single crystal, which diffracted to 1.85 Å was grown from reservoir solution number 47 from the commercial screen Crystal Screen: Macromolecular Crystallization Kit (Hampton Research) of *S. thermophilum* PP_i-PFK. The reservoir solution comprised of 0.1 M sodium acetate trihydrate, buffered at pH 4.6, and 2 M ammonium sulfate.

2.3.6 X-ray Data Collection

All diffraction experiments were carried out at Massey University using a Rigaku RU200 Cu-Kα rotating-anode generator with capillary-focusing optics coupled with a Rigaku Raxis IIC image plate detector. The power setting was 1.1 kW and the beam was collimated to 0.1 mm.

2.3.6.1 Data set: 2.2 Å resolution

The *S. thermophilum* PP_i-PFK crystal (see section 2.2) was transferred from the mother liquor into an equivalent solution with 25% additional glycerol, constituting the cryoprotectant. After soaking for approximately 1 minute the crystal was mounted in a fiber loop and flash-frozen within a stream of liquid nitrogen vapor at 114 K (Oxford Cryosystems). Three oscillation photographs were taken at phi = 15°, 60°, and 150° with an exposure time of 30 minutes. Diffraction spots (Bragg spots) were observed to approximately 2 Å. The spots could clearly be resolved with little overlap. Based upon

the misorientation indices obtained from the three oscillation photographs data set collection was started at phi = 60°. Setting the crystal to film distance at 130 mm, 300 oscillation photographs were taken over 150° with an individual exposure time of 30 min and oscillation range (Δphi) of 0.5°.

2.3.6.2 Data set: 1.85 Å resolution

The mother liquor turned opaque when flash-frozen within a stream of liquid nitrogen vapor at 114 K (Oxford Cryosystems). A cryoprotectant solution was prepared by adding glycerol to the reservoir solution to give a concentration of 15%. The large crystal (see section 2.6) was mounted in a fiber loop and flash-frozen within a stream of liquid nitrogen vapor at 114 K (Oxford Cryosystems).). Setting the crystal to film distance at 80mm, 3 oscillation photographs were taken at phi = 15°, 60°, and 150°. Spots could clearly be resolved with little overlap. Based upon the misorientation index data set collection was started at phi = 60°. 312 oscillation photographs were collected over 156° with an individual exposure time of 30min and an oscillation range (Δphi) of 0.5°.

2.3.7 Crystallographic Data Processing

The diffraction images of the respective data sets were auto-indexed and batch refined using DENZO (Otwinowski, Z., Minor, W. 1997). DENZO indexes the diffraction pattern of individual oscillation photographs, refines crystal and detector parameters, and integrates diffraction maxima. The following methodology applies to processing of both data sets. A peak search of the three initial oscillation photographs determined the spots to be used in the auto-indexing subroutine. Auto-indexing gave an unambiguous primitive monoclinic (P2) Bravais lattice solution for both data sets based upon comparison of the distortion index (a percentage value indicating the degree of distortion of the unit cell parameters required to fit the lattice) for each of the fourteen Bravais lattices. Based upon the Bravais lattice manually assigned DENZ assigns a profile to the predicted Bragg reflections. The profiles were defined as circular with a radius of 0.4 mm for all predicted reflections. The predicted reflections aligned with the spots indicating a correct solution. Once the predicted and observed reflections matched the positions parameters defining crystal orientation, unit cell, mosaicity, and detector orientation were

refined: initial refinement is typically executed at a medium resolution limit with subsequent rounds of refinement at higher resolutions. The initial frames of the 1.85 Å data set were refined at 2.7 Å, 2 Å, and finally 1.85 Å. The positions were refined until the predicted decrease in χ^2 (the average ratio, squared, of the error in fitting, divided by the expected error) was near zero. Intensities were measured for the fitted profiles. Values determined from refinement of the parameters for the initial oscillation photographs were defined for the batch refinement of all oscillation photographs. Partial reflections, a reflection that is predicted to be only partially recorded in the positional refinement were used in the batch refinement.

SCALEPACK (Otwinowski, Z., Minor, W. 1997) was use to scale the individual films and merge equivalent intensity measurements in the data. Postrefinement optimized the global crystal mosaicity and unit cell dimensions, yielding better scale factors and merging statistics. The individual oscillation photographs were subsequently reprocessed (batch refined) with the new mosaicity and unit cell dimensions defined and SCALEPACK rerun. SCALEPACK applies Bayesian rejection of spurious observations (hkl). If the probability of an observation being an outlier is higher than a specified value it is rejected. The rejection probability should approximate the number of outliers divided by the number of observations. Therefore the higher the redundancy of the data the less likely the inappropriate rejection of an outlier. The rejection probability was specified as 0.001. Iterative rejections were carried out until χ^2 approached 1. Structure factor amplitudes were calculated from the merged intensity data output from SCALEPACK using TRUNCATE by first converting the intensity output into an .mtz format using SCALEPACK2MTZ. TRUNCATE and SCALEPACK2MTZ are from the CCP4 version 3.5 suite of programs (CCP4 [Collaborative Computational Project, Number 4] 1994). TRUNCATE employs a Wilson plot to derive an overall scale factor (the exponential of the negative x intercept) and temperature (B) factor (equal to the negative slope of the fall-off in mean intensity) for the diffraction intensity data. The program also uses a Bayesian statistical procedure formulated by French and Wilson (1978) to inflate the intensity values of weak reflections, especially negative intensities and convert the standard errors of the intensities to the standard errors of the structure factor amplitudes

2.3.8 Molecular Replacement

Calculation of electron density (ρ[xyz]) maps requires that structure factor amplitudes and phases be known.

$$\mathbf{F}(hkl) = \sum_{j=1}^{N} fj_{hkl} \exp[2\pi i\,(hx_j + ky_j + lz_j)] = |F(hkl)|\exp\,[2\pi i\,(\alpha_{hkl})]$$

(Eqn. 2.2)

The structure factor amplitude (F[hkl]) is a geometrical complex sum of the atomic scattering magnitudes (fj_{hkl}), given in electrons for a spherical atom j centered at x_j, y_j, z_j at some resolution specified by the reciprocal vector hkl. The structure factor equation (Eqn. 2.2) is comprised of an amplitude (|F[hkl]|) and phase (exp[2πiαhkl]) component.

$$I(hkl) \propto |F(hkl)|^2$$

(Eqn. 2.3)

F(hkl) is proportional to the square root of the measured diffraction intensity (I[hkl]) (Eqn. 2.3). Note however that only the scalar magnitude of F(hkl), |F(hkl)| can be determined from the intensities, hence the phase component must be determined by other means: molecular replacement is one such method. The structure factor amplitudes, F(hkl), for a polypeptide can be determined from its atomic coordinates (x, y, and z) within some arbitrary unit cell. Molecular replacement approximates the scattering of an unknown molecule from the calculated scattering of a search model, a molecule of known three-dimensional structure. For $F(hkl)_{obs}$ to equal $F(hkl)_{calc}$ the atomic model of the search model must closely resemble the unknown structure and be in the correct orientation and position within the unit cell. A molecule whose three dimensional structure has been determined, which shares high sequence similarity with that of the unknown molecule can be used as a search model, based upon the premise that the primary amino acid sequence dictates the folded, tertiary structure of a polypeptide. The orientation and position of the search model within the unit cell is defined by six parameters, three of which specify orientation (rotation [α, β, and γ]) and three which

specify position (translation [x, y, and z]). Note that the position need only be defined by two parameters for a polar space group. The approximate orientation and position parameters (approximate because the search model is only an approximation of the unkown) can be determined separately via the Patterson synthesis (Eqn. 2.4): a Fourier map calculated with the squares of the structure factor amplitudes. The Patterson map corresponds to an inter-atomic vector map defined by the unit vector u,v,w. Each peak in the map corresponds to a vector between atoms in the crystal and the intensity of the peak is the product of the electron densities of each atom.

$$\mathbf{P}(u,v,w) = \sum |F(hkl)|^2 \exp[2\pi i\, hu + kv + lw)$$

(Eqn. 2.4)

2.3.8.1 Rotation function

The Patterson synthesis is partially independent of the translation of the atomic model relative to the unit cell axes and symmetry operators. Because inter-atomic vectors can be sorted into those that are of intra-molecular origin and those that arise from inter-molecular vectors, we can take advantage of the intra-molecular vectors as they should be independent of translation and only depend on the model orientation specified by the Eulerian rotation angles α, β, and γ. Hence the rotation function Patterson synthesis is calculated in such a way that contributions from inter-molecular vectors are minimized. The rotation function is used to compare the Patterson calculated from the search model (being careful to avoid inter-molecular vectors) with the Patterson from the observed data. A product function is used to quantify the overlap between the observed and calculated Patterson syntheses by sequentially summing the product or overlap between the observed and model vectors at each vector (u,v,w). This results in a maximum when the two Pattersons overlap, i.e. when the search model is in a similar orientation to the molecule in the observed unit cell. The rotation function R(**C**) is defined as;

$$R(\mathbf{C}) = \int_v P_{obs}(\mathbf{u})\, P_{calc}(\mathbf{Cu})\mathbf{du}$$

(Eqn. 2.5)

Where **C** is a matrix which rotates the models coordinates with respect to the unit cell, $P_{obs}(\mathbf{u})$ is the Patterson function of the crystal and $P_{calc}(\mathbf{Cu})$ is the self-convolution function of the rotated model. R(**C**) will have a maximum when the peaks of the two functions superimpose (Giacovazzo, C. *et al.* 2002), indicative of the search model orientation closely approximating that of the molecule in the crystal lattice, hence we have a solution to the rotation function.

Sometimes a clear solution is not evident, particularly if the structure of the search model is a poor approximation to our unknown molecule. If so then several parameters that affect the Rotation Function synthesis may be modified. These include the radius of integration that selects which inter-atomic vectors are used for the overlap calculation, the resolution cutoff of the crystallographic data, and the size of the model unit cell. For a search model, which is a good approximation to the structure being determined, the data resolution is often not crucial. Often data between 6 and 4 Å will yield an acceptable solution. Systematic errors are introduced by low resolution data as the search model does not take into account solvent, which will affect the observed data. Alternatively to altering the resolution a B-value may be applied to weaken (positive B) or strengthen (negative B) the high-resolution terms. Application of normalized structure factors, or E-values functions similarly to a negative B-value, strengthening the high-resolution data. The size of the unit cell of the search model and the angular step size, by which the model is rotated, can also be altered to give a clearer solution.

2.3.8.2 Translation function

The rotation function determines the orientation of the model. The translation function determines the absolute position within a unit cell. Resultant to a translation of the model within the unit cell is an equivalent translation of all symmetry related molecules (no translation is required for P1, i.e. when only one molecule is present in the unit cell), which alters the inter-molecular vectors. When the calculated Patterson cross-vectors superimpose with those of the observed Patterson cross-vectors the model is in the correct position within the unit cell.

$$T(t) = \int_v P_{obs}(\mathbf{u})\, P_{calc}(\mathbf{u},\mathbf{t})\, d\mathbf{u}$$

(Eqn.2.6)

Where $P_{calc}(\mathbf{u},\mathbf{t})$ is the calculated Patterson cross-vector's and $P_{obs}(\mathbf{u})$ the observed Patterson cross-vectors. The translation function (T) will have a maximum when the Patterson cross-vectors superimpose (Giacovazzo, C. *et al.* 2002). Note that the quality of the solution to the translation function is dependent upon the search model having the correct orientation (i.e. a correct rotation function solution).

2.3.9 AMORE

The PP_i-PFK structure from the Spirochete *B. burgdorferi* has been solved and refined at 2.55 Å resolution (Moore, S. *et al.* 2002; PDB 1KZH). The amino acid sequence of *B. burgdorferi* PP_i-PFK shares 56% identity with that of *S. thermophilum* PP_i-PFK. Subunit A of the *B. burgdorferi* PP_i-PFK homodimer with all solvent molecules removed was used as the search model. Residues that differed in identity between *B. burgdorferi* and *S. thermophilum* PP_i-PFK other than conservative substitutions (Val/Thr/Ile, Thr/Ser, Gln/Glu, Asn/Asp, Phe/Tyr) were replaced with Ala as all residues, with the exception of Gly, share a Cβ atom whose orientation is strictly stereochemically defined. Rotation and translation searches using the search model were carried out using AMORE (Navaza, J. 1994). An orthorhombic unit cell was defined for the search model with dimensions a = 125, b = 90, and c = 80 and crystallographic data within the resolution range 15 Å - 4 Å was used for the rotation and translation functions. An integration radius of 35 Å was stipulated. PDBSET from the CCP4 version 3.5 suite of programs 5 (CCP4 [Collaborative Computational Project, Number 4] 1994) was used to transform the model as dictated by the Eulerian angles and fractional coordinates calculated by AMORE (Navaza, J. 1994). Subunit A of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model was used as the search model for the 1.85 Å data set. The unit cell and crystallographic resolution range defined were identical to the 2.2 Å data set.

2.3.10 Restrained Crystallographic Refinement

A maximum likelihood target was used in all refinement steps, monitored by cross-validation, over a resolution range 50 - 2.2 Å and 50 - 1.85 Å, respectively. CNS protein, water, and ion topology and parameter libraries were used throughout the refinement (Brunger, A. *et al.* 1998). All manual rebuilding of the atomic model was carried out using TURBO-FRODO (Roussel, A., Cambillau, C. 1991).

2.3.10.1 Cross-validation

The free R factor (R_{free}) was employed to monitor the macromolecular refinement process (Brunger, A. 1992). The R_{free} statistic is calculated from a set of reflections independent of the refinement process. The omitted set is comprised of a randomly selected 5% of the total measured reflections. Meaningful refinement must take into consideration not only the drop in the R factor (R) but also a concomitant decrease in R_{free}. A drop in R without a concomitant reduction in R_{free} is indicative of over-fitting of the data. The minimization of R is the criterion in the refinement.

$$wR = \frac{\Sigma[w(|F_o| - |F_c|)]^2}{\Sigma w|F_c|^2}$$

(Eqn. 2.7)

2.3.10.2 CNS Rigid-body Refinement

Rigid body refinement of the initial model was carried out. Each subunit of the dimer was initially refined as a rigid body followed by refinement with the individual domains, found from the rotation and translation searches as rigid bodies.

2.3.10.3 CNS Restrained Crystallographic Conjugate Gradient Minimization

Subsequent to rigid body refinement the model was submitted to stereochemically restrained refinement with a maximum likelihood target. Minimization refinement is advantageous in that it takes into account phase errors in the calculated structure factors. Bond lengths and angles of the atomic model are maintained at sensible values based

upon a protein parameter file (protein_rep.param) read into the program derived from Cambridge Data Base model structures (Engh, R.A. and Huber, R. 1991).

2.3.10.4 Restrained Atomic Temperature factor Refinement

Random thermal motion causes the position of an atom to fluctuate about an equilibrium position of minimal energy. The extent of the fluctuation is directly proportional to the absolute temperature. The fluctuation causes aberration of the atoms ability to scatter, resulting in a more diffuse electron density with increasing values of $\sin\theta/\lambda$.

$$B = 8\pi^2 U(\text{Å}^2)$$

(Eqn. 2.8)

Where U is the mean square Gaussian displacement of an atom with respect to its equilibrium position. The atomic temperature (see Eqn. 2.8) factor for each individual atom in the model was refined with a maximum likelihood target. The B factor was tightly restrained initially. The target rms deviation of temperature factors for main-chain and side-chain atoms was set to 1.0 and 2.0, respectively. The target rms deviation of temperature factors for main-chain and side-chain angles was set to 2.0 and 2.5, respectively. The weight for B factor restraints was altered accordingly to achieve the target rms deviations. In the latter rounds of refinement these target constraints were relaxed.

A bulk solvent correction was applied for both minimization and B factor refinement. A mask around the molecule, calculated from coordinates from selected atoms defined the region outside of this as solvent. The solvent density level and solvent B factor were determined automatically.

2.3.10.5 Atomic Model Building and Electron Density Interpretation

After rigid-body and crystallographic conjugate gradient minimization refinement, SIGMAA maps; (2*m|F_o| - D|F_c|)^exp(i phi_calc) (2|F_o| - |F_c|) and (m|F_o| - D|F_c|)^exp(i phi_calc) (|F_o| - |F_c|) were calculated using phase information from the molecular replacement model. D is an estimate of error in the structure from coordinate errors and

m is the figure-of-merit (FOM) defined as the mean cosine of the phase error. As m approaches one the phase error approaches zero. The respective models were subjected to iterative rounds of manual rebuilding followed by minimization and B factor refinement. Manual rebuilding involved placement of incorrectly located atoms (indicated by negative difference density peaks within $|F_o| - |F_c|$ maps) into reasonable electron density in $2|F_o| - |F_c|$ maps whilst maintaining reasonable geometry i.e. bond lengths and angles. Missing portions of the structure, water molecules, and sulfate ions were built into positive difference density peaks within the $|F_o| - |F_c|$ maps.

Chapter 3

Results: Data Quality and Atomic Model Statistics

3.1 Introduction

S. thermophilum PP_i-PFK crystals diffracting to 2.2 Å and 1.85 Å were crystallized from 2 M $(NH_4)_2SO_4$, 0.1 M NaCl, and 0.1 M Hepes (pH 7.5) and from 0.1 M sodium acetate trihydrate (pH 4.6), 2 M ammonium sulfate, respectively. Diffraction data sets were collected and processed as outlined in section 2.3.6 and 2.3.7. The phase problem for both data sets was solved by molecular replacement (see section 2.3.8 and 2.3.9).

3.2 Space group and Unit Cell

Auto-indexing of the 2.2 Å and 1.85 Å data set diffraction images using DENZO (Otwinowski and Minor 1997) indicated a primitive monoclinic cell for both, assumed to belong to the P 2_1 space group, with distinct unit cell dimensions (see Table 3.1 and Figure 3.1).

	S. thermophilum PP_i-PFK	
	1.85 A Data set	2.2 A Data set
Unit cell		
a	66.804	61.876
b	103.311	96.503
c	83.547	103.650
α	90	90
β	102.861	94.765
γ	90	90
Space group	$P2_1$	$P2_1$
V_{cell}	5.62×10^5 Å^3	6.18×10^5 Å^3
V_m	2.32 Å^3.Da^{-1}	2.55 Å^3.Da^{-1}

Table 3.1: Unit cell dimensions and space group of the *S. thermophilum* PP_i-PFK 1.85 Å and 2.2 Å unit cells.

$$V_{cell}\,[\text{primitive monoclinic}] = a.b.c.\sin\beta \qquad \text{(Eqn. 3.1)}$$

$$M_r\,[\textit{S. thermophilum}\ PP_i\text{-PFK}] = 60.5\ \text{kDa}.$$

$$V_m = V_{cell} / z.M_R \qquad \text{(Eqn. 3.2)}$$

The Matthews' coefficient (V_m) for the 1.85 Å and 2.2 Å unit cell was calculated to be 2.32 $Å^3.Da^{-1}$ and 2.55 $Å^3.Da^{-1}$, respectively with z = 4, indicating a dimer in the asymmetric unit of the respective unit cells, as there are two unique positions in a primitive monoclinic Bravais lattice (P2 or $P2_1$). The estimated overall atomic temperature factors (B) and scale factors for the 1.85 Å and 2.2 Å data sets were 22.3 $Å^2$ and 27.9 $Å^2$ and 135.6 and 227.8 respectively, determined from the Wilson plots calculated by TRUNCATE (see section 2.3.7). The difference in the temperature factors is reflective of the difference in resolution (Å). The greater the thermal motion (measured by B) the lower, in general the resolution to which a crystal normally diffracts. The difference in B factor suggests fewer crystal contacts within the 2.2 Å unit cell, facilitating greater thermal motion.

The diffraction images of the respective data sets were auto-indexed and batch refined using DENZO and SCALEPACK was used to scale the individual films and merge equivalent intensity measurements in the data as decribed in section 2.3.7 (Otwinowski, Z., Minor, W. 1997). Six iterations of rejection for the 1.85 Å data set and four iteration for the 2.2 Å data set were executed until χ^2 (defined in section 2.3.7) approached 1. For final refinement statistics of the respective data sets see Table 3.2a and 3.2b.

$P2_1$ 2.2 Å Data set					
Resolution shell (Å)	No. of measured reflections	Completeness (%)	I / σ	Redundancy	R-merge
50.00 - 3.76	12023	96.2	24.0	3.08	0.060
3.76 - 2.99	11893	96.6	18.1	2.97	0.086
2.99 - 2.61	12077	98.2	8.6	3.00	0.149
2.61 - 2.37	12163	99.0	4.6	3.00	0.223
2.37 - 2.20	12165	99.2	2.8	2.97	0.298
All hkl	60321	97.8	13.2	3.00	0.094

Table 3.2a: Data refinement statistics for the *S. thermophilum* PP_i-PFK P 2_1, 2.2 Å data set. Rejected reflections (19675) are not included.

$P2_1$ 1.85 Å Data set					
Resolution shell (Å)	No. of measured reflections	Completeness (%)	I / σ	Redundancy	R-merge
50.00 - 3.99	9505	98.6	25.3	3.1	0.054
3.99 - 3.16	9508	99.6	25.6	2.8	0.062
3.16 - 2.76	9468	99.4	20.9	2.5	0.075
2.76 - 2.51	9395	99.0	15.4	3.1	0.088
2.51 - 2.33	9347	98.5	12.2	3.1	0.098
2.33 - 2.19	9305	98.2	9.5	3.1	0.112
2.19 - 2.08	9253	95.7	7.2	3.1	0.136
2.08 - 1.99	9182	95.2	5.1	3.1	0.179
1.99 - 1.92	9192	95.0	3.5	3.1	0.252
1.92 - 1.85	9118	96.5	2.3	3.0	0.367
All hkl	93273	98.2	16.5	3.1	0.076

Table 3.2b: Data refinement statistics for the unique *S. thermophilum* PP_i-PFK P 2_1, 1.85 Å data set. Rejected reflections (5853) are not included.

3.3 AMORE: 2.2 Å solution

The temperature factor for all atoms in the model were reset to 27 (value calculated from TRUNCATE) based upon the overall value derived from the Wilson plot, and atomic occupancies set to 1.

3.3.1 Rotation function

Peak	Eulerian Angles			Peak
no.	α	β	γ	(σ)
1	23.50	59.80	47.02	10.6
2	333.35	39.60	241.00	10.1
3	138.78	72.67	203.18	8.2
4	41.97	41.22	146.13	7.6

Table 3.3: Results of the AMORE rotation search.

Data between 15 and 4 Å were used in the computation. An integration radius of 35 Å was stipulated. Consistent with the prediction of a dimer in the asymmetric unit and a primitive monoclinic unit cell two unambiguous peaks were evident at heights of approximately 10.6 and 10.1σ, clearly distinct from the next highest peak at 8.2σ (see Table 3.3).

3.3.2 Translation function

Translation searches were carried out within a resolution range 15 Å and 4 Å for each of the best four rotation solutions in a $P2_1$ space group (see Table 3.4).

Solution	Eulerian Angles			Fractional coordinates			C.C.	Rfac
	α	β	γ	Tx	Ty	Tz		%
SOLUT_1	23.50	59.80	47.02	0.1618	0.0000	0.1852	0.193	54.6
SOLUT_2	* 333.35	39.60	241.00	0.7115	0.8510	0.6872	0.309	51.2
SOLUT_3	* 138.78	72.67	203.18	0.2379	0.9790	0.5308	0.17.0	55.4
SOLUT_4	* 41.97	41.22	146.13	0.7203	0.9357	0.1712	0.17.2	55.5

Table 3.4: Amore translation solutions for each of the four best rotation solutions. * Solutions calculated with SOLUT_1 fixed.

$$\text{C.C.} = \text{Correlation Coefficient} = \frac{\{ |F_o|^2 * |F_c|^2 - \{|F_o|^2\}\{|F_c|^2\} \}}{\text{sqrt} (\{|F_o|^4 - \{|F_o|^2\}^2\} \{|F_c|^4 - \{|F_c|^2\}^2\})}$$

(Eqn.3.3)

Eqn. 3.3: Correlation coefficient derived from observed and calculated structure factors within the translation function. { denote the average over a defined resolution range.

Note the translation along the y axis (Ty) is equal to zero for the first solution as the absolute y axis coordinate value is arbitrary in a polar space group such as $P2_1$. The top two rotation function solutions produced the top two solutions for the translation search, which suggests that they were indeed the correct solutions. The two solutions were clearly distinct, ranked based upon their correlation coefficients. Note how solutions SOLUT_3 and SOLUT_4 give poor correlation coefficients and high R factors, indicating that they are false solutions to the molecular replacement problem. Solutions 1 and 2 represent the B and A subunits, respectively of the *S. thermophilum* PP_i-PFK dimers in opposing asymmetric units. A 2_1 symmetry operation about the y axis of the first solution model using the refined unit cell dimensions from SCALEPACK generated subunit A in the other asymmetric unit, forming a dimer with the second solution model and vice versa. The ability to effect a solution with sensible symmetry related contacts between dimers and sensible contacts across the dimer interface validate the molecular replacement solution in a $P2_1$ cell (see Figure 3.1).

3.4 AMORE 1.85 Å solution

The temperature factors for all atoms in the model were reset to 22 (calculated from TRUNCATE) based upon the overall value derived from the Wilson plot, and atomic occupancies were set to 1.

3.4.1 Rotation function

Peak	Eulerian Angles			Peak
no.	α	β	γ	(σ)
1	315.07	85.75	194.82	15.0
2	152.80	86.69	69.84	13.1

Table 3.5: Results from the AMORE rotation search.

Data between 15 and 4 Å were used in the computation. Consistent with the prediction of a dimer in the asymmetric unit in a primitive monoclinic unit cell two unambiguous peaks were evident at heights of approximately 15.0σ and 13.1σ (see Table 3.5). Note that four solutions were shown in Table 3.3 and 3.4 to highlight the true solutions of the searches.

3.4.2 Translation function

Translation searches were carried out within a resolution range 15 and 4 Å for each of the rotation solutions (see Table 3.6).

Solution	Eulerian Angles			Fractional coordinates			C.C	Rfac
	α	β	γ	Tx	Ty	Tz		%
SOLUT_1	315.07	85.75	194.82	0.2222	0.0000	0.4205	0.25	51.6
SOLUT_2 *	152.80	86.69	69.84	0.7430	0.0255	0.1255	0.40	46.7

Table 3.6: Amore translation solutions for each of the rotation solutions. * Solution calculated with SOLUT_1 fixed.

In agreement with the rotation function the top two solutions produced the top two solutions for the translation search. The two solutions were clearly distinct, ranked based upon their correlation coefficients. The two best solutions are related by a screw-dyad (2_1) symmetry axis. Solutions 1 and 2 represent the A and B subunits, respectively of the

S. thermophilum PP_i-PFK dimers in opposing asymmetric units. A symmetry operation parallel to y of the first solution model, labeled subunit B using the refined unit cell dimensions from SCALEPACK, generated subunit B in the other asymmetric unit, forming a dimer with the second solution model, labeled subunit A and vice versa. The ability to effect a solution with sensible crystal contacts and sensible contacts across the dimer interface validates the choice of $P2_1$ as the space group for the molecular replacement solution (see Figure 3.1).

3.5 Crystal Packing

Both unit cells belong to a $P2_1$ space group with a dimer in the asymmetric unit, yet the unit cell volume (V_{cell}) of the 1.85 Å data set is approximately 10% smaller than the 2.2 Å data set, indicating tighter molecular packing in the crystal and therefore more extensive crystal contacts, as hypothesized earlier (see Table 3.7 and 3.8).

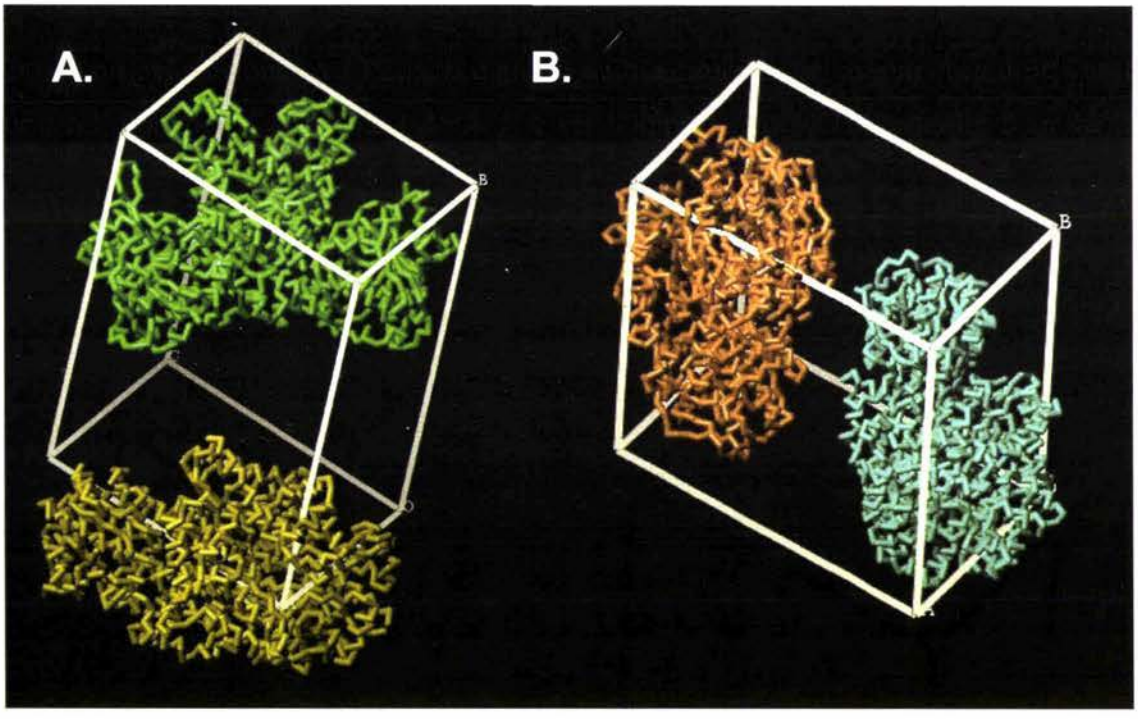


Figure 3.1: Diagram of the *S. thermophilum* PP_i-PFK homodimer in each asymmetric unit of the 1.85 Å (A.) and 2.2 Å (B.) data sets unit cell ($P2_1$). Figure produced with TURBO-FRODO (Roussel and Cambillau 1991).

Symmetry Related Hydrogen Bonding					
Parent Molecule Atom	Secondary structure	Symmetry Related Neighbor	Secondary structure	Distance Å	Symmetry operation
P16B O	α1–α2	R63A NE	β2	2.93	-x, y-1/2, -z+1
P16B O	α1–α2	R63A NH2	β2	2.83	-x, y-1/2, -z+1
V15B N	α1–α2	S65AOG	β2–β3	2.91	-x, y-1/2, -z+1
E25B OE2	α2–β1	Q23A NE2	α2	3.1	-x, y-1/2, -z+1
R71B NH2	β2–β3	Q191A NE2	α8	3.3	x, y-1, z
R71B NH2	β2–β3	Q191A O	α8	2.99	x, y-1, z
R43B NH2	α3	R403B NH2	α17	3.04	-x, y+1/2, -z
D326B OD1	α13	K334A NZ	α13	2.89	-x+1, y-1/2, -z
S350A OG	α14	G491B O	β14–β15	2.65	-x+1, y+1/2, -z
E409A OE1	α17	R511B NH2	α21	3.26	x+1, y, z
E410A OE1	α17	R511B NH1	α21	3.21	x+1, y, z
A406A O	α17	R511B NH2	α21	3.24	x+1, y, z
G411B O	α17–β11	R498B NH1	β15–α21	2.61	-x, y-1?2, -z

Table 3.7: Crystal contacts made by the *S. thermophilum* PP_i-PFK 2.2 Å atomic model. No direct interactions exist between the asymmetric units. All direct contacts shown exist between dimers from different unit cells. Highlighted is an asymmetric crystal contact between α14 of the helical domain (subunit A) and the β14-β15 β-hairpin (subunit B).

Symmetry Related Hydrogen Bonding

Parent Molecule Atom	Secondary Structure	Symmetry Related Neighbor	Secondary Structure	Distance Å	Symmetry operation
Q35A NE1	β1–α3	D293B OD2	α12	2.87	x-1, y, z
P36A O	β1–α4	R403B NH2	α17	2.89	x-1, y, z
E44A OE2	α3	E409B OE2	α17	2.84	x-1, y, z
K47A NZ	α3	E409B OE2	α17	3.02	x-1, y, z
K47A NZ	α3	E410B OE1	α17	3.15	x-1, y, z
K47A NZ	α3	E410B OE2	α17	3.07	x-1, y, z
R48A NE	α3	D168B OD1	α7–β6	3.33	x-1, y, z-1
R48A NH2	α3	G166B O	α7–β6	3.33	x-1, y, z-1
D120A OD2	α5–β5	E347B OE2	α14	2.52	-x, y+1/2, -z
K163A NZ	α7	Q35B OE1	β1–α3	2.85	x-1, y, z-1
H165A NE2	α7	E347B OE2	α14	3.10	-x, y+1/2, -z
Q192A NE1	α8	G55B O	α3–β2	3.05	x-1, y, z-1
E212A OE1	β7–α12	T283B OG1	α11–α12	3.07	x-1, y, z
E215A OE1	β7–α12	N281B OD1	α11–α12	2.87	x-1, y, z
S285A OG	α12	R63B NH2	β2	2.69	-x, y+1/2, -z+1
S285A OG	α12	R28B NH1	α2–β1	2.97	-x, y+1/2, -z+1
S285A OG	α12	R28B NE	α2–β1	3.16	-x, y+1/2, -z+1
D289A OD2	α12	R63B NE	β2	2.50	-x, y+1/2, -z+1
K464A NZ	β12	N281B O	α11–α12	2.90	x-1, y, z
N465A OD1	β12–β13	E279B O	α11	3.13	x-1, y, z
N465A ND2	β12–β13	E279B O	α11	2.96	x-1, y, z
R468A NH1	β12–β13	R278B O	α11	3.28	x-1, y, z
R468A NE	β12–β13	E279B OE2	α11	3.17	x-1, y, z
R468A NH1	β12–β13	E279B O	α11	3.01	x-1, y, z
M493A N	β15	Q192B O	α8	3.06	x-1, y, z-1
D503A OD2	β15–α21	N359B N	α15	2.99	x-1, y, z
E505A OE2	β15–α21	A355B O	α14	2.78	x-1, y, z
E505A OE1	β15–α21	S357B O	α14–α15	3.28	x-1, y, z
E505A OE1	β15–α21	S357B N	α14–α15	2.88	x-1, y, z
K510A NZ	α21	E327B OE1	α13	2.99	x-1, y, z
Q551A NE1	α23	N281B OD1	α11–α12	3.15	x-1, y, z

Table 3.8: Crystal contacts made by *S. thermophilum* PP_i-PFK 1.85 Å atomic model. Crystal contacts are more significant than the 2.2 Å atomic model as a result of the reduced unit cell volume. Highlighted are asymmetric crystal contact formed with the helical domain of subunit B. Underlined interactions are directly attributable to protonated carboxylate groups.

Recall from section 3.1 that the *S. thermophilum* PP_i-PFK crystals diffracting to 2.2 Å and 1.85 Å were crystallized at pH 7.5 and pH 4.6, respectively. The p*Ka* of the terminal carboxylate group of aspartic and glutamic acid residues is 4.4, therefore at a pH of 4.6 a proportion will be protonated, as evidenced in Table 3.8.

3.6 Refinement of the *S. thermophilum* PP_i-PFK Atomic Models

The *S. thermophilum* PP_i-PFK 2.2 Å and 1.85 Å dimers generated from the molecular replacement solutions exhibit a pseudo dyad symmetry axis owing to structural difference between the A and B subunits of the *S. thermophilum* PP_i-PFK homodimer. It was therefore possible to reduce the number of refinable parameters for both atomic models by including this non-crystallographic symmetry (NCS) information in restraining the refinement. The main chain and side chain atoms of the NCS groups were restrained. As the refinement progressed the NCS was continually modified (see Table 3.9).

NCS restraints Residue #	1.85 Å Model RMS difference * Å	2.2 Å Model RMS difference * Å
18-22	0.0510	0.0770
31-35	0.0270	0.0620
57-62	0.0560	0.0550
74-102	0.0480	0.0430
133-144	0.0900	0.0610
154-162	0.0590	0.0520
169-175	0.0390	0.0290
178-190	0.0670	0.0410
437-463	0.0360	0.0400
476-486	0.0400	0.0740
Average rms difference	0.0513	0.0530

Table 3.9: Final non-crystallographic symmetry (NCS) groups between subunits A and B for the *S. thermophilum* PP_i-PFK 2.2 Å and 1.85 Å atomic models. * RMS difference between selected regions (NCS groups) of the A and B subunits.

The average rms displacement for the NCS groups of the 1.85 Å and 2.2 Å atomic models is 0.0513 Å and 0.0530 Å, respectively compared to an average rms xyz displacement of 0.867 Å and 0.878 Å for all protein residues, respectively. This however can be misleading as structural deviation between regions under NCS constraint can be mitigated via an inflation of the respective B factors. The average residue B factor for NCS groups of the 1.85 Å and 2.2 Å atomic models is 21.72Å^2 (23.86 Å^2) and 27.29 Å^2 (32.04 Å^2) respectively which is significantly lower than the average residue B factor for all protein residues. It can therefore be concluded that the NCS definition is valid for both models. Close inspection of electron density verified this.

3.6.1 Model building: 2.2 Å model

After rigid-body and crystallographic conjugate gradient minimization refinement of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model, $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ type maps were calculated using phases from the refined molecular replacement model.

Immediately apparent were the reordering of specific surface loops and approximately the entire α-helical domain (residues 30-41 and 62-70 [N-terminal insert], 318-375 [helical domain], 398-415, and 485-506[β14-β15 β-hairpin]) of both *S. thermophilum* PP_i-PFK subunits relative to the *B. burgdorferi* PP_i-PFK A subunit. In these regions of the model, continuous $2|F_o| - |F_c|$ electron density at 1σ was absent and both positive and negative difference density peaks ($|F_o| - |F_c|$), contoured at -3 and 3σ, respectively, were observed. These regions were therefore excised from the atomic model. The large and small domains of the model exhibited a higher degree of structural conservation with the *B. burgdorferi* PP_i-PFK search model, as evidenced by continuous $2|F_o| - |F_c|$ electron density contoured at 1σ, calculated from the molecular replacement phases. Within these regions positive difference density peaks contoured at 3σwere observed in close proximity to the Cβ atoms of substituted Ala residues, facilitating the rebuilding of the correct residue in its correct conformation into the model (see Figure 3.2). Difference electron density peaks were also observed for residues that were conserved between *B. burgdorferi* and *S. thermophilum* PP_i-PFK, indicating an altered conformation of the amino acid side chain. Positive difference density indicating the conformation of all the missing amino acid side chains was not immediately evident or was ambiguous. As the

refinement progressed electron density appeared that was attributable to water molecules and sulfate ions. The initial criteria for building a water molecule into the model was a spherical 4σpeak in the $|F_o|$ - $|F_c|$ map and sensible hydrogen bonding contacts to the polypeptide, i.e. between 2.4 and 3.3 Å. Towards the latter stages of refinement the criteria were relaxed to spherical $3\sigma|F_o|$ -$|F_c|$ and 1σ $2|F_o|$ - $|F_c|$ peaks and sensible hydrogen bonding contacts. The criterion for a sulfate ion to be built into the model was a large 4σ $|F_o|$ -$|F_c|$ peak with a roughly tetrahedral shape. The sulfate oxygen atoms also had to be making sensible electrostatic or hydrogen bonding contacts with the polypeptide.

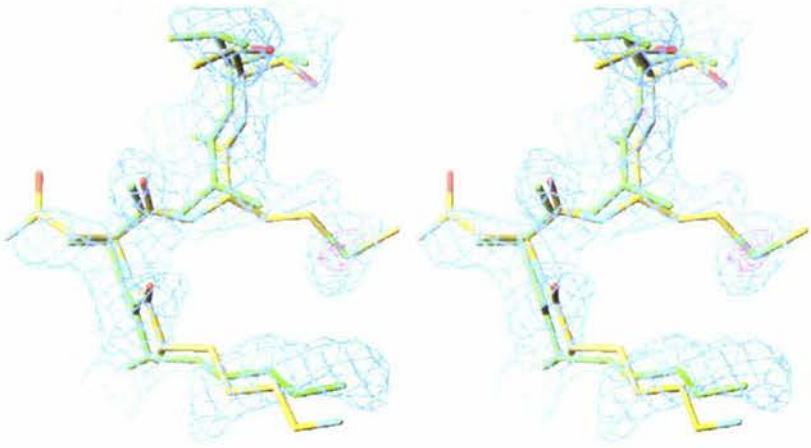


Figure 3.2: Initial electron density maps of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model are shown in blue ($2|F_o|$ - $|F_c|$), contoured at 1σ and purple ($|F_o|$ - $|F_c|$), contoured at 3σ). The molecular replacement solution model is shown in green. The positive difference electron density ($|F_o|$ - $|F_c|$) allows the modeling of Met282 B, absent in the initial model (nitrogen = blue, carbon = yellow, oxygen = red, sulfur = green).

3.6.2 Model building: 1.85 Å model

As subunit A of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model was used as the search model for the 1.85 Å AMORE solution only residues at the active site were replaced with an Ala residue. The dimeric AMORE solution was submitted to rigid body refinement

and initial $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ maps were calculated. Inspection of the initial electron density maps indicated a significant displacement of the helical domain, and the C-terminal domain to a lesser degree of both subunits A and B relative to the *S. thermophilum* PP_i-PFK 2.2 Å atomic model. The helical domain and β14-β15 β-hairpin of both subunits were subsequently excised from the *S. thermophilum* PP_i-PFK 1.85 Å model prior to the ensuing round of refinement (337-392 [α-domain], 487-499 [β14-β15]). The same criteria for the modeling of water and sulfate ions was applied, as previously defined (3.6.1).

3.7 Iterative Restrained Refinement

Improvement of the atomic models is concomitant with improvement of the atomic models crystallographic phases, therefore the quality of the respective electron density maps improved with each round of refinement. The appearance of interpretable electron density allowed further incorporation of the aforementioned excised loops and missing amino acid side chains. Iterative cycles of crystallographic x y z and B factor refinement and manual rebuilding of the 1.85 Å and 2.2 Å atomic models was terminated when the R factor reached 0.1923 (R_{free} = 0.2035 [FOM = 0.88794]) and 0.199 (R_{free} = 0.229 [FOM = 0.85447]), respectively. At no point during the refinement did R and R_{free} significantly diverge indicating that the data were not significantly over refined.

3.8 *S. thermophilum* PP_i-PFK: Quality of the Final Atomic Models

The *S. thermophilum* PP_i-PFK is a homodimer. Each subunit is comprised of residues 2-554. Continuous electron density ($2|F_o| - |F_c|$) contoured at 1σ was observed for the entire backbone of each subunit with the exception of the N-terminal methionine residue for both atomic models. Electron density was absent or was not interpretable for a number of amino acid side chains in both subunits (see Table 3.10 and 3.11). These residues were replaced with an alanine. The majority of these side chains were situated upon surface loops of the enzyme and are not in close proximity to, or presupposed to function in catalysis or substrate binding at the active site. The side chain conformation was not

interpretable for three residues thought to function at the active site, K148B, R253B, and H382B of subunit B of the 1.85 Å atomic model. K148B and R253B (MG**R**) are situated at the active site and His382B is positioned upon the 380-390 β-hairpin, which is likely to interact with Asp177 (GGD**D**) upon a conformational change positioning the loop into the active site (Moore, S.A. *et al.* 2002). In both atomic models the number of residues in subunit B with no interpretable side chain electron density is disproportionate to subunit A (see Table 3.10). This will result in an underestimation of the average B factor for subunit B. Note that the majority of uninterpretable residues are situated upon the helical domain (315B-390B). In the 2.2 Å atomic model the average main chain B factor for the helical domain of subunit A is 48.8 $Å^2$ compared to 59.6 $Å^2$ for subunit B. The greater disorder of the helical domain of subunit B relative to that of subunit A suggests that its movement within the crystal is less restricted. The crystal contacts formed by the respective helical domains are asymmetric. α14 of the helical domain of subunit A forms hydrogen bonding and van der Waals crystal contacts with the β14-β15 β-hairpin of subunit B of the neighboring dimer. This contact is not formed by the helical domain of subunit B (see Table 3.7). The fewer contacts formed by the helical domain of subunit B may account for its disorder relative to that of subunit A. Disorder of the helical domain is illustrated in Figure 3.5. Tables 3.10 and 3.11 summarize the temperature factor statistics for the respective *S. thermophilum* PP_i-PFK atomic models.

3.8.1 Active Site

Homologous to prokaryotic ATP-PFKs the active site of PP_i-PFKs is formed at a cleft between the N- and C-terminal domains. Examination of the active site of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model during refinement revealed tetrahedral positive difference density peaks ($|F_o| - |F_c|$) in the region expected to be occupied by the phosphate moieties of F16bP and the PP_i β-phosphate of subunit A and the F16bP 6-phosphate binding site of subunit B. Sulfate ions were modeled into the electron density with sensible bonding interactions. With an occupancy of 1, their real space correlation coefficients refined to 0.929 (SO_4 901 [6-phosphate, subunit A]), 0.961 (SO_4 902 [1-phosphate, subunit A]), 0.857 (SO_4 906 [β-phosphate, subunit A]), and 0.948 (SO_4 905[6-phosphate, subunit B]) with average temperature factors of 35.5, 40.3, 64.1 and 42.6

respectively. Tetrahedral positive difference density ($|F_o|$ - $|F_c|$) indicative of bound sulfate ions at the F16bP 6-phosphate binding site was observed in both subunits of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model.

1.85 A Model				2.2 A Model			
Subunit A	Seq. #	Residue	Subunit B	Subunit A	Seq. #	Residue	Subunit B
	1	M		-	1	M	-
ALA	2	T	ALA		2	T	ALA
ALA	71	R			47	K	ALA
	148	K	ALA		48	R	ALA
ALA	150	E			63	R	ALA
	164	K	ALA		103	R	ALA
	253	R	ALA		153	E	ALA
ALA	334	K	ALA	ALA	164	K	
ALA	335	E	ALA		332	L	ALA
ALA	337	E	ALA		334	K	ALA
	338	V	ALA		337	E	ALA
ALA	340	N	ALA		338	V	ALA
ALA	341	R	ALA		339	F	ALA
ALA	342	I			340	N	ALA
ALA	343	D	ALA	ALA	341	R	ALA
ALA	347	E			342	I	ALA
	355	K	ALA	ALA	343	D	ALA
	362	H	ALA	ALA	346	R	ALA
	378	D	ALA		347	E	ALA
	379	R	ALA		348	R	ALA
	382	H	ALA		350	S	ALA
	405	K	ALA		355	K	ALA
	414	K	ALA		365	S	ALA
ALA	488	K	ALA	ALA	389	R	
ALA	489	R	ALA	ALA	414	K	ALA
ALA	490	H	ALA	ALA	488	K	ALA
	492	K	ALA	ALA	489	R	ALA
	493	M	ALA		490	H	ALA
ALA	494	K	ALA	ALA	492	K	ALA
	498	R	ALA	ALA	494	K	ALA
ALA	554	E					

Table 3.10: Residues in the final *S. thermophilum* PP_i-PFK 1.85 Å and 2.2 Å atomic models with absent or uninterpretable electron density for the side chain. These residues have been modeled as Ala residues. Residues **bold** and underlined replaced with an Ala residue are putatively involved in substrate binding and/or catalysis or interact with residues which are.

	S. thermophilum PP_i-PFK			
	1.85 Å Model		**2.2 Å Model**	
	A	B	A	B
No. atoms in chain	4219	4161	4239	4184
No. of residues in chain	553	552	553	553
No. of side chain residues	434	420	442	426
B average for main chain $Å^2$	22.73	22.79	29.85	32.78
B average for side chain $Å^2$	25.20	24.83	31.84	33.66
B average for whole chain $Å^2$	23.91	23.74	30.80	33.20
rms deviation of residue B for main-chain Å	0.88	0.87	1.03	0.93
rms deviation of residue B for side-chain Å	1.69	1.67	1.85	1.67

Table 3.11: Synopsis of B factor statistics for the *S. thermophilum* PP_i-PFK 1.85 Å and 2.2 Å atomic models.

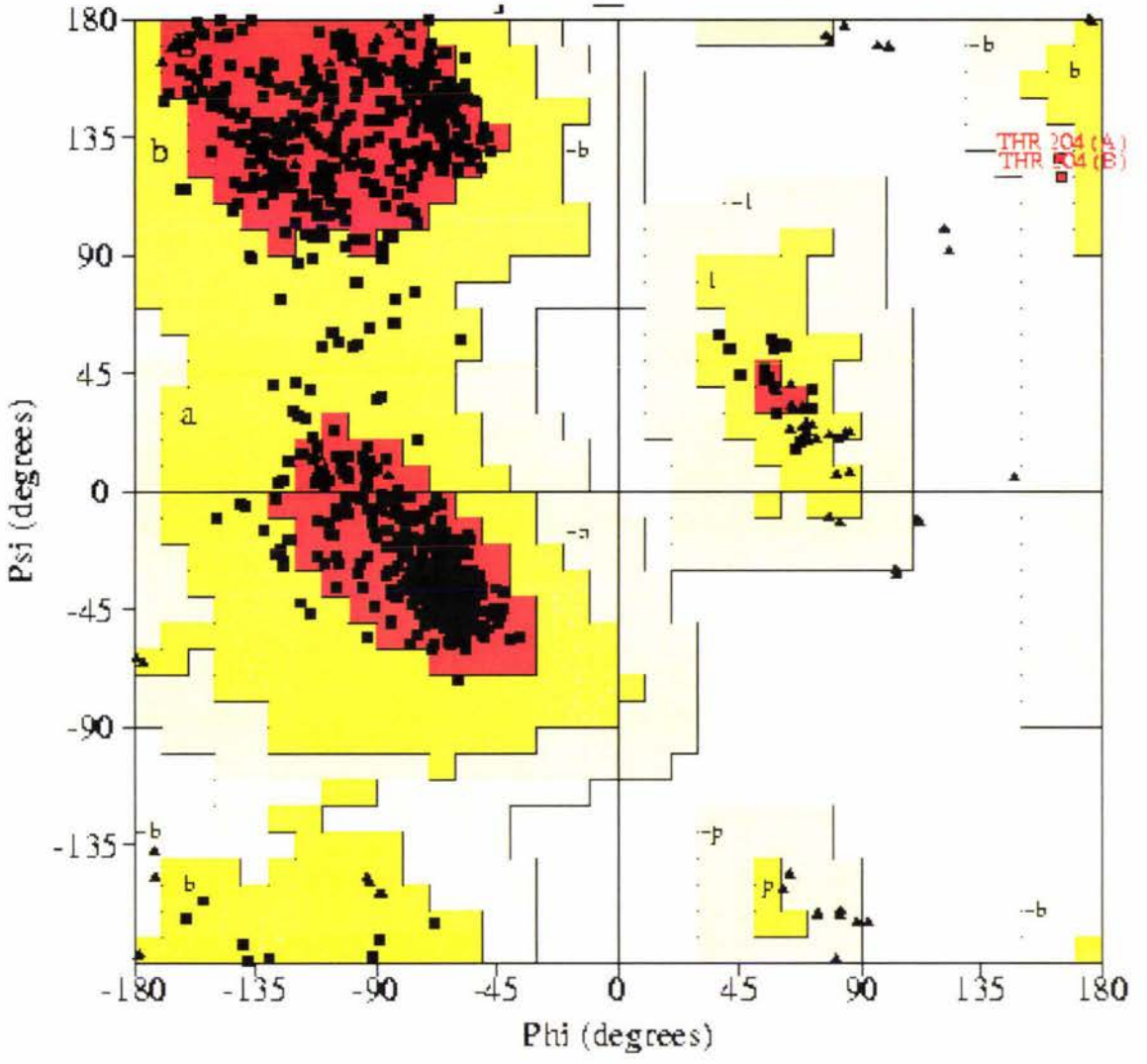
Refinement was tightly restrained. For the 1.85 Å and 2.2 Å atomic models the rms deviation from ideal bond lengths is 0.006 Å and 0.006Å and the rms deviation from ideal bond angle 1.235° and 1205°, respectively.

		S. thermophilum PP_i-PFK	
	Target rms $Å^2$	1.85 Å model rms $Å^2$	2.2 Å model rms $Å^2$
Bonded main-chain atoms	1.5	1.29	1.36
Bonded side chain atoms	2	2.05	2.19
Angle main-chain atoms	2	1.85	2.02
Angle side chain atoms	2.5	2.89	2.98

Table 3.12: B factor rms deviation targets for B factor refinement and statistics for the final atomic models.

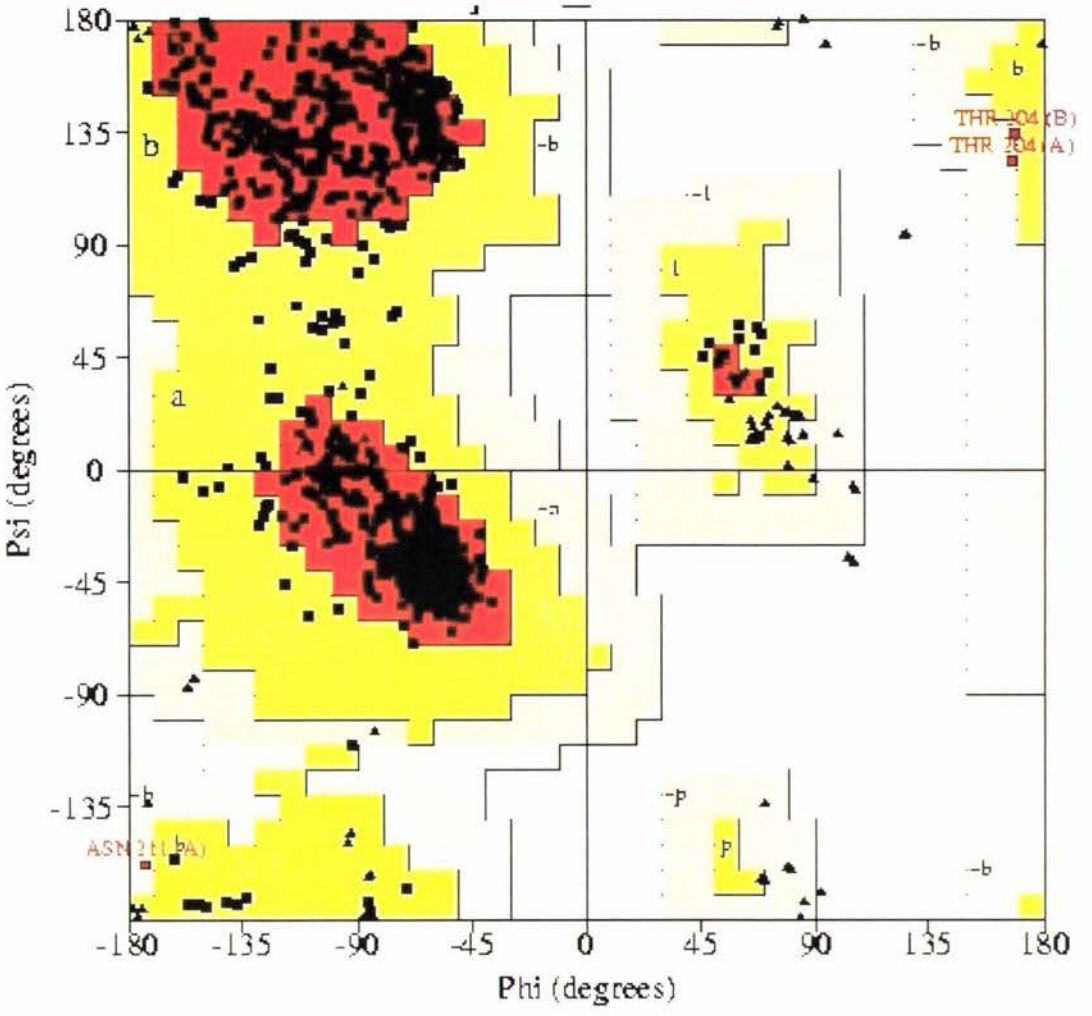
Main-chain torsion angles were not restrained during refinement of *S. thermophilum* PP_i-PFK atomic models. The stereochemical quality of the atomic models was assessed using PROCHECK (Laskowski, R. *et al.* 1993). Ramachandran plots of the *S. thermophilum* PP_i-PFK 1.85 Å (see Figure 3.3) and 2.2 Å (see Figure 3.4) atomic models illustrates that

91.5% and 91.1% of residues are in their most favored region, 8.3% and 8.6% in additional allowed regions, 0.2 % and 0.3% in generously allowed regions, and 0.0% in disallowed regions, respectively. The *S. thermophilum* PP_i-PFK atomic models are above average for all stereochemical parameters assessed by PROCHECK (Laskowski, R. *et al.* 1993) for structures determined and refined at 1.85 Å and 2.2 Å resolution, respectively. It can therefore be concluded that the stereochemical quality of the *S. thermpophilum* PP_i-PFK atomic models is very high.



Most favored regions [A,B,L]	91.50%
Additional allowed regions [a,b,l,p]	8.3
Generously allowed regions [~a,~b,~l,~p]	0.2
Disallowed regions	
No. Pro residues	56
No. Gly residues	94

Figure 3.3: Ramachandran plot of the refined *S. thermophilum* PP_i-PFK homodimer 1.85 Å model. Glycine residues are represented by triangles. Figure produced with PROCHECK (Laskowski, R. *et al.* 1993).



Most favored regions [A,B,L]	91.10%
Additional allowed regions [a,b,l,p]	8.6
Generously allowed regions [~a,~b,~l,~p]	0.3
Disallowed regions	
No. Pro residues	56
No. Gly residues	94

Figure 3.4: Ramachandran plot of the refined *S. thermophilum* PP_i-PFK homodimer 2.2 Å model. Glycine residues are represented by triangles. Figure produced with PROCHECK (Laskowski, R. *et al.* 1993).

A measure of the fit of the atomic models to the electron density can be determined by calculating the real space correlation coefficients (see Eqn.3.3). The average correlation coefficient for all protein atoms in the *S. thermophilum* PP_i-PFK 1.85 Å model is 0.9483 (where 1 is a perfect fit) with a standard deviation of 0.046 as compared to 0.7336 and 0.129 respectively for the initial AMORE solution model. The real space correlation coefficient, calculated with CNS (Brunger, A. *et al.* 1998), as a function of residue number for each subunit of the final *S. thermophilum* PP_i-PFK 1.85 Å atomic model is shown in Figure 3.5 and a synopsis of the data for the initial and final models is given in Table 3.13. A synopsis of the *S. thermophilum* PP_i-PFK 2.2 Å atomic models correlation coefficients for the initial and final data is given in Table 3.14.

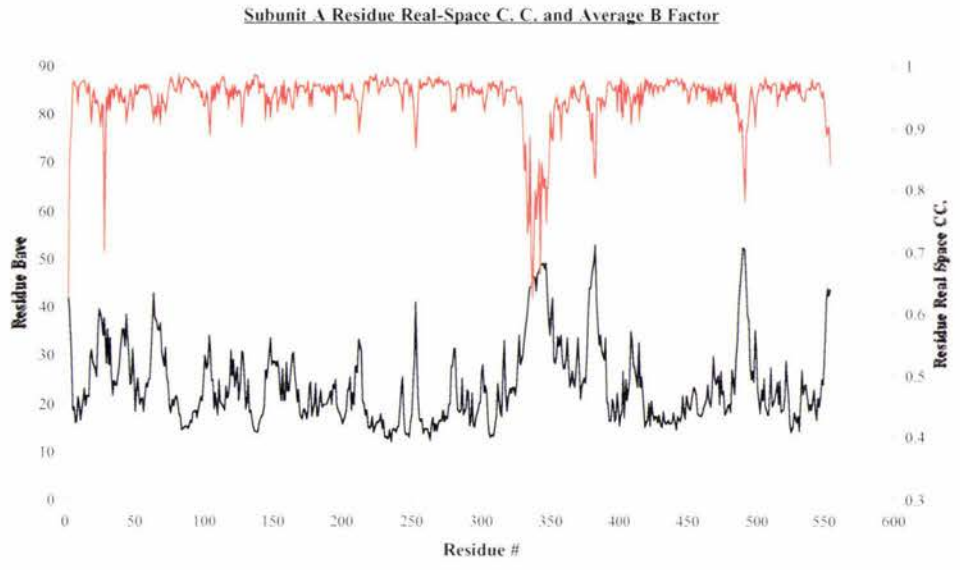
	Initial Model			Final Model		
	Subunit B	Subunit A	All	Subunit B	Subunit A	All
CC average	0.7106	0.7566	0.7336	0.9488	0.9478	0.9483
CC Std. Dev.	0.1416	0.1108	0.1291	0.0442	0.0481	0.0462
R average	0.2894	0.2434	0.2664	0.0512	0.0522	0.0517
R Std. Dev.	0.1416	0.1108	0.1292	0.0442	0.0481	0.0462

Table 3.13: Synopsis and comparison of the correlation coefficients (CC.) and R values for the initial AMORE solution model and the final *S. thermophilum* PP_i-PFK 1.85 Å model. Note values are for protein atoms only.

	Initial Model			Final Model		
	Subunit B	Subunit A	All	Subunit B	Subunit A	All
CC average	0.7032	0.7072	0.7052	0.9437	0.9507	0.9472
CC Std. Dev.	0.1383	0.1456	0.1420	0.0380	0.0317	0.0351
R average	0.2968	0.2928	0.2948	0.0563	0.0493	0.0528
R Std. Dev.	0.1383	0.1456	0.1420	0.0380	0.0317	0.0351

Table 3.14: Synopsis and comparison of the real space correlation coefficients (CC.) and R values for the initial and final *S. thermophilum* PP_i-PFK 2.2 Å model. Note values are for protein atoms only.

A.



B.

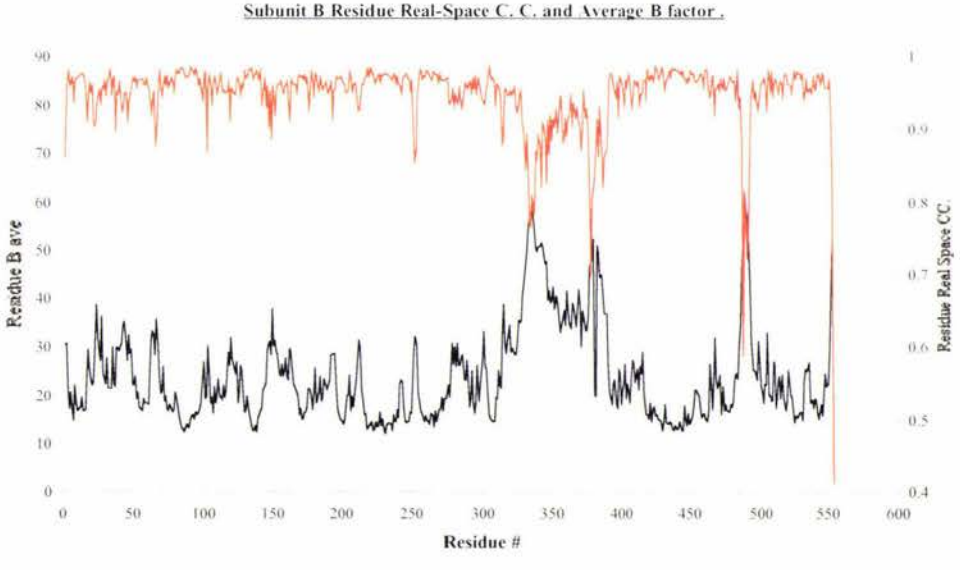


Figure 3.5: Real space correlation coefficient (red) and average residue B factor (black) as a function of residue number for subunit A (**A.**) and B (**B.**) of the *S. thermophilum* PP_i-PFK 1.85 Å model.

The plot of the real space correlation coefficient and average residue main chain B factor for subunit A and B of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model indicates that the helical domain of both subunits is disordered relative to the N- and C-terminal domains, feasibly as a result of movement within the crystal. An almost identical degree of disorder is seen for the 2.2 Å model (Data not shown). The disorder of the helical domain of subunit B is greater than that of subunit A as a result of asymmetric crystal contacts. The helical domain of subunit A forms more crystal contacts. The implications of helical domain movement will be discussed in Chapter 4.

Comparison of the *S. thermophilum* PP_i-PFK atomic model $2|F_o| - |F_c|$ electron density maps illustrates the superior quality of the 1.85 Å $2|F_o| - |F_c|$ electron density maps (see Figure 3.6 and 3.7).

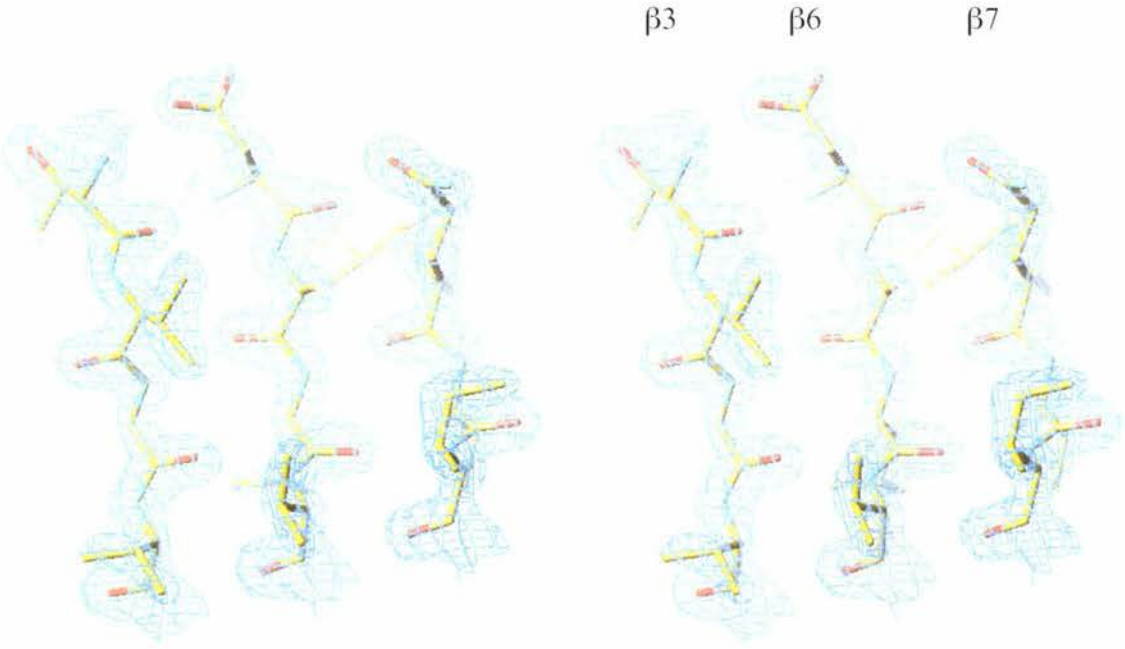


Figure 3.6: Stereo view of the final $2|F_o| - |F_c|$ maps contoured at 2σ for strands β3 (74-77B), β6 (168-171B), and β7 (196-199) of the Subunit B large domain β-sheet of the final *S thermophilum* PP_i-PFK 1.85 Å model.

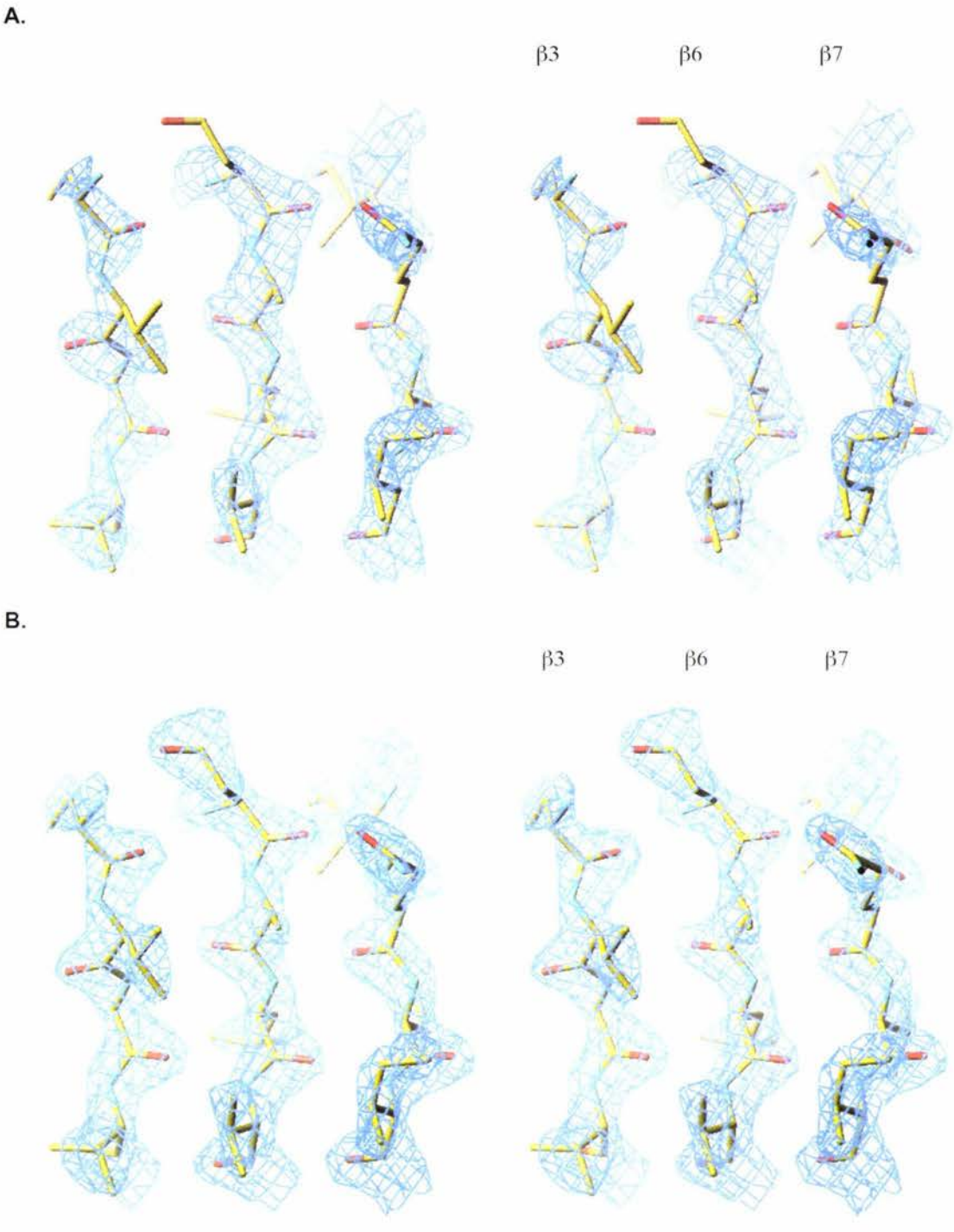


Figure 3.7: Stereo view of the initial (**A.**) and final (**B.**) $2|F_o| - |F_c|$ maps contoured at 2σ for strands β3 (74A-77A), β6 (168A-171A), and β7 (196A-199A) of the Subunit A large domain β-sheet of the final *S thermophilum* PP_i-PFK 2.2 Å atomic model (carbon = yellow, nitrogen = blue, oxygen = red).

Results: Synopsis

The structure of a PP_i-PFK from the thermophilic bacterium, *S. thermophilum* PP_i-PFK has been solved by molecular replacement in two distinct conformations at 1.85 Å and 2.2 Å resolution, respectively.

No. protein atoms	8380
No. water molecules	723
No. sulfate atoms	60
R_{free}(50-1.85 Å, 4650 reflections [4.9%])	0.1923
R(50-1.85 Å, 87943 reflections [93.2%])	0.2035
Residues in most favored Ramachandran regions	91.5
Residues in additional allowed Ramachandran regions	8.3
Residues in disallowed Ramachandran regions	0
Average B factor main-chain atoms $Å^2$	22.76
Average B factor side-chain atoms $Å^2$	22.96
Average B factor solvent $Å^2$	29.32
Average B factor sulfate ion atoms $Å^2$	49.42
RMS bond length Å	0.006
RMS bond angles degrees	1.235
Average RS correlation coefficient	0.948
Std. Dev. RS correlation coefficient	0.0461

Table 3.15: Synopsis of refinement and model statistics for the *S. thermophilum* PP_i-PFK homodimer 1.85 Å model.

No. protein atoms	8369
No. water molecules	697
No. sulfate atoms	50
Rfree(50-2.2 Å, 3039 reflections [4.9%])	0.2288
R(50-2.2 Å, 57232 reflections [92.8%])	0.1991
Residues in most favored Ramachandran regions	91.1%
Residues in additional allowed Ramachandran regions	8.6%
Residues in disallowed Ramachandran regions	0.0%
Average B factor main-chain atoms $Å^2$	31.31
Average B factor side-chain atoms $Å^2$	30.06
Average B factor solvent $Å^2$	33.58
Average B factor sulfate ion atoms $Å^2$	63.3
RMS bond length Å	0.006
RMS bond angles degrees	1.205
Average Real Space correlation coefficient	0.9415
Std. Dev. Real Space correlation coefficient	0.0413

Table 3.16: Synopsis of refinement and model statistics for the *S. thermophilum* PP_i-PFK homodimer 2.2 Å model.

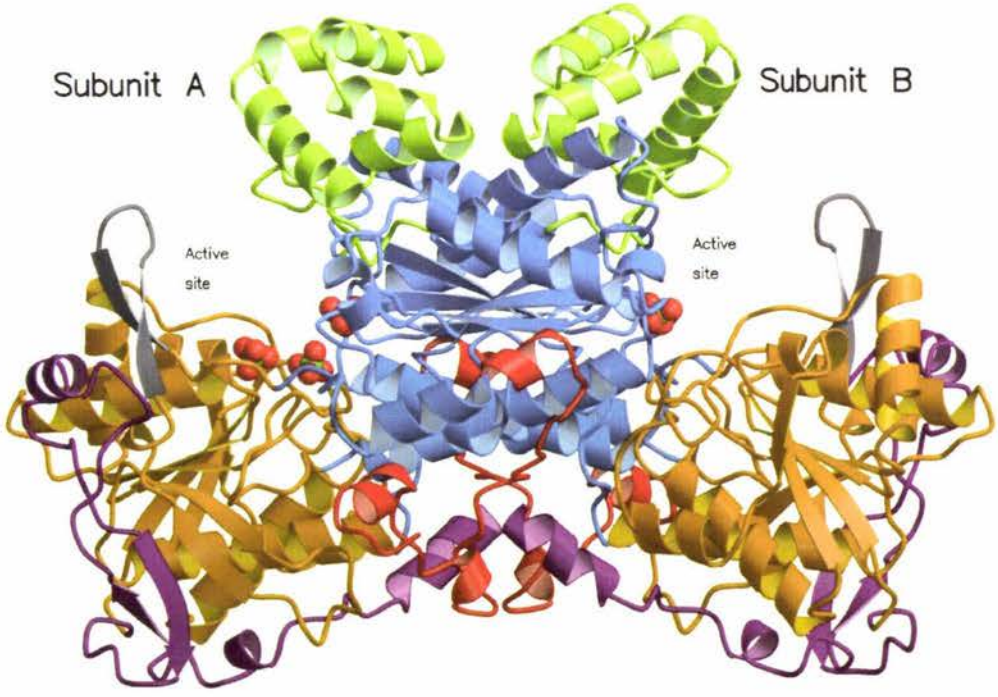
Chapter 4

Spirochaeta thermophilum PP_i-PFK: Structure and Function

4.1 Quaternary Structure

The PP_i-PFK from the thermophilic bacterium *S. thermophilum*, a homologue of the plant PP_i-PFK β-subunit assembles into a homodimer (see Figure 4.1). The 554-residue subunit has a molecular mass of 61080 $g \cdot mol^{-1}$ and a theoretical isoelectric point of 6.22, as calculated by the ProtParam tool (http://kr.expasy.org/tools/protparam.html). The *S. thermophilum* PP_i-PFK 1.85 Å atomic model has a measured solvent accessible area of 38842 $Å^2$. A total surface area of 5353 $Å^2$ is buried at the dimer interface, as calculated by AREAIMOL (CCP4 version 3.5 [Collaborative Computational Project, Number 4. 1994]).

A.



B.

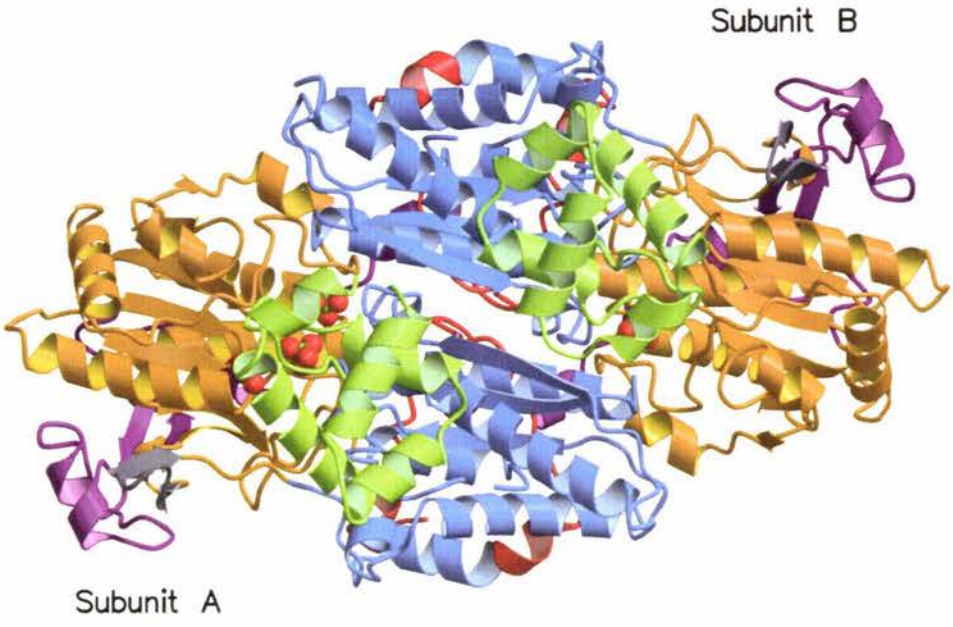


Figure 4.1: Structure of *S. thermophilum* PP_i-PFK. Figures prepared using MOLSCRIPT (Kraulis 1991). **A.** Ribbon diagram of *S. thermophilum* PP_i-PFK. The pseudo two-fold symmetry axis is vertical, parallel to the page, through the centre of the molecule. The small domain is shown in blue and the large domain in orange. The N-terminal insert is shown in purple, the C-terminal insert in red, the β14-β15 insert in grey, and the helical domain in green. **B.** Ribbon diagram of *S. thermophilum* PP_i-PFK looking down the pseudo 2-fold symmetry axis.

4.2 Tertiary Structure

The PP_i-PFK of *S. thermophilum* shares 18.4% sequence identity with the prokaryotic ATP-PFK of *E. coli*. Superposition of the Cα trace of *E. coli* ATP-PFK (1PFK) with subunit A of *S. thermophilum* PP_i-PFK (see Table 4.1a) highlights conservation of the N- and C-terminal domain tertiary structure. Based upon superposition of the N-terminal domain it is evident that the C-terminal domain of *S. thermophilum* PP_i-PFK is displaced relative to *E. coli* ATP-PFK (1PFK) (see Figure 4.3). The consequence of this displacement with respect to the conformation of the F6P/F16bP binding site will be discussed further in later sections of this chapter. The aforementioned and subsequent structural superpositions were based upon backbone atoms (N, Ca, C, and O) and were executed using LSQKAB (CCP4 [Collaborative Computational Project, Number 4] 1994).

	S. thermophilum PP_i-PFK	
	1.85 Å Model	**2.2 Å Model**
N-terminal (N)	0.967 Å	1.001 Å
C-terminal (C)	1.628 Å	1.629 Å
All (N + C)	1.872 Å	1.711 Å

Table 4.1a: Rms difference of backbone atom positions (N, Cα, C, O) based upon superposition of the N- and C-terminal domains of subunit A of *E. coli* ATP-PFK (1PFK) with subunit A of the *S. thermophilum* PP_i-PFK 1.85 Å and 2.2 Å atomic models. Superposition was executed using LSQKAB (CCP4 [Collaborative Computational Project, Number 4] 1994).

	E. coli ATP-PFK	*S. thermophilum* PP_i-PFK
N-terminal domain	5-30	76-101
	35-56	108-129
	79-115	152-188
	119-129	198-208
	255-285	434-464
	139-159	220-240
C-terminal domain	160-190	242-272
	200-220	290-310
	240-253	417-430

Table 4.1b: Residues of the N- and C-terminal domains of *E. coli* ATP-PFK and *S. thermophilum* PP_i-PFK used in the structural superposition which are structurally equivalent.

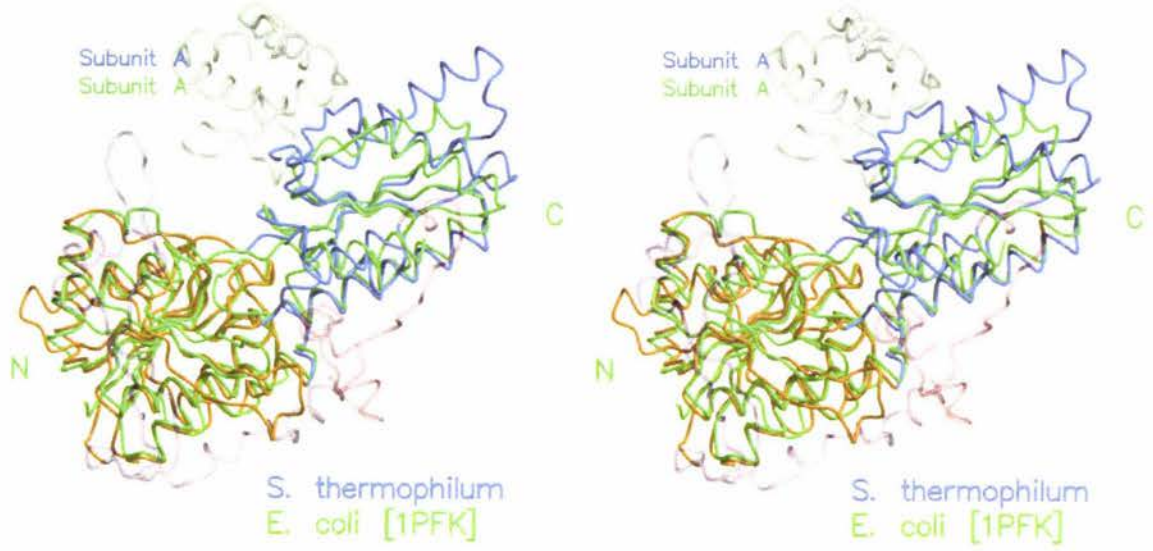
Like *E. coli* ATP-PFK the N-terminal domain exhibits a Rossmann-like fold; a core three layered α/β/α domain of βα repeats (see Figure 4.5b). The βα topology of the Rossmann fold is interrupted by a number of inserts, which *S. thermophilum* PP_i-PFK shares with *B. burgdorferi* PP_i-PFK and other plant PP_i-PFK β-subunit homologues. Relative to the prokaryotic ATP-PFK of *E. coli* (1PFK) *S. thermophilum* PP_i-PFK exhibits three major

insertions, one of which forms a third, autonomous helical domain (see Figure 4.6 a, b, and c) (Moore *et al.* 2002). The N-terminal domain is comprised of residues 1-218 and 436-502. The C-terminal domain is comprised of residues 219-315, 392-435, and 503-554. Residues 316-391 form the helical domain (see Figure 4.6b). The Cα backbone traverses between the N- and C-terminal domains three times, a feature common to the prokaryotic ATP-PFK of *E. coli* (Shirakihara, Y. and Evans, P 1988), *B. stearothermophilus* (Schirmer, T. and Evans, P. 1990), and the PP_i-PFK of *B. burgdorferi* (Moore, S.A. *et al.* 2002). Determination of a PP_i-PFK structure allows for a definitive structure based sequence alignment of PFK A PFKs (see Figure 4.2).

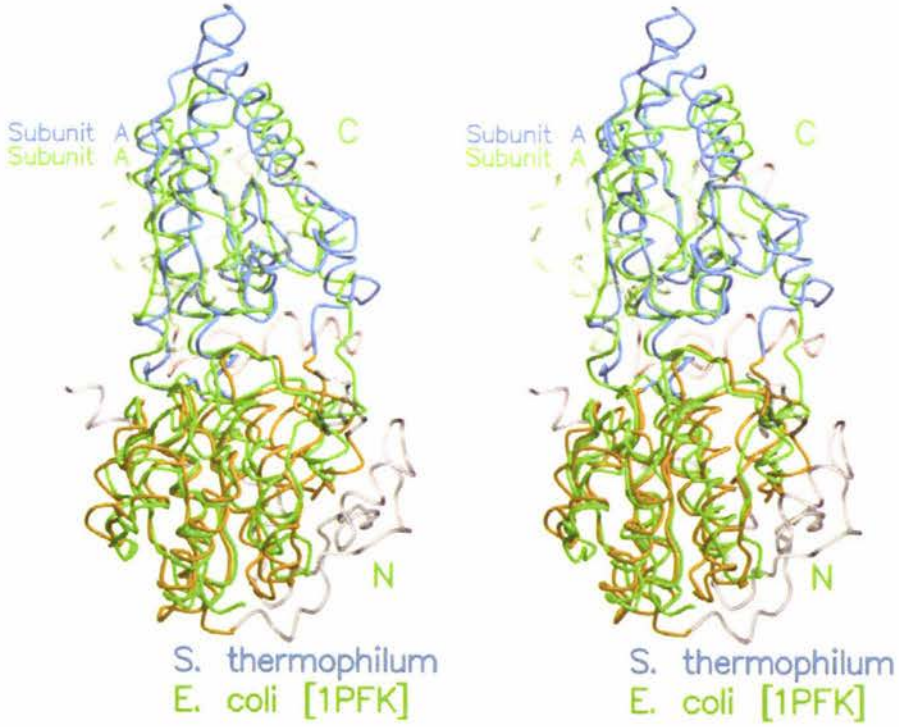
Figure 4.2: (over page) Structure based sequence alignment of the *S. thermophilum*, *B. burgdorferi*, and *E. histolytica* PP_i-PFK with the prokaryotic ATP-PFKs of *E. coli* and *B. stearothermophilus*. Coloring of the respective domains and insertions is consistent with Figure 4.1. Figure produced with ALSCRIPT (Barton 1993).


S.thermophilum
B.burgdorferi
E.histolytica
B.stearothermophilum
E.coli
Allosteric
catalysis
ATP/PPi
Catalysis
F6P
Allosteric
F6P
F6P
F6P

A.


Subunit A
Subunit A
C
N
S. thermophilum
E. coli [1PFK]
Subunit A
Subunit A
C
N
S. thermophilum
E. coli [1PFK]

B.


Subunit A
Subunit A
C
N
S. thermophilum
E. coli [1PFK]
Subunit A
Subunit A
C
N
S. thermophilum
E. coli [1PFK]

C.

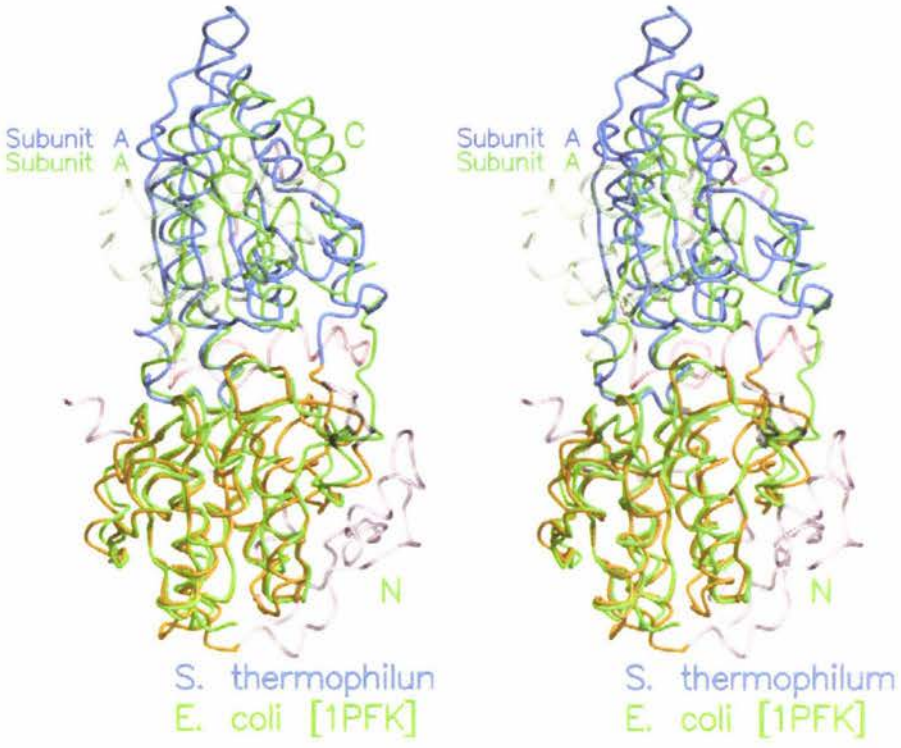


Figure 4.3: Superposition of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model (coloring consistent with Figure 4.1) with that of *E. coli* ATP-PFK (green). **A**. Superposition of the subunit A N- and C-terminal domains of *E. coli* ATP-PFK with those of *S. thermophilum* PP_i-PFK illustrating structural conservation of the domains Rossmann-like fold. Insertions are transparent for clarity. **B**. Same diagram as **A**. viewed approximately down the pseudo dyad symmetry axis. **C**. Superposition of the subunit A N-terminal domain of *E. coli* ATP-PFK with that of *S. thermophilum* PP_i-PFK illustrating the relative rigid-body displacement of the C-terminal domain. Figures prepared using MOLSCRIPT (Kraulis 1991).

4.3 Secondary Structure

The secondary structure of the *S. thermophilum* PP_i-PFK is comprised of 15 β-strands labeled β1-15 and 24 α-helices labeled 1-13, 13a, and 14-23 (see Figure 4.6a), labeled to maintain consistency with the *B. burgdorferi* PP_i-PFK structure (Moore, S.A. *et al.* 2002). Helices 1, 6, 7, 12, 17, 19, and 22 end in a 3_{10} turn, helices 4, 5, and 13 start with a 3_{10} turn, and helices 14 and 15 start and end with a 3_{10} turn. Helix 13a ends with two 3_{10} turns. Corresponding residues, 335A-340A, in *B. burgdorferi* PP_i-PFK exhibit α-helical torsion angles but do not form α-helical hydrogen bonds. Helices 2 and 20 are comprised of a single 3_{10} turn. Helix 21 is kinked, characterized by a 3_{10} turn centered at residue 515. It is feasible that the helix is kinked to maintain a salt bridge, between Arg516 and Glu276 and hydrophobic packing of the C-terminus of helix 22 with strand 9 and helix 10, which would otherwise be solvent accessible. Residues 426-429 have been labeled as helix 18 for consistency with the *B. burgdorferi* PP_i-PFK atomic model but, although they have α-helical torsion angles, they do not form helical hydrogen bonds. Figure 4.4 is a hydrophobicity plot of *S. thermophilum* PP_i-PFK, using the Eisenberg scale (Eisenberg *et al.* 1984) illustrating the degree of hydrophobicity of the secondary structural elements.

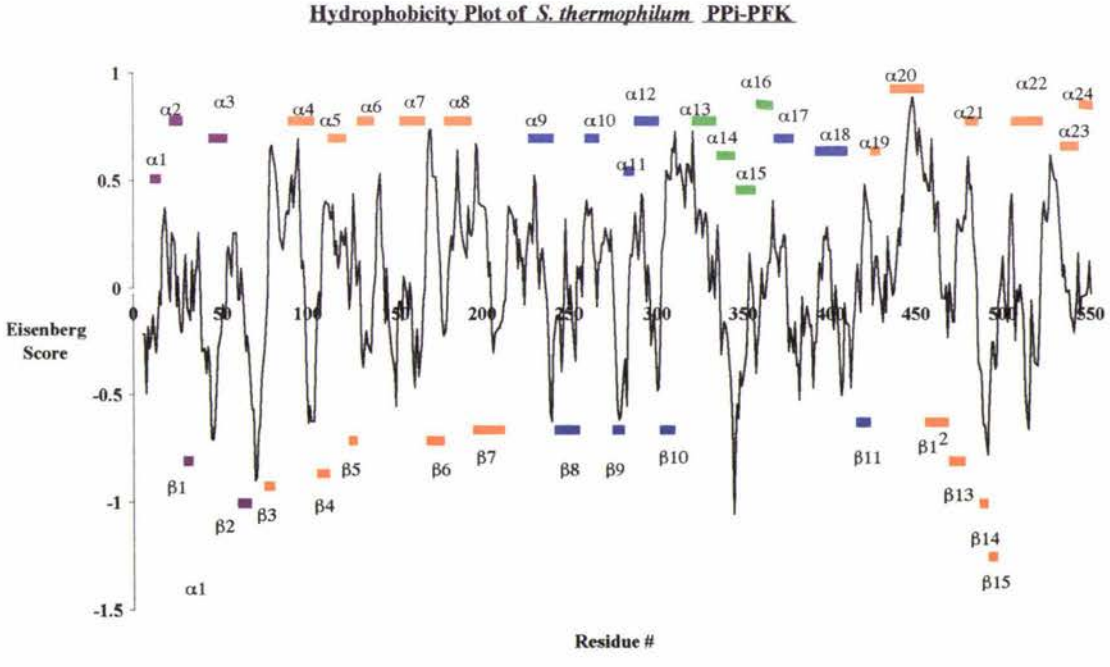


Figure 4.4: Hydrophobicity plot of *S. thermophilum* PP_i-PFK, calculated using the Eisenberg scale (Eisenberg *et al.* 1984). Secondary structural elements are shown. Coloring is consistent with Figure 4.2.

The central β-sheet of the N-terminal domain consists of 9 strands; β3-7, β12, and β13. Strands β14 and β15 form a β-hairpin. The central β-sheet of the C-terminal domain consists of 4 strands, β8-10 and β11 (see Figure 4.5a).

A.

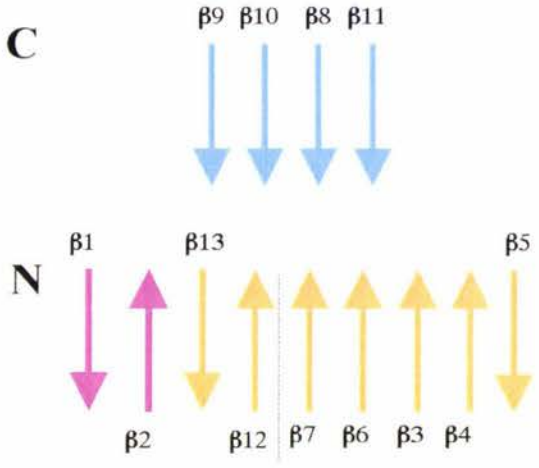


Figure 4.5a: Topology of the N- and C-terminal domain beta-sheets of *S. thermophilum* PP_i-PFK.

B.

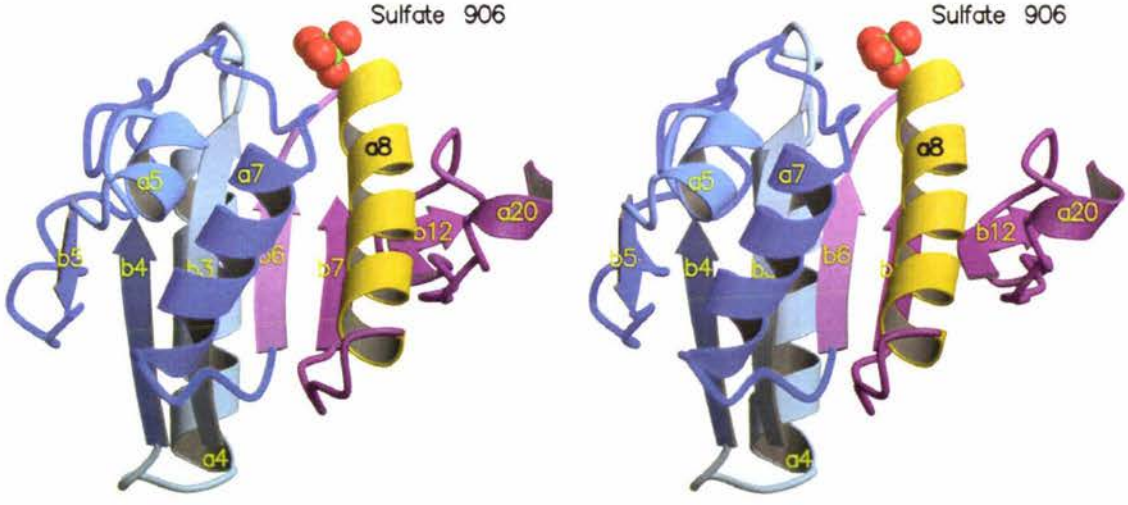
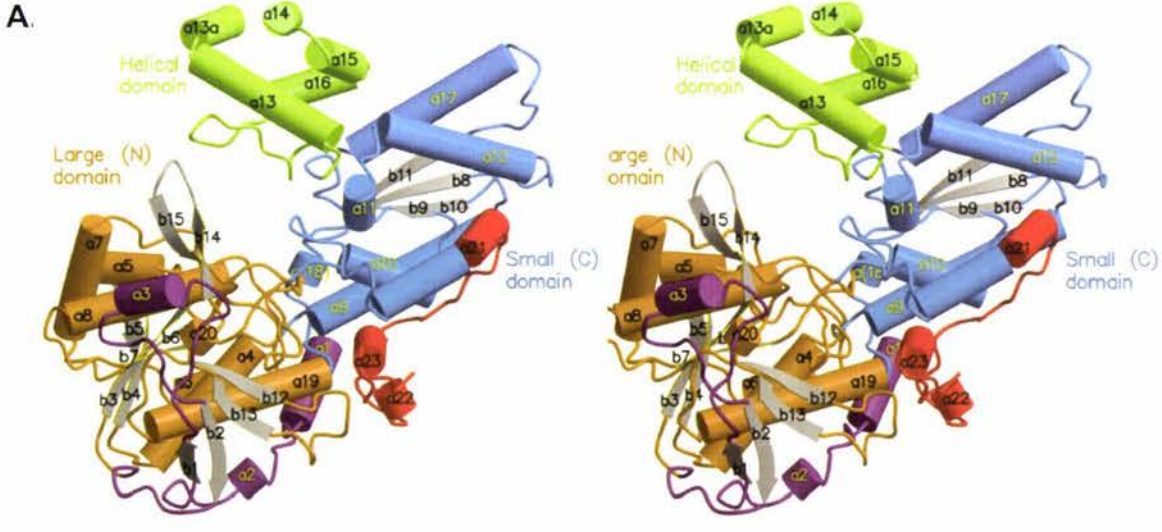


Figure 4.5b: Ribbon diagram of the Rossmann-like fold of the *S. thermophilum* PP_i-PFK N-terminal domain. A sulfate ion is shown bound at the N-terminus of the pyrophosphate-binding α-helix. Compare to the Rossmann-like folds of *E. coli* ATP-PFK and ALDH (see Figure 1.4). Figure prepared using MOLSCRIPT (Kraulis 1991).

4.4 Insertions

S. thermophilum has three major insertions relative to the prokaryotic ATP-PFK of *E. coli* (see Figure 4.6c). An extension of the N-terminus by 70 residues adds 3 α-helices (α1, α2, and α3) and 2 β-sheets (β1 and β2), which extend the N-terminal domain β-sheet to 9 strands. An extension of the C-terminus by 38 residues extends α-helix 13 of *E. coli* ATP-PFK (1PFK) by 6 residues and forms two additional α-helices (α22 and α23). An insertion between β-sheet H and α-helix 9 of *E. coli* ATP-PFK adds 5 helices (α13-17) and a β-hairpin, which form an autonomous third, α-helical domain. There are ten smaller insertions, relative to the ATP-PFK of *E. coli*. These include the addition of α11, β14, β15, the extension of α8 and the α8-β7 loop, α17 and the α17-β11 loop, the β12-β13 loop, and four single or double insertion's; a double insertion in the α1-B loop (α4-β4), a single insertion in the α(3)-α4a loop (α6-α7), a double insertion in the E-α6 loop (β7-α9), and a single insertion in α6 (α9-β8) (see Figure 4.2).

The sequence alignment of the *S. thermophilum* and *B. burgdorferi* PP_i-PFKs is virtually continuous (see Figure 4.2). The N- and C-termini of *S. thermophilum* PP_i-PFK are extended by a single residue and a single residue insertion is present in the α13a-α14 loop of the helical domain. *B. burgdorferi* PP_i-PFK has a single residue insertion in the β2-β3 loop and a three-residue insertion, which extends the 380-390 β-hairpin.



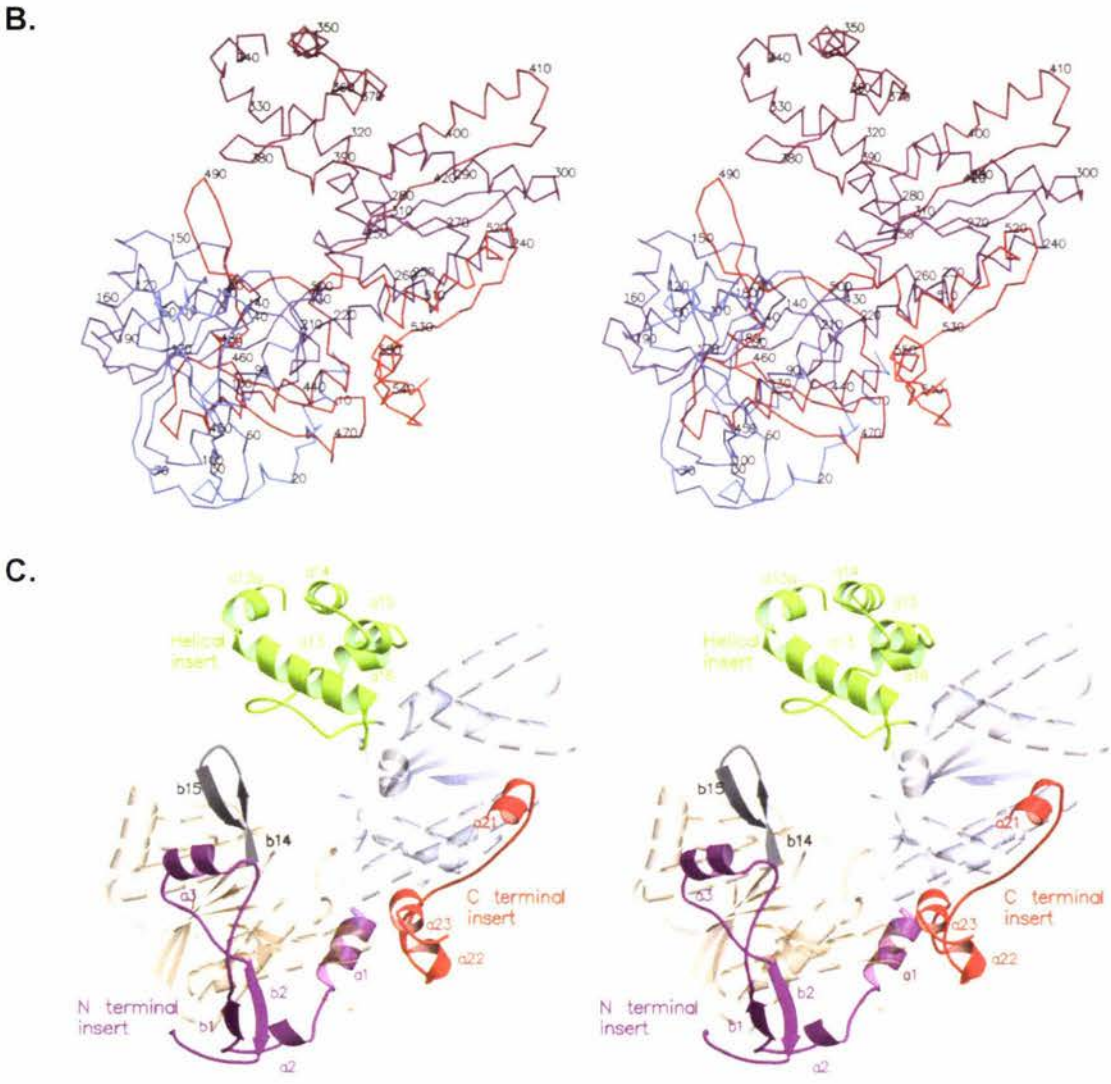


Figure 4.6: Secondary structural elements of *S. thermophilum* PP_i-PFK.

A. A stereo diagram of subunit A of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model. Secondary structural elements are labelled. Regions homologous to the large and small domain of *E. coli* ATP-PFK are coloured blue and orange, respectively (purple = N terminal insert, red = C terminal insert, and green = autonomous helical domain insert).

B. A stereo Cα diagram of subunit A of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model in an identical orientation as **A**. The colouring is ramped from the N-terminus (blue) to the C-terminus (red). Every tenth residue is labelled.

C. A stereo diagram of subunit A of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model identifying the major insertions relative to the prokaryotic ATP-PFK of *E. coli* (helical insert = green, N terminal insert = purple, C terminal insert = red, and β14-β15 insert grey). Regions homologous to the large and small domain of *E. coli* ATP-PFK are χολουρεδ blue and orange, respectively and are transparent. All diagrams (**A-C**) were prepared using MOLSCRIPT (Kraulis 1991) and Raster3D (Merrit and Bacon 1997).

4.5 Structural Asymmetry

Despite structural conservation of the individual domains the tertiary structure of the *B. burgdorferi* PP_i-PFK subunits is asymmetrical (Moore *et al.* 2002). Superposition of the A and B subunits based upon regions of the N-terminal domain homologous to *E. coli* ATP-PFK yields an rms displacement for all main chain atoms of 2.46 Å, calculated using DIFRES (CCP4 version 3.5 [Collaborative Computational Project, Number 4] 1994) (see Table 4.2b). Relative to subunit B, based upon superposition of the N-terminal domain, the helical and C-terminal domains of subunit A have undergone independent rigid body displacements towards the active site. The N-terminal domain β14-β15 β-hairpin of subunit A has also shifted towards the active site with a concomitant shift of helices α3 and α8 (see Figure 4.8a). Resultant of the asymmetry is closure of the active site of subunit A whilst that of subunit B remains open. Displacement of the helical domain into the active site results in interactions between residues of the helical domain 380-390 β-hairpin with conserved residues (GGD**D**) involved in substrate binding and catalysis (Moore *et al.* 2002). The significance of the aforementioned asymmetry will be discussed in later sections of this chapter with respect to the *S. thermophilum* PP_i-PFK atomic models and catalysis/substrate binding. Both *S. thermophilum* PP_i-PFK atomic models exhibit a relative 'open' and 'closed' subunit asymmetry, to a lesser degree than the *B. burgdorferi* structure (1KZH). The subunit structural asymmetry of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model can be approximately characterized as displacement of the C-terminal and helical domains of subunit A as a rigid body, relative to superposition of the N-terminal domain (see Figure 4.8b). Displacement of the C-terminal and helical domains relative to the N-terminal domain involves a rotation towards the active site about an axis (z-axis) through G219A, perpendicular to the pseudo 2-fold symmetry by approximately 4° concomitant with a rotation towards the dimer interface about an axis (y axis) through G219A, parallel to the pseudo 2-fold symmetry axis by 3°. Superposition of the respective domains illustrates that the structure of the N- and C-terminal domains is conserved between the *S. thermophilum* PP_i-PFK 1.85 Å and *B, burgdorferi* PP_i-PFK atomic models (see Table 4.2a). Superposition of the N- and C-terminal domains concomitantly illustrates that their spatial orientation is also conserved, hence the 'open' and 'closed' asymmetry of the *S. thermophilum* 1.85 Å atomic model is a

close approximation of the *B. burgdorferi* PP_i-PFK with the exception of the helical domain, which is not displaced into the active site in the *S. thermophilum* PP_i-PFK 1.85 Å structure (see Table 4.3, Figure 4.8d).

PP_i-PFK		Domain N (all)	N (core)	C	α-Helical
B. burgdorferi 1KZH	Δ rms Å	0.716	0.208	0.348	0.972
S. thermophilum 1.85 Å model	Δ rms Å	0.598	0.216	0.345	0.926
S. thermophilum 2.2 Å model	Δ rms Å	0.306	0.105	0.197	0.527

Table 4.2a: Independent superposition of the N-terminal, C-terminal, and helical domains for both subunits of the dimer for *B. burgdorferi* PP_i-PFK, *S. thermophilum* PP_i-PFK 1.85 Å, and *S. thermophilum* PP_i-PFK 2.2 Å atomic models. N (core) is defined as the core of the domain that is structurally equivalent to that of *E. coli* ATP-PFK (see Table 4.1b). Domain structure within the relative models is highly conserved with the exception of the helical domain. Superposition was executed using LSQKAB (CCP4 [Collaborative Computational Project, Number 4] 1994).

	Superposition	*B. burgdorferi*	*S. thermophilum* 1.85 A Model	*S. thermophilum* 2.2 A Model
rms difference between subunits for all residues	all	1.978 Å	0.878 Å	0.922 Å
	[N-core]	2.46 Å [0.208 Å]	1.64 Å [0.216 Å]	1.17 Å [0.105 Å]

Table 4.2b: Comparison of the asymmetry between subunits of the three structures. Rms difference between subunits of the respective atomic models based upon superposition of all residues or regions of the N-terminal domain homologous to that of *E. coli* ATP-PFK (N [core]). The *B. burgdorferi* PP_i-PFK exhibits the greatest asymmetry. Superposition was executed using LSQKAB (CCP4 [Collaborative Computational Project, Number 4] 1994).

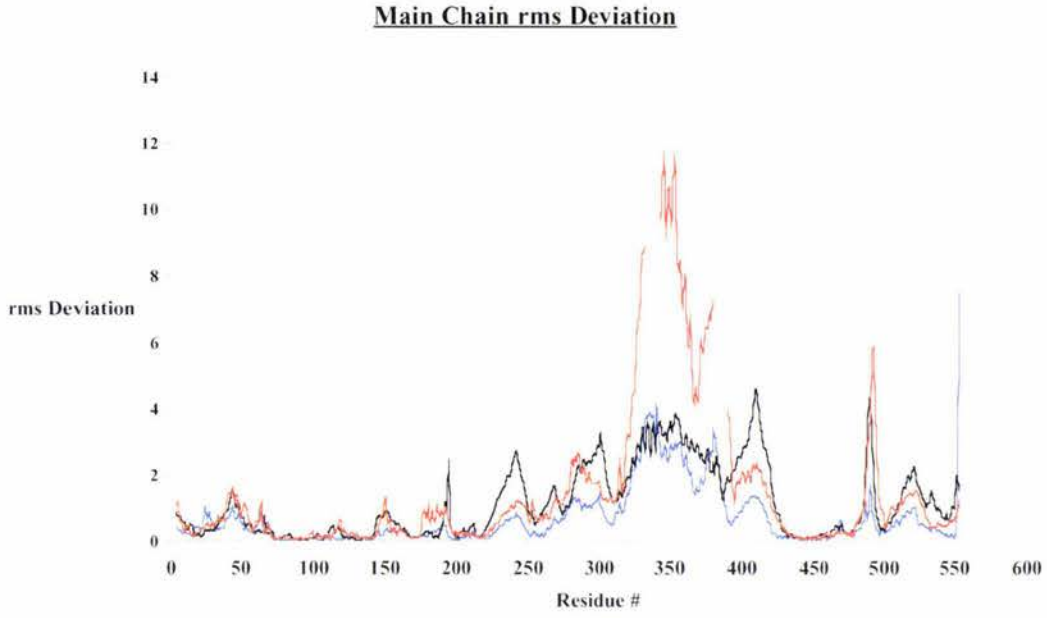


Figure 4.7a: Graphical representation of the rms deviation (Å) between subunits A and B based upon superposition of the large domain (70-147, 97-210, 432-462) for the three PP_i-PFK atomic models; *B. burgdorferi* PP_i-PFK (red), *S. thermophilum* PP_i-PFK 2.2 Å (blue), and *S. thermophilum* PP_i-PFK 1.85 Å (black).

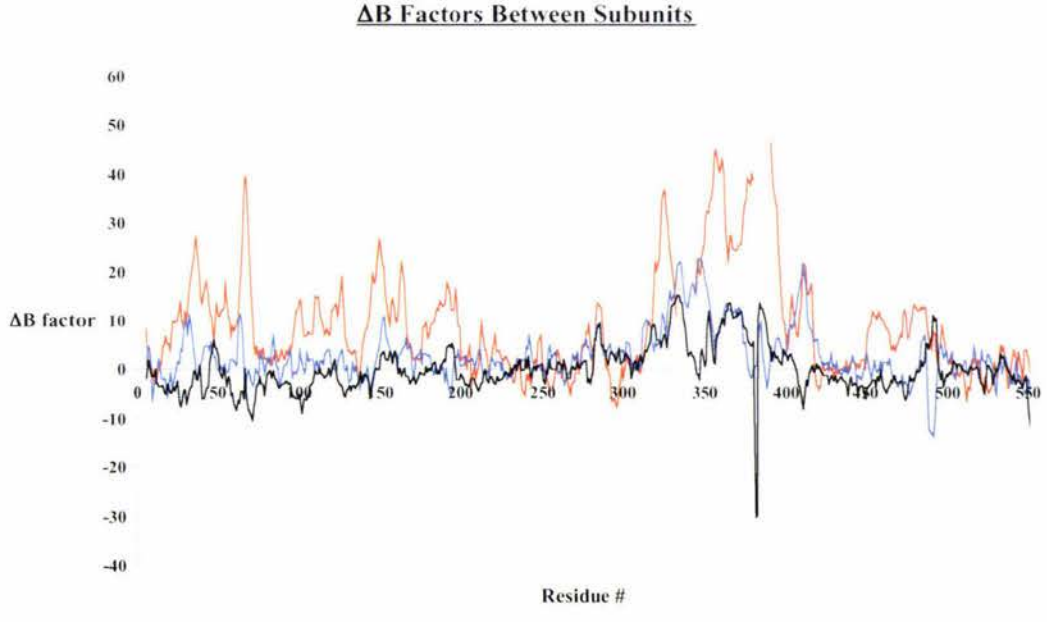


Figure 4.7b: Graphical representation of the difference in B factor between subunits A and B of the three PP_i-PFK atomic models; *B. burgdorferi* PP_i-PFK (red), *S. thermophilum* PP_i-PFK 2.2 Å (blue), and *S. thermophilum* PP_i-PFK 1.85 Å (black).

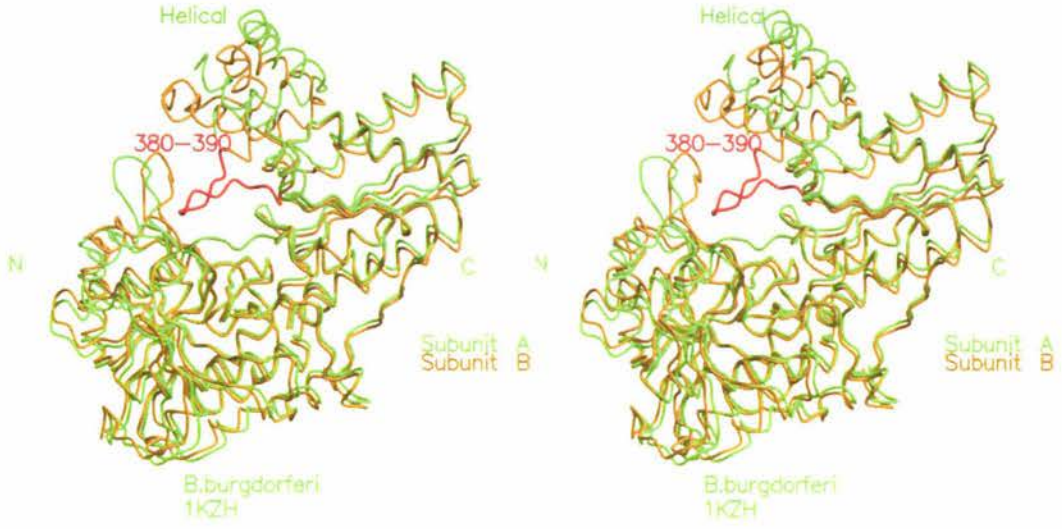
Domain	*S. thermophilum* 1.85 Å Structure Subunit	*B. burgdorferi* 1kzh Subunit	rms displacement Å
All			
(5-550)	A	A	2.056
	A	B	1.857
	B	A	2.16
	B	B	1.766
N-core			
70-147	A	A	0.601
197-210	A	B	0.599
432-462	B	A	0.577
	B	B	0.561
N-all			
5-210	A	A	1.503
435-500	A	B	1.22
	B	A	1.211
	B	B	1.115
C			
220-310	A	A	0.403
392-432	A	B	0.429
	B	A	0.514
	B	B	0.559
N-core + C			
	A	A	0.532
	A	B	0.684
	B	A	0.705
	B	B	0.656
Helical			
320-342	A	A	1.214
344-380	A	B	1.082
	B	A	1.138
	B	B	0.995

Table 4.3: Structural comparison of the *S. thermophilum* PP_i-PFK 1.85Å structure with the *B. burgdorferi* PP_i-PFK structure based upon superposition of the respective domains. Superposition was executed using LSQKAB (CCP4 [Collaborative Computational Project, Number 4] 1994).

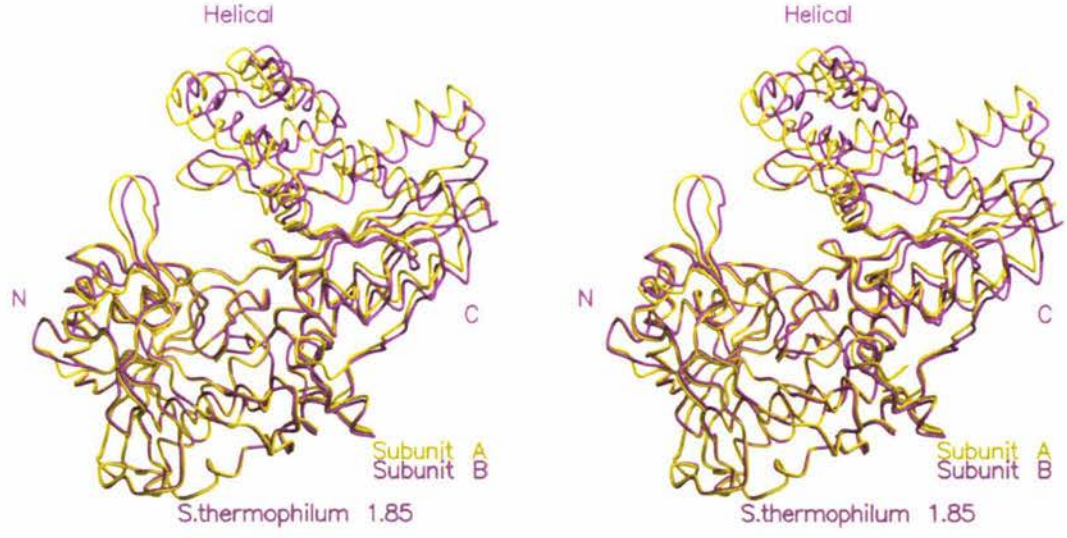
Domain	1.85 Å Structure Subunit	2.2 Å Structure Subunit	rms displacement Å
ALL			
5-550	A	A	1.8
	A	B	2.503
	B	A	1.485
	B	B	2.204
N-core			
70-147	A	A	0.239
197-210	A	B	0.27
432-462	B	A	0.143
	B	B	0.161
N-all			
5-210	A	A	0.561
435-500	A	B	0.641
	B	A	0.316
	B	B	0.382
C			
220-310	A	A	0.433
392-432	A	B	0.438
	B	A	0.296
	B	B	0.32
Helical			
320-342	A	A	0.512
344-380	A	B	0.438
	B	A	0.876
	B	B	0.984

Table 4.4: Structural comparison of subunits A and B from the 2.2 Å and 1.85 Å S. thermophilum PP_i-PFK structures. The data indicates that both subunits of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model bares a closer structural resemblance to subunit B of the 1.85 Å atomic model. Superposition was executed using LSQKAB (CCP4 [Collaborative Computational Project, Number 4] 1994).

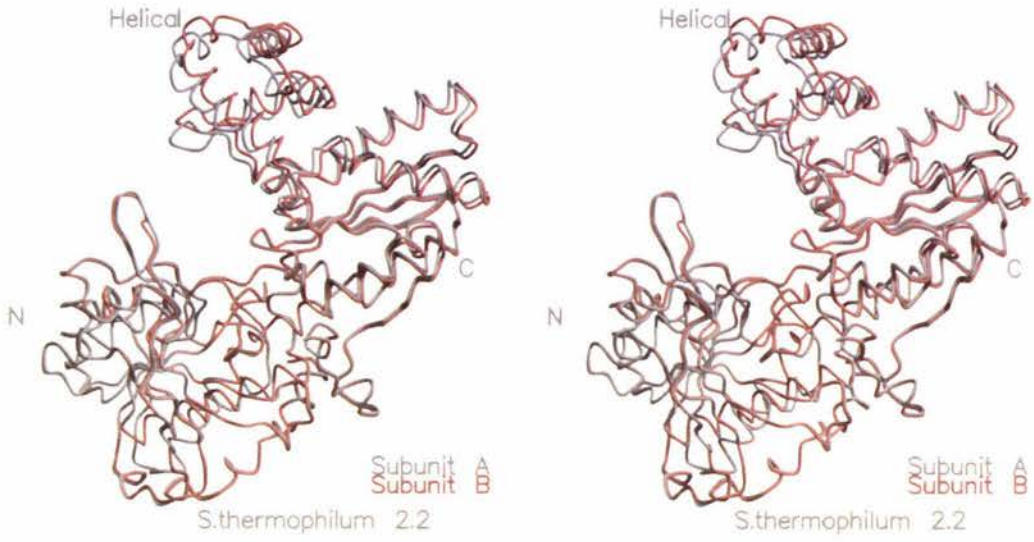
A.


Helical
380–390
N
C
Subunit A
Subunit B
B.burgdorferi
1KZH
Helical
380–390
N
C
Subunit A
Subunit B
B.burgdorferi
1KZH

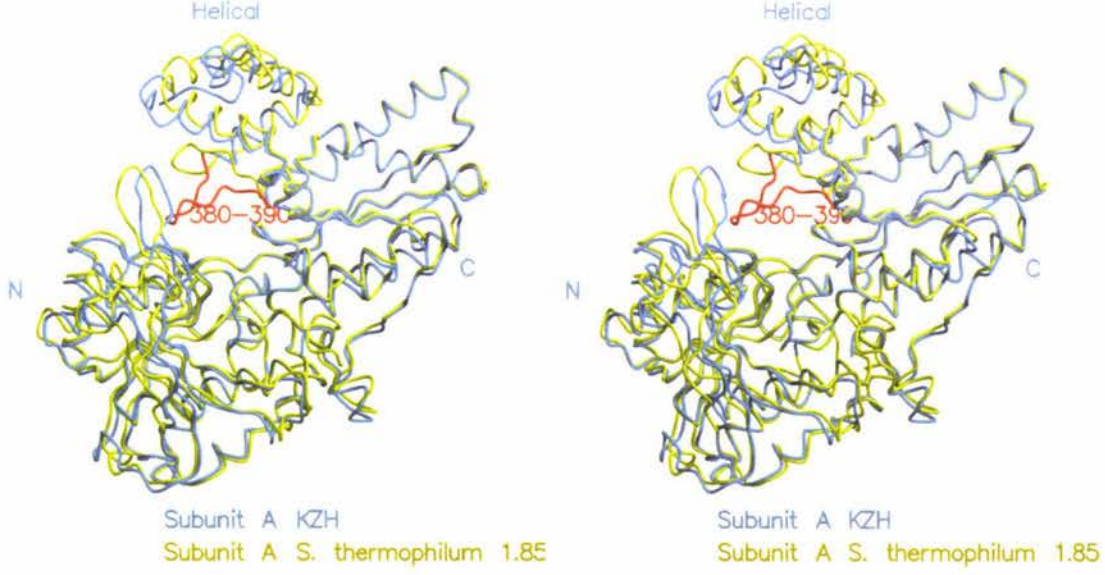
B.


Helical
N
C
Subunit A
Subunit B
S.thermophilum 1.85
Helical
N
C
Subunit A
Subunit B
S.thermophilum 1.85

C.


Helical
N
C
Subunit A
Subunit B
S.thermophilum 2.2
Helical
N
C
Subunit A
Subunit B
S.thermophilum 2.2

D.


Helical
380–390
N
C
Subunit A KZH
Subunit A S. thermophilum 1.85
Helical
380–390
N
C
Subunit A KZH
Subunit A S. thermophilum 1.85

E.

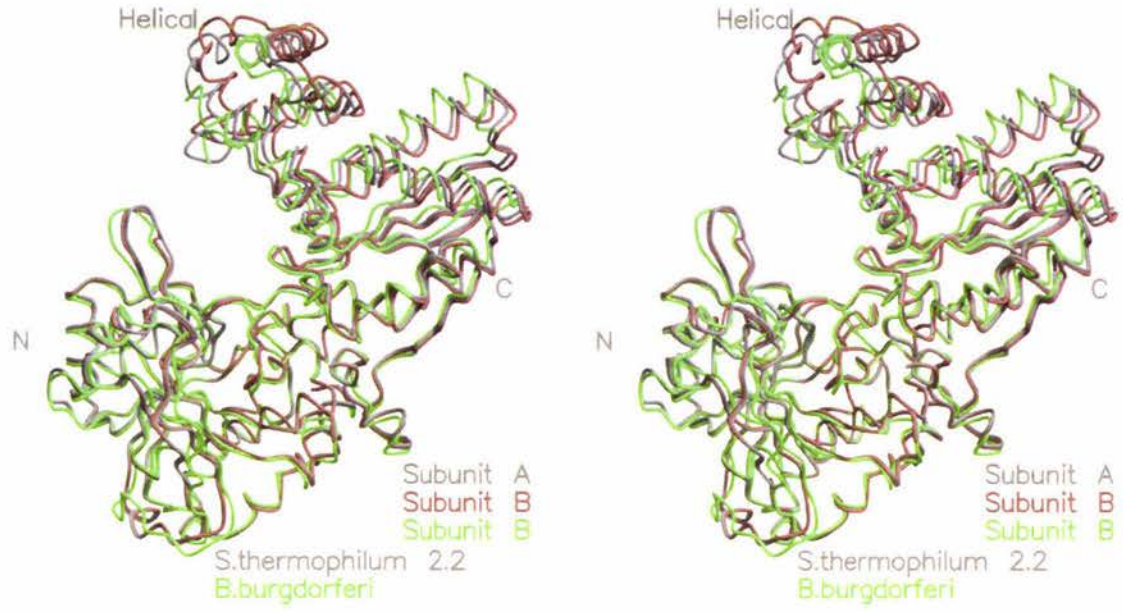


Figure 4.8: Structural subunit asymmetry of *B. burgdorferi* (1KZH[Moore *et al.* 2002]) and *S. thermophilum* PP_i-PFKs. Figures prepared using MOLSCRIPT (Kraulis 1991) and Raster3D (Merrit and Bacon 1997).

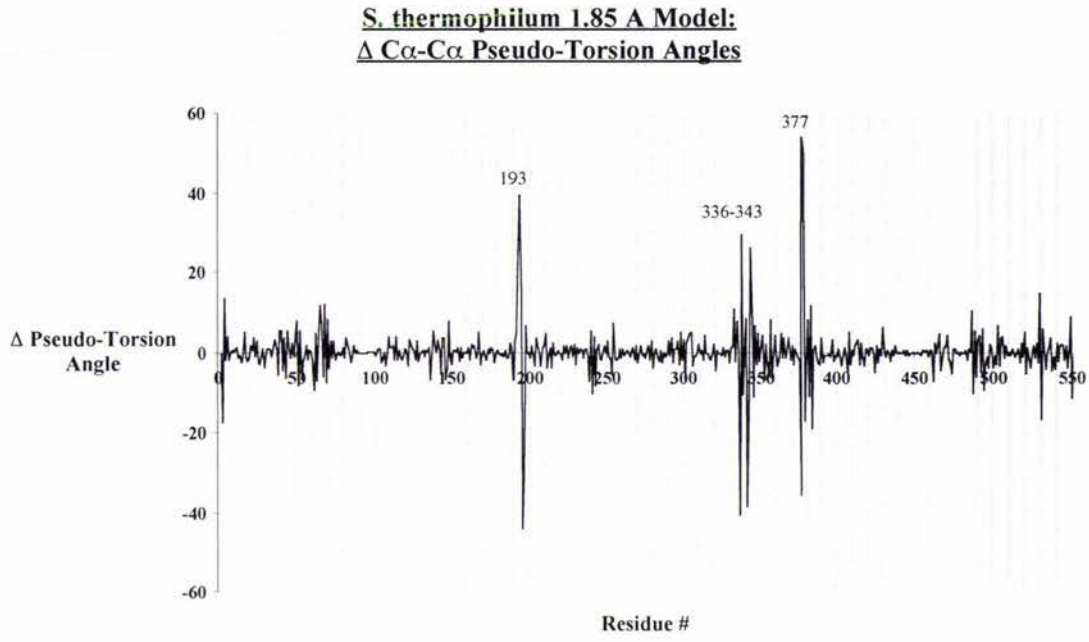
A. Superposition of subunits A (orange) and B (red) of the *B. burgdorferi* PP_i-PFK based upon the N-terminal domain illustrating the 'open' and 'closed' asymmetry of the subunits. **B**. Superposition of subunits A (yellow) and B (purple) of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model based upon regions identical to **A.** illustrating subunit asymmetry. **C.** Superposition of subunits A (grey) and B (salmon) of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model based upon regions identical to **A.** illustrating subunit asymmetry. D. Superposition of subunit A (*S. thermophilum* PP_i-PFK 1.85 Å [yellow]) with subunit A of *B. burgdorferi* PP_i-PFK (blue) based upon regions identical to **A. E**. Superposition of subunits A (grey) and B (salmon) of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model with the 'open' subunit, B, of *B. burgdorferi* PPi-PFK (green) based upon regions identical to **A.** illustrating that both subunits of *S. thermophilum* are in a 'open' conformation.

Neither subunit of the *S. thermophilum* 1.85 Å atomic model is closed as that of its *B. burgdorferi* PP_i-PFK counterpart due to the relative conformation of the 380-390 and β14-β15 β-hairpins into the active site (see Figure 4.8d). Although the *S. thermophilum* PP_i-PFK 2.2 Å atomic model exhibits an 'open' and 'closed' subunit structural asymmetry (see Figure 4.8c), relative to the *S. thermophilum* PP_i-PFK 1.85 Å and *B. burgdorferi* PP_i-PFK atomic models both of its subunits are in an 'open' conformation (see Figure 4.8e). The subunit structural asymmetry of the *S. thermophilum* PP_i-PFK 2.2 Å differs from that of the 1.85 Å atomic model. Relative to superposition of the N-terminal domain the subunit asymmetry can be approximately characterized as independent rigid-body displacements of the C-terminal and helical domains of subunit A relative to subunit B (see Table 4.2a). The displacement of the C-terminal domain relative to the N-terminal domain involves a rotation about the z-axis by approximately 2.5°. The displacement of the helical domain is not a simple rotation. The helical domain is rotated about an axis through the Cα of Glu393, parallel to the z-axis by approximately 4° concomitant with a rotation about an axis through the Cα of Glu393, parallel with the y-axis by approximately 3°. It is important to note that the subunit asymmetry of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model involves a rotation of the C-terminal domain, relative to the N-terminal domain about the y-axis, towards the dimer interface. This is not observed in the *S. thermophilum* PP_i-PFK 2.2 Å atomic model. The repercussions of this with respect to the dimer interface will be discussed further in section 4.7. The aforementioned displacement of the respective domains is relatively small. Deviation in the displacement of the small domain is concluded to be significant as it results in the altered orientation of the MGR motif into the active site (discussed further in section 4.9). Deviation in the displacement of the respective helical domains is concluded to be a result of random conformational flexibility, as demonstrated by its disorder i.e. high B factor and low real space correlation coefficients (see Figure 4.7b and Figure 3.5). See Table 4.3 for helical domain rms differences between the three structures.

It is apparent from superposition of the respective domains of the 1.85 Å atomic model that the C-terminal domain of subunit A is shifting as a rigid-body (see Table 4.2a). A plot of the difference in the Cα-Cα pseudo-torsion angles (defined for residue i by the 4

atoms Cα [i-1] to Cα [i+2]) between subunits A and B of the 1.85 Å structure (standard deviation = 6.518°) does not identify a distinct hinge movement at the interface of the N- and C-terminal domains, which would be indicated by a peak in the Cα-Cα pseudo-torsion angle plot. The distinct hinge movements observed at Gly193 (α8-β7 surface loop [see Figure 4.9a]) is a direct result of asymmetric crystal contacts and is unlikely to be physiologically relevant. It is therefore concluded that conformational variance of the C-terminal domain does not arise from a distinct hinge movement. The peak observed at Met377 is representative of a hinge movement of the 380-390 β-hairpin of subunit A, relative to that of subunit B. Distinct hinge movements are also observed in the region of 336-343, and Met377 of the helical domain (see Figure 4.9a). Similar hinge movements within the helical domain are observed in a ΔCα-Cα pseudo-torsion angle plot for the *S. thermophilum* PP_i-PFK 2.2 Å atomic model (structure (standard deviation = 5.189°) (see Figure 4.9b). Superposition of the helical domains from the *S. thermophilum* PP_i-PFK 1.85 Å and 2.2 Å atomic models (see Table 4.4) indicates that their conformation differs to a slightly larger extent than the N- and C-terminal domains. The helical domain of subunit B of both *S. thermophilum* PP_i-PFK atomic models is disordered with respect to the A subunit, as illustrated by a plot of real-space correlation coefficients and average residue B factors (see Figure 3.5), particularly in the region of α13 and α14. In subunit A of the 2.2 Å atomic model α13 is comprised of a single helical turn followed by a single 3_{10} turn. In subunit B however α13 is comprised of a single 3_{10} turn. Based upon positive difference density it appears that residues 342B and 343B, which form the turn between α13 and α14, have alternative conformations. This indicates that this region in subunit B is relatively free to move within the crystal, which can be attributed to asymmetric crystal contacts (see Table 3.7). The helical domain of subunit B of *B. burgdorferi* PP_i-PFK, like that of the *S. thermophilum* PP_i-PFK atomic models is disordered relative to the rest of the molecule preventing the complete tracing of the polypeptide (residues 333-343 and 380-390 are missing [Moore *et al.* 2002]).

A.



B.

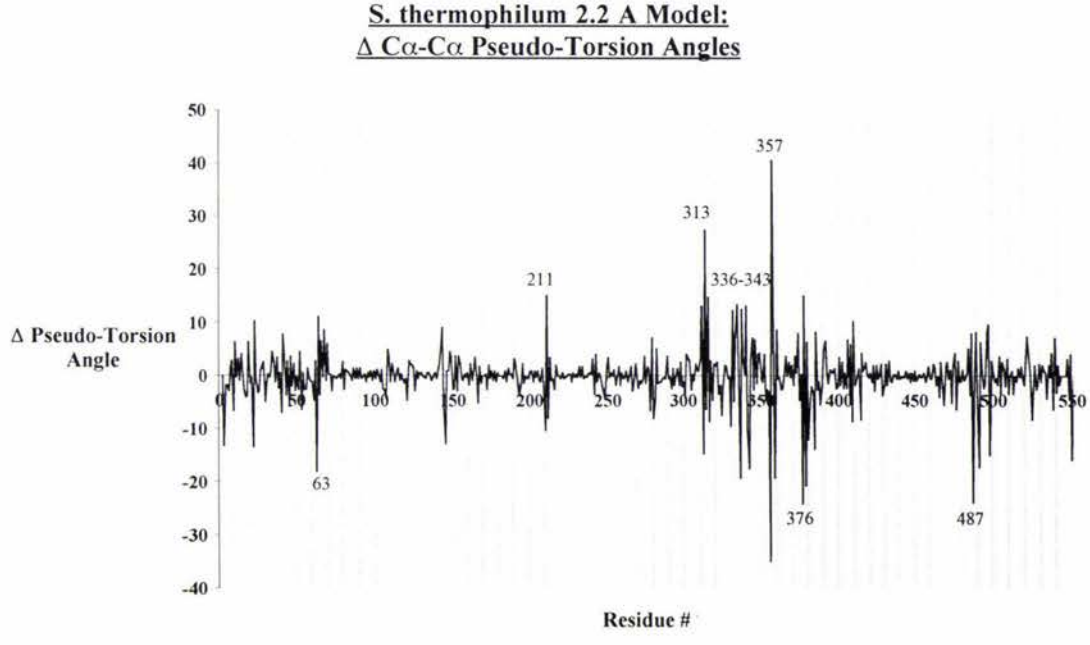


Figure 4.9: Differences in the Cα-Cα pseudo-torsion angles between subunits A and B of the respective S. thermophilum PP_i-PFK atomic models (**A** 1.85 Å model, **B** 2.2 Å model). A peak indicates a local hinge movement.

Note in the ΔCα-Cα pseudo-torsion angle plot of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model the absence of peaks at Glu487, Asn211, and Gly313, relative to the 2.2 Å plot (see Figure 4.9b). Neither Asn211 nor Gly313 is positioned upon a surface loop and therefore can not be dismissed directly as a result of asymmetric crystal contacts. The peak at Asn211, though at the interface of the N- and C-terminal domains is not representative of a domain hinge movement, rather of a local conformational deviation between subunits resultant from the distinct orientation of the C-terminal alpha helix, α23, which differs between subunits due to asymmetric crystal contacts. The peak at Gly313 is representative of a hinge movement between the C-terminal and helical domains. The absence of this peak in the 1.85 Å plot validates the prior conclusion that subunit asymmetry of the respective *S. thermophilum* PP_i-PFK atomic models is distinct. It also indicates that subunit asymmetry of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model involves displacement of the C-terminal and helical domains as a rigid body.

4.6 Conformational change

Based upon superposition of the N-terminal domain core (residues 70-147, 197-210, 432-462) the conformational change of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model, relative to that of the 2.2 Å model entails a rigid-body displacement of the C-terminal and helical domains of both subunits. The displacement of the subunit A helical domain can be characterized as a rotation around the y-axis, towards the dimer interface by approximately 6.5° concomitant with a rotation about the z-axis, towards the active site by approximately 4°. The helical domain is further rotated towards the active site around an axis through the Cα of Glu393A, parallel to the z-axis by approximately 8°(see Figure 4.10a and b). The conformational change of the *S. thermophilum* PP_i-PFK 1.85Å B subunit, relative to that of the 2.2 Å atomic model entails a rigid-body displacement of the C-terminal domain, which can be characterized as a rotation about the z-axis, towards the active site by approximately 4°. The helical domain is further rotated towards the active site around an axis through the Cα of Glu393A, parallel to the z-axis by approximately 11° (see Figure 4.10c and d). The conformational change of the C-terminal domain of subunit B relative to the N-terminal domain does not involve a rotation about

the y-axis, towards the dimer interface. The conformational change illustrated by the *S. thermophilum* PP_i-PFK 1.85 Å and 2.2 Å atomic models has significant repercussions upon the conformation of the F6P/F16bP 6-phosphate binding site (see section 4.9.1). Surprisingly, unlike prokaryotic ATP-PFKs the repercussions of the conformational change are not elicited via deviation of the dimer interface. Superposition of the 1.85 Å and 2.2 Å *S. thermophilum* PP_i-PFK structures based upon the N-terminal domain concomitant with the C-terminal domain of the opposing subunit illustrates that the orientation of the N-terminal domain relative to the C-terminal domain of the opposing subunit is conserved (see Table 4.5a, Figure 4.11a). How is this possible in view of the conformational change? It is possible via the conservation of contacts at the dimer interface.

Domain Superimposed	rms displacement Å
B Subunits N-core and A Subunits C-domain	0.438
A Subunits N-core and B Subunits C-domain	0.433

Table 4.5a: Structural comparison between the 1.85Å and 2.2Å *S. thermophilum* PP_i-PFK structures reveals that the orientation of the C-terminal domain relative to the N-terminal domain of the opposing subunit in the dimer is highly conserved despite the conformational change. Superposition was executed using LSQKAB from the CCP4 version 3.5 suite of programs (Collaborative Computational Project, Number 4, 1994).

Domains Superimposed	rms displacement Å
B Subunits N-core and A subunits C-domain	0.791
A Subunits N-core and B Subunits C-domain	0.747

Table 4.5b: Structural comparison between the *S. thermophilum* PP_i-PFK 1.85 Å and *B. burgdorferi* PP_i-PFK atomic models reveals that the orientation of the C-terminal domain relative to the N-terminal domain of the opposing subunit in the dimer is also relatively conserved. Superposition was executed using LSQKAB from the CCP4 version 3.5 suite of programs (Collaborative Computational Project, Number 4, 1994).

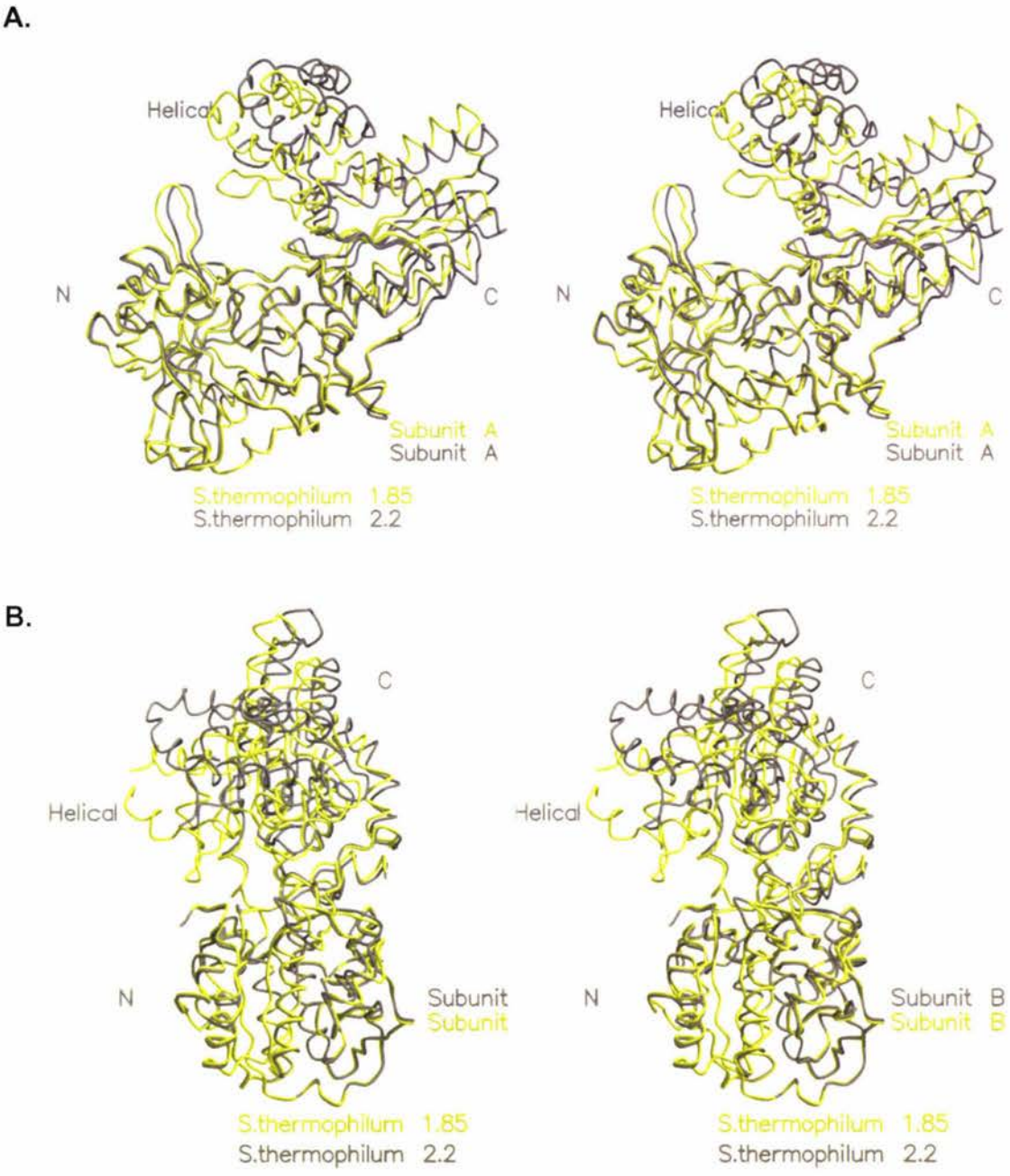
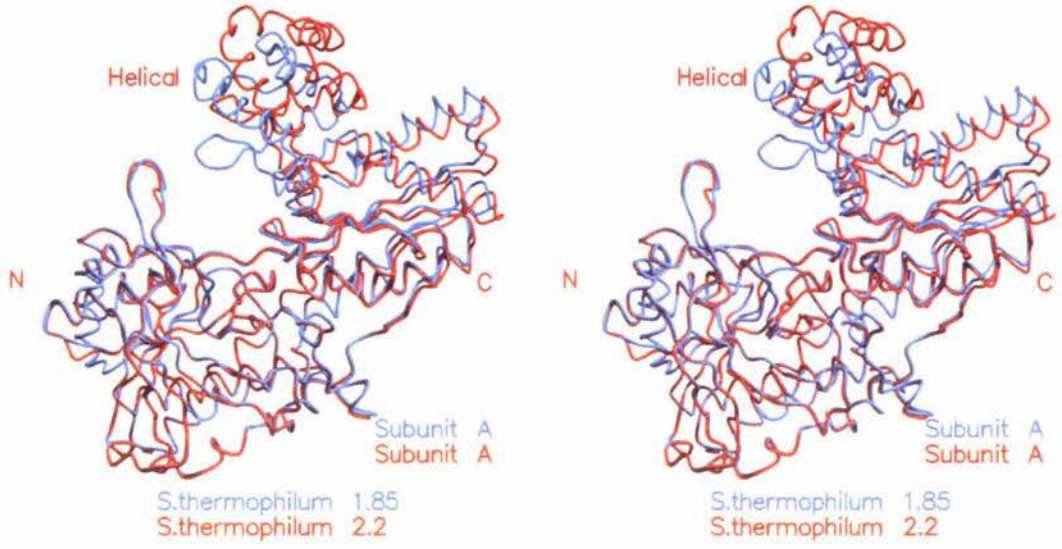


Figure 4.10: Conformational change of the PP_i-PFK from the thermophilic bacterium *S. thermophilum*. Figures prepared using MOLSCRIPT (Kraulis 1991).

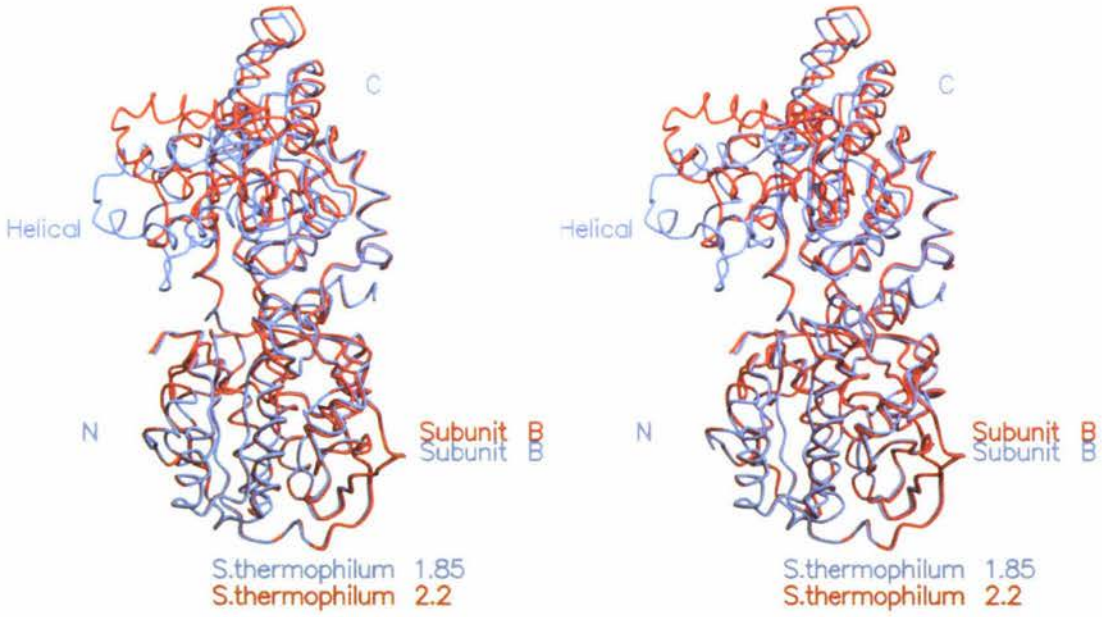
A. Superposition of subunit A of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model (yellow) with subunit A of the 2.2 Å atomic model (blue) based upon regions of the N-terminal domain homologous to the N-terminal domian of *E. coli* ATP-PFK (1PFK [Shirakihara and Evans 1988]).

B. Superposition of subunit A of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model (yellow) with subunit A of the 2.2 Å atomic model (blue) identical to **A.**, viewed down the pseudo 2-fold symmetry axis.

C.



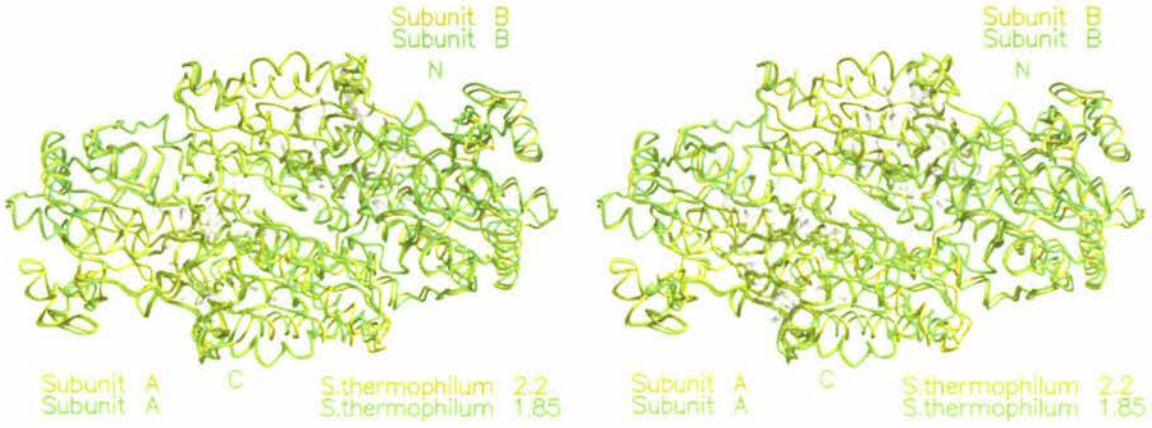
D.



C. Superposition of subunit B of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model (blue) with subunit B of the 2.2 Å atomic model (red) based upon regions identical to **A.**

D. Superposition of subunit B of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model (blue) with subunit A of the 2.2 Å atomic model (red) identical to **C.** viewed down the pseudo 2-fold symmetry axis.

A.



B.

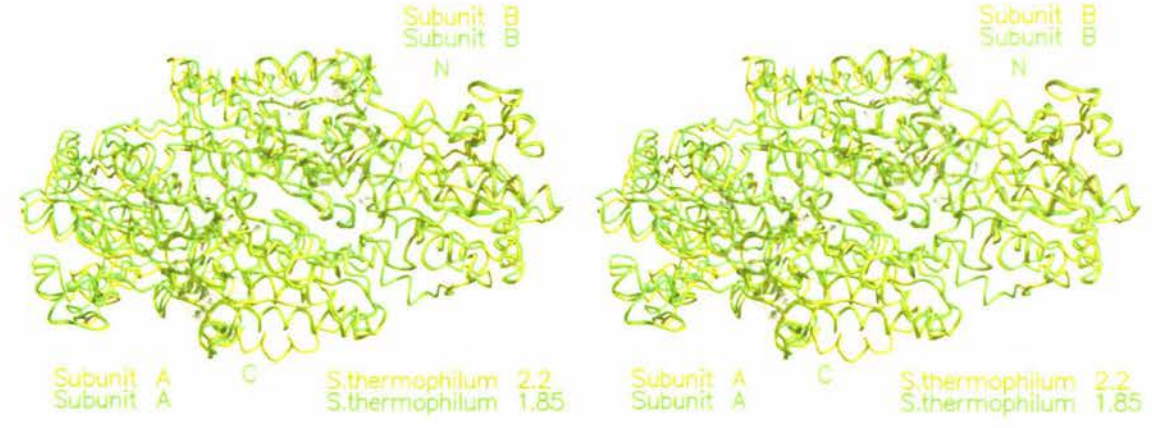


Figure 4.11: Superposition of the *S. thermophilum* PP_i-PFK 1.85 Å (green) and 2.2 Å (yellow) atomic models illustrating the conserved orientation of the N-terminal domain relative to the C-terminal domain of the opposing subunit.

A. Superposition of the *S. thermophilum* PP_i-PFK 1.85 Å (green) and 2.2 Å (yellow) atomic models based upon the N-terminal domain of the A subunits and the C-terminal domain of the B subunits highlighting the conserved orientation of the N-terminal domain to the C-terminal domain of the opposing subunit.

B. Superposition of the *S. thermophilum* PP_i-PFK 1.85 Å (green) and 2.2 Å (yellow) atomic models based upon the N-terminal domain of the B subunits and the C-terminal domain of the A subunits highlighting the conserved orientation of the N-terminal domain to the C-terminal domain of the opposing subunit.

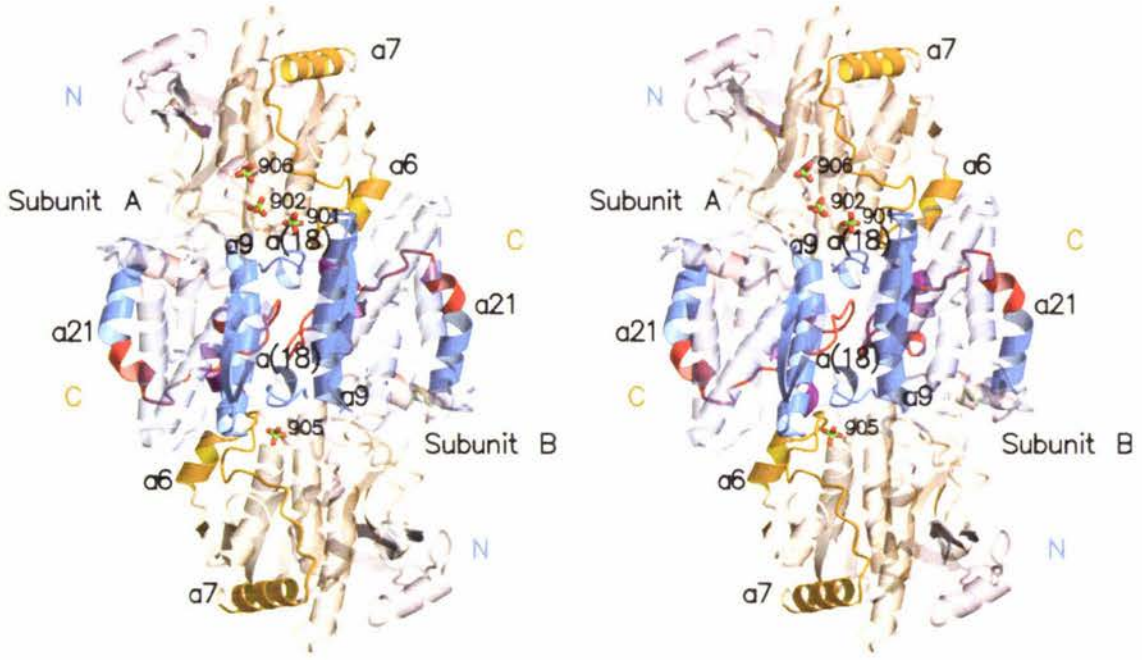
4.7 Dimer Interface

Residue	Atom	Secondary Structure	Residue	Atom	Secondary Structure	Distance
R3A	NH1	N-term	H267B	NE2	α10–β9	2.86
	NH1	N-term	L264B	O	α10	3.02
	NH2	N-term		O	α10	2.91
S5A	OG	N-term	F527B	O	α21–α22	2.56
G137A	O	α6–α7	R236B	NE	α9	3.01
F139A	N	α6–α7	S240B	OG	α9	2.98
D140A	OD1	α6–α7	R236B	NH2	α9	2.80
	OD2	α6–α7		NE	α9	2.84
	N	α6–α7	S240B	OG	α9	3.31
N233A	ND2	α9	E427B	O	α(18)	2.88
D237A	OD1	α9	R429B	NH2	α(18)	2.95
	OD2	α9		NE	α(18)	2.82
Y244A	OH	α9–β8	R389B	NH2	α16–α17	3.10
I531A	O	α21–α22	Q532B	NE2	α21–α22	2.91
Q532A	NE2	α21–α22	1531B	O	α21–α22	2.93
Q532A	OE2	α21–α22	Y533B	N	α21–α22	2.96
Y533A	N	α21–α22	Q532B	OE1	α21–α22	2.91
L264A	O	α10	R3B	NH2	N-term	2.96
F527A	O	α21–α22	S5B	OG	N-term	2.55
R236A	NE	α9	G137B	O	α6–α7	3.13
	NE	α9	D140B	OD2	α6–α7	2.91
	NH2	α9		OD2	α6–α7	3.31
	NH2	α9		OD1	α6–α7	2.95
S240A	OG	α9	F139B	N	α6–α7	3.11
	OG	α9	D140B	OD2	α6–α7	3.3
E427A	O	α(18)	N233B	ND2	α9	2.87
R429A	NE2	α(18)	D237B	OD2	α9	2.77
	NH2	α(18)		OD2	α9	3.23
	NH2	α(18)		OD1	α9	2.98
	NH2	α(18)	S241B	OG	α9–β8	3.30
Q532A	NE2	α21-α22	I531B	O	α21-α22	2.91
I531A	O	α21-α22	Q532A	NE2	α21-α22	2.93
Y533A	N	α21-α22	Q532A	OE2	α21-α22	2.96
Q532A	OE1	α21-α22	Y533A	N	α21-α22	2.91

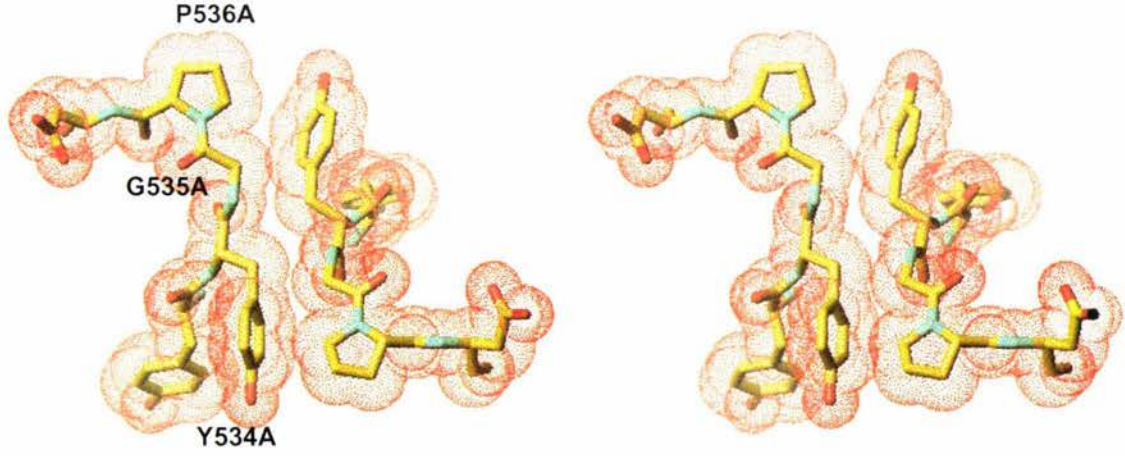
Table 4.6: Interactions between subunits of the *S. thermophilum* PP_i-PFK 1.85Å model across the dimer interface. Interactions are highly symmetrical and conserved relative to the *S. thermophilum* PP_i-PFK 2.2Å model. The majority of inter-subunit contacts are formed between the N-terminal domain and the C-terminal domain of the opposing subunit. Highlighted are asymmetric contacts.

Subunits of the *S. thermophilum* PP_i-PFK homodimer are related by a pseudo dyad symmetry axis (see Figure 4.1). Packing at the dimer interface predominantly involves the N terminus, α6- α7 loop, α21- α22 loop, α(18), and the C-terminus of α9 (see Table 4.6 and Figure 4.12a). Extensive van der Waals contacts are made between Ala530, Tyr534, Gly535, and Pro536 of the α21-α22 loop and their pseudo dyad axis symmetry related counterparts (see Figure 4.12b).

A.



B.



C.

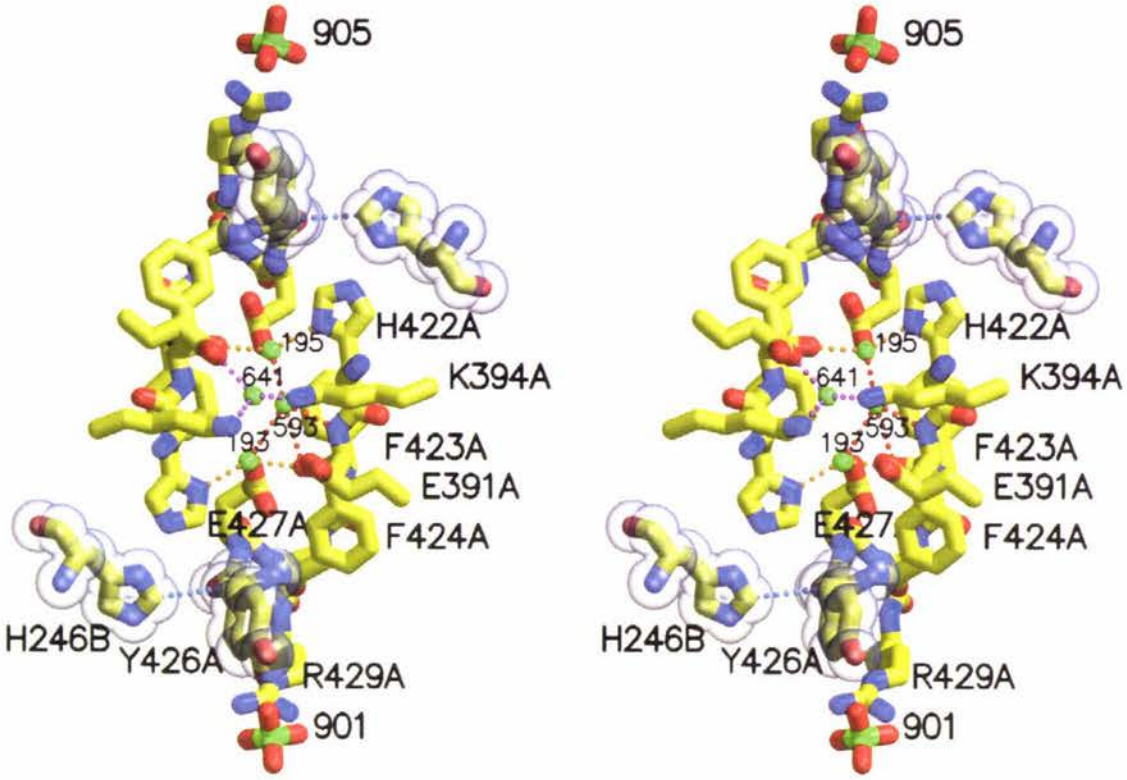


Figure 4.12: Dimer interface of the *S. thermophilum* PP_i-PFK.

A. Ribbon diagram of the *S. thermophilum* PP_i-PFK 1.85 Å structure illustrating the secondary structure elements involved in dimer interface contacts. Figure prepared with MOLSCRIPT (Kraulis 1991) and Raster3D (Merrit and Bacon 1997).

B. Diagram of the van der Waals contacts between Ala530, Tyr534, Gly535, and Pro536 of the α21-α22 loop and their pseudo dyad axis symmetry related counterparts.

C. Ball-and-stick diagram of the water-mediated hydrogen bonds β-strand β11 and its symmetry related counterpart. Figures **B.** and **C.** prepared with TURBO-FRODO (Roussel and Cambillau 1997).

The characterized R to T-state transition of *B. stearothermophilus* ATP-PFK involves expulsion of a layer of water molecules between the two, symmetry related β-strands, I, facilitating formation of direct interactions (Schirmer and Evans 1990). The expulsion is thought to coincide with closure of the active site (Shirakihara and Evans 1988). The dimer interface of *S. thermophilum* PP_i-PFK most closely resembles that of the R-state of *E. coli* ATP-PFK (1PFK). The two symmetry related β-strands, β11, form indirect interactions via a layer of water molecules (see Figure 4.12c). The hydrophobic layer formed by Leu247 and Ile250 beneath the 2 β-strands I of *E. coli* ATP-PFK (1PFK) is not conserved. Rather Ile250 is substituted with Glu427, which further coordinates the layer of water molecules (see Figure 4.12c).

Despite the conformational change; the rigid body displacement of the C-terminal domain, the inter-subunit contacts across the dimer interface of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model are highly symmetrical and conserved, relative to the 2.2 Å model. Inter-subunit contacts are generally formed between the N-terminal domain and the C-terminal domain of the opposing subunit (see Table 4.6). Therefore the N-terminal domain and the C-terminal domain of the opposing subunit can be thought of as rigid bodies and the conformational change as the reorientation of one rigid body with respect to the other (see Figure 4.11 and Table 4.5a). As a direct result rotation of the C-terminal domain of subunit A towards the N-terminal domain of subunit A dictates that the N-terminal domain of subunit B moves with it to conserve the dimer interface i.e. the N-terminal domain of subunit B is effectively rotating towards the C-terminal domain of subunit B. Therefore displacement of the C-terminal domain with respect to N-terminal domain of one subunit dictates displacement of the other to conserve the dimer interface.

The conformation of the 6-phosphate binding site and thus the orientation of F6P/F16bP into the active site is stringently defined by the dimer interface and is relatively conserved between structures. The F6P/F16bP 6-phosphate binding site predominantly involves the FG**Y**EA**R** motif (residues 424-429) and Lys243 and Tyr244 from the opposing subunit. In the *S. thermophilum* PP_i-PFK 2.2Å model Arg429 and Lys243 of both subunits forms salt-bridge contacts, and Tyr244 forms a hydrogen bond with a sulfate ion bound at the F6P/F16bP 6-phosphate binding site. The orientation of Arg429 is strictly defined,

sandwiched between and forming van der Waals contacts with the GGXXPG loop (residues 81-86) and the aromatic ring of Tyr426. NH1 and Nε of Arg429 of both subunits forms salt-bridge contacts with OD1 and OD2, respectively of Asp237 from the opposing subunit. Asp237 is exclusively conserved in large subunit (≈550 amino acids) PFKs from bacteria, protists, and plants (group II long clade family members [see Figure 1.10 and Figure 1.12]). An acidic residue (Asp/Glu) is exclusively conserved in all group II family members. Interestingly a Thr residue is exclusively conserved at this position in prokaryotic and eukaryotic ATP-PFKs and their close relatives (group I and III family members [see Figure 1.10]). OG1 of Thr156 of *E. coli* ATP-PFK (1PFK) forms a strong hydrogen bond with Asp12. Asp12 is positioned upon the GG**D**XPG loop (residues 10-15) and is exclusively conserved in group I and III family members. The Cβ of Thr156 also forms a weak van der Waals contact with the Cβ of Arg252, which is homologous to Arg429 of *S. thermophilum* PP_i-PFK. OG of Ser241 in both subunits of the *S. thermophilum* PP_i-PFK 2.2Å model forms hydrogen bonds with NH1 of Arg429 from the opposing subunit and OD2 of Asp237. Asp237, Ser241, Lys243, and Tyr244 are positioned at the C-terminus of the trans-domain α-helix, α9 and the proceeding loop of the C-terminal domain. Any reorientation of α9 relative to the N-terminal domain of the opposing subunit would cause significant aberration of the F6P/F16bP 6-phosphate binding site (see Figure 4.14). Superposition of the C-terminal domain of subunits A and B of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model illustrates that the Cα of Ser241A is displaced by 0.5 Å (see Figure 4.14). The shift of the Cα is concomitant with the altered conformation of OG of Ser241A. Ser241B shares a conformation similar to that of Ser241 of both subunits of the *S. thermophilum* PP_i-PFK 2.2Å atomic model. It forms a weak hydrogen bond with NH2 of Arg429A (3.3 Å) but does not form a hydrogen bond with Asp237B. OG of Ser241A is rotated away from the Arg429B and forms a hydrogen bond with the sulfate ion bound at the F6P/F16bP 6-phosphate binding site of subunit B (see Table 4.14). Overall the conformational mobility of *S. thermophilum* PP_i-PFK; the displacement of the C-terminal and helical domains has very little effect on the dimer interface. In prokaryotic ATP-PFKs a conformational change at the A:D subunit interface is crucial to the R- to T-state transition.

4.8 Active Site Geometry

The active site lies in a cleft formed at the interface of the N- and C-terminal domains. The spatial orientation of the conserved motifs; **M**G**R** (residues 251-253), GG**D D** (residues 174-177), P**K**TI**D**G**D** (residues 202-208), and FG**Y**EA**R** (residues 424-429) involved in substrate binding and catalysis are conserved between ATP- and PP_i-PFKs. The motifs preside predominantly upon loops at the C-terminus of the respective β-sheets. Analogous to PFK A ATP-PFKs F6P/F16bP is predominantly bound by the C-terminal domain and the phosphoryl donor, in this case PP_i by the N-terminal domain. Superposition of the subunit A active site of *E. coli* ATP-PFK with subunit A of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model identifies the binding sites of three sulfate ions, bound at the active site to be analogous to the F16bP 6-phosphate group (sulfate 901 and 906 [see Figure 4.14]), F16bP 1-phosphate group (sulfate 902 [see Figure 4.14]), and the ADP β-phosphate (sulfate 906 [see Figure 4.24]). The sulfate ion bound at the F6P/F216bP 6-phosphate binding site is predominantly bound by Arg429 (Arg252 in *E. coli* ATP-PFK), Arg146 (Arg72 in *E. coli* ATP-PFK), and Lys243 and Tyr244 (Arg162 in *E. coli* ATP-PFK) from the opposing subunit (see Figure 4.14). Subunit B of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model and subunits A and B of the 2.2 Å atomic model only exhibit a sulfate ion bound at the F6P/F16bP 6-phosphate binding site. Binding is highly conserved.

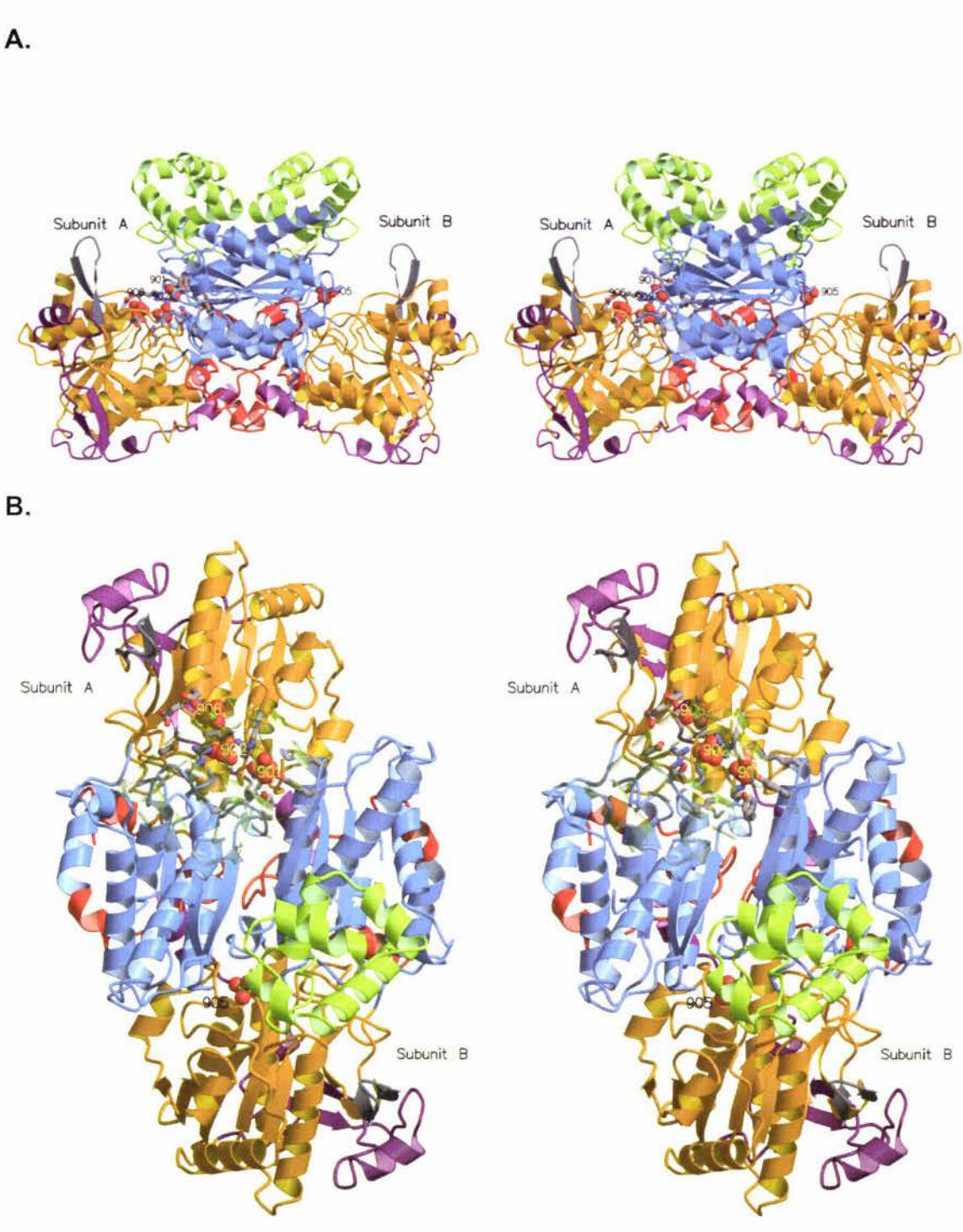


Figure 4.13: Ribbon diagram of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model illustrating the relative orientation of bound sulfate ions within the active site os subunits A and B. **A.** Diagram of *S. thermophilum* PP_i-PFK 1.85 Å atomic model viewed with the pseudo dyad axis parallel to the plane of the page.

B. Diagram of *S. thermophilum* PP_i-PFK 1.85 Å atomic model viewed down the pseudo dyad axis. The helical domain (green) of subunit A is transparent for clarity. **A.** and **B.** were drawn with MOLSCRIPT (Kraulis 1991) and Raster3D (Merritt and Bacon 1997).

4.9 F6P/F16bP binding site

The F6P/F16bP binding site is predominantly formed by the conserved **M**G**R** and FG**Y**EA**R** motifs, contributed by the C-terminal domain. Lys243 and Tyr244, positioned upon the α9-β8 loop of the C-terminal domain from the opposing subunit project into the active site and play an integral role in binding the F6P/F16bP 6-phosphate moiety (see Figure 4.14).

A.

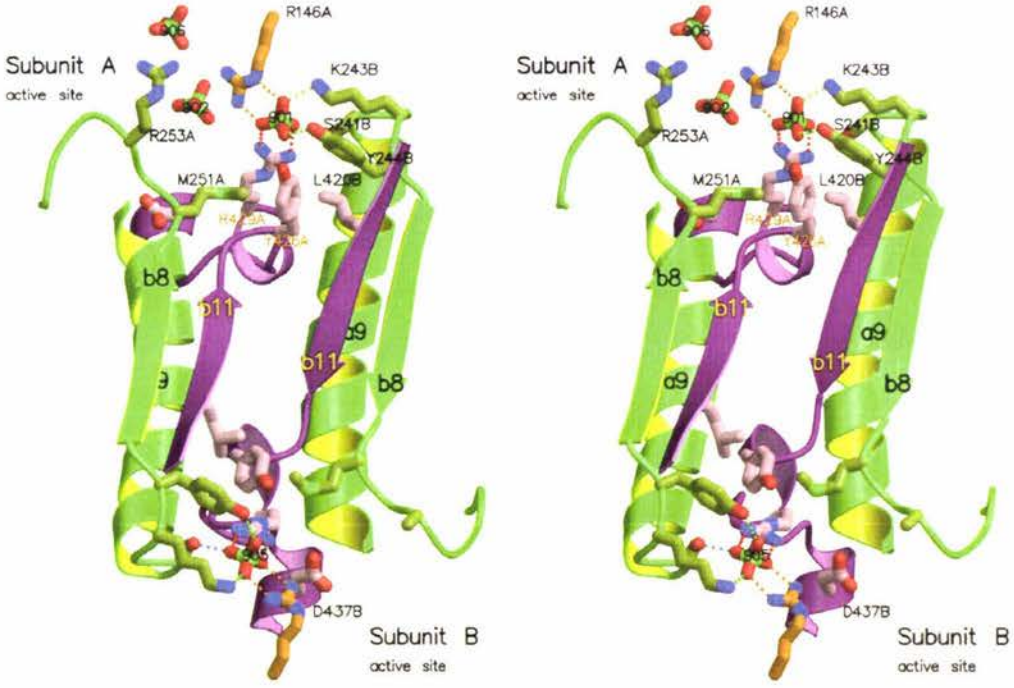


Figure 4.14: Ribbon diagram of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model active sites viewed down the pseudo dyad symmetry axis. Identical residues are shown at both active sites, though only residues of subunit A are labeled. Binding of the sulfate ion at the F6P/F16bP 6-phosphate binding sites is highly conserved with the exception of the conformational change of Ser241A, resulting in the formation of a hydrogen bond (blue) with the 6-phosphate sulfate ion (905) at the active site of subunit B and the altered conformation of Arg146B relative to Arg146A. Figure prepared with MOLSCRIPT (Kraulis 1991) and Raster3D (Merritt and Bacon 1997).

B.

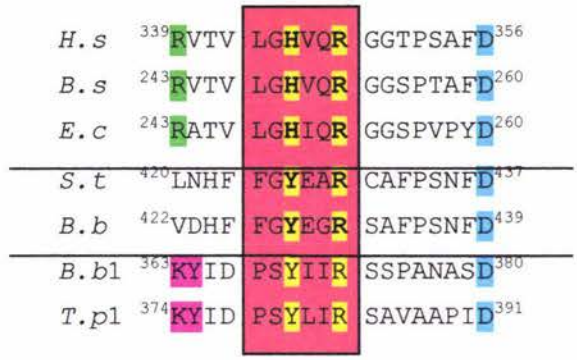


Figure 4.14b: Structure-based sequence alignment of the F6P/F16bP binding motif of *S. thermophilum* PP_i-PFK (*S.t*), *B. burgdorferi* PP_i-PFK (*B.b*), *B. burgdorferi* (pfk1) (*B.b*1), *T. pallidum* (pfk1) (*T.p*1), *H. sapiens* ATP-PFK (liver-type), *B. stearothermophilus* ATP-PFK (*B.s*), and *E. coli* ATP-PFK (*E.c*). See Figure 4.2.

	Sulfate	atom	Residue	Atom	Distance Å
	6-phosphate				
Subunit A					
	901	O4	K243B	NZ	2.88
	901	O4	R146A	NE	3.11
	901	O3	R146A	NH2	2.79
	901	O3	R429A	NH1	2.82
	901	O2	R429A	NH2	2.98
	901	O1	Y244B	OH	2.77
Subunit B					
	903	O2	K243A	NZ	2.66
	903	O2	R146B	NH1	2.72
	905	O1	R146B	NH2	2.81
	903	O1	R429B	NH1	2.75
	903	O4	R429B	NH2	3.05
	903	O4	S241A	OG	2.74
	903	O3	Y244A	OH	3.15

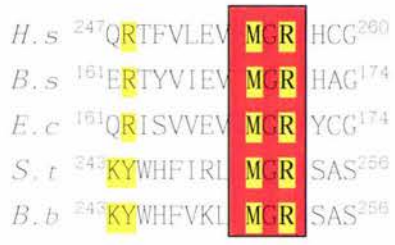
Table 4.7a: Contacts made by sulfate ions bound at the putative F6P/F16bP 6-phosphate binding site of subunits A and B of the *S. thermophilum* PP_i-PFK 1.85Å structure. Note that the contacts are not symmetrical. Differences are shaded. See Figure 4.14.

	Sulfate	atom	Residue	Atom	Distance Å
	1-phosphate				
Subunit A					
	902	O2	G82A	N	3.29
	902	O2	X499		2.70
	902	O3	T204A	OG1	2.82
	902	O3	X507		2.74
	902	O4	G82A	N	2.82
	902	O4	K203A	NZ	3.07

Table 4.7b: Contacts made by the sulfate ion bound at the likely F16bP 1-phosphate binding site of subunit A of the *S. thermophilum* PP_i-PFK 1.85Å structure. See Figure 4.14.

The **M**G**R** motif is positioned upon the β8-α10 loop. Immediately preceding β-strand β8 is the conserved Lys243 and Tyr244, which contribute to the F6P/F16bP binding site of the opposing subunit. The FG**Y**EA**R** motif is positioned upon helix α(18), which is preceded by β-strand β11 of the C-terminal domain β-sheet. Tyr244 aligns with Arg162 of *E. coli* ATP-PFK (see Figure 4.2). Lys243, however does not align with Arg243 of *E. coli* ATP-PFK. Hydrophobic packing of β-strand β11 with helix α9 is extensive and highly conserved within the group II long clade PFKs. Though not strongly conserved Leu420 packs between and forms exquisite van der Waals contacts with the aromatic rings of Tyr244 and Tyr426 of the FG**Y**EA**R** motif (see Figure 4.15b). A hydrophobic residue is broadly conserved at this position in Group II PFK family members. Interestingly Leu420's position upon β-strand β11 is homologous to an exclusively conserved Arg (Arg243 in *E. coli* ATP-PFK) in group I and III PFKs involved in binding the F6P 6-phosphate group across the dimer interface in concert with Arg162 (see Figure 4.2 and Figure 1.5b).

A.



B.

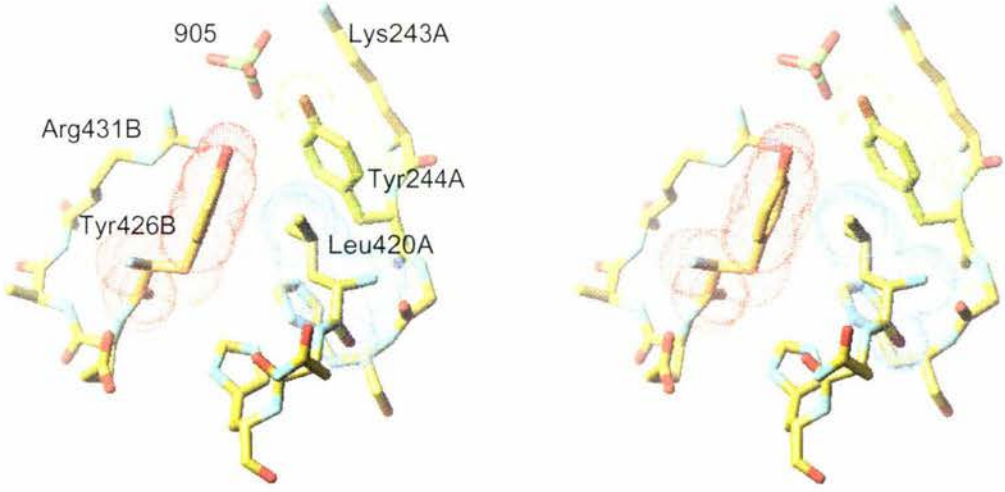


Figure 4.15: **A.** Structure-based sequence alignment of the MGR motif of *S. thermophilum* PP_i-PFK (*S.t*), *B. burgdorferi* PP_i-PFK (*B.b*), *H. sapiens* ATP-PFK (liver-type), *B. stearothermophilus* ATP-PFK (*B.s*), and *E. coli* ATP-PFK. See Figure 4.2. **B.** Ball-and-stick diagram of the sulfate ion (905) bound at the F6P/F16bP binding site of subunit B of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model (carbon = yellow, nitrogen = blue, oxygen = red, and sulfur = green). van der Waals surfaces are shown for Tyr426B (red), Tyr244A (green), and Leu420A (blue). Drawn with TURBO-FRODO (Roussel and Cambillau 1991).

Both Arg162 and Arg243 are implicated in the allosteric regulation of ATP-PFKs (Berger and Evans 1990). Interestingly group II x clade family members (see Figure 4.10), including the ATP-PFK of *E. histolytica* and the homologous putative ATP-PFK of *T. pallidum* and *B. burgdorferi* do not exhibit the distinctive $K^{243}Y^{244}$ and ensuing large hydrophobic residues of the long clade. They do however exhibit a highly conserved Lys residue, which aligns with Arg243 of *E. coli* ATP-PFK and Leu420 of *S. thermophilum* PP_i-PFK, which is immediately followed by a highly conserved Tyr (within the x clade) (see Figure 4.15a). The lack of conservation of Arg162 and Arg243, implicated in the R- to T-state transition of prokaryotic ATP-PFKs, characterized in *B. stearothermophilus* (Schirmer and Evans 1990) indicates that it is implausible PP_i-PFKs undergo such a conformational transition. The lack of conservation of residues involved in formation of the F6P/F16bP binding site across the dimer interface is suggested to be a direct effect of the displacement of the C-terminal domain witnessed between *E. coli* ATP-PFK and *S. thermophilum* PP_i-PFK (see Figure 4.3). The C-terminal domain of group II x clade family members may not exhibit this displacement as they exhibit a conserved Lys which aligns with Arg243 of *E. coli* ATP-PFK and a theoretical allosteric binding site (see Figure 1.13). Displacement of the C-terminal domain results in conformational deviation of the trans-domain α-helix α9 relative to helix α6 of *E. coli* ATP-PFK [1PFK]). The structure at the N-terminus of the α-helix is highly conserved. An exclusively conserved Asp, Asp437, forms the helical N-cap in *S. thermophilum* PP_i-PFK, which proceeds the FG**Y**EA**R** motif (see Figure 4.14). In the *E. coli* ATP-PFK (1PFK) structure helix α6 (139-160) is decidedly kinked as opposed to helix α9 of the *S. thermophilum* PP_i-PFK and *B. burgdorferi* PP_i-PFK atomic models. The *E. coli* ATP-PFK peptide oxygen atom of residues 139, 141, 142, and 146-150 form 3_{10} α-helical hydrogen bonds with the peptide nitrogen atoms of 142, 144, 145, and 149-153, respectively. Residues 151,152, 155, and 156, though exhibiting α-helical torsion angles do not form α-helical hydrogen bonds. All ATP-PFK structures in the Protein Data Bank (PDB [www.rcsb.org/pdb/]) exhibit this kinked conformation of the trans-domain α-helix (1PFK, 2PFK, 3PFK, 4PFK, and 6PFK). Structure-based sequence alignment (see Figure 4.2 and Figure 4.16) reveals a single residue insert at the C-terminus of the *S. thermophilum* and *B. burgdorferi* PP_i-PFK trans-domain α-helix (α9) relative to that of prokaryotic ATP-

PFKs. The tertiary structure of the C-terminal domain β-sheet is conserved; superposition of the respective β-strands yields an rms displacement of 0.783 Å. Helix α6 of prokaryotic ATP-PFKs is kinked to accommodate the single residue deletion, therefore maintaining the tertiary structure of the β-sheet.

A.

```
S.t 219GFDTATKVYAELIGNIARDAISSRK
B.b 219GFDSATKIYSELIGNLCRDAMSTKK
B.s 138GFDTALNTVIDAIDKIRDTATS-HE
E.c 138GFFTALSTVVEAIDRLRDTSSS-HQ
```

B.

```
S.t 219GFDTATKVYAELIGNIARDAISSRK
B.b 219GFDSATKIYSELIGNLCRDAMSTKK
B.s 138GFDTALNTVIDAI-DKIRDTATSHE
E.c 138GFFTALSTVVEAI-DRLRDTSSSHQ
```

Figure 4.16: **A**. Structure-based sequence alignment of *S. thermophilum* PP_i-PFK (*S.t*), *B. burgdorferi* PP_i-PFK (*B.b*), *B. stearothermophilus* ATP-PFK (*B.s*), and *E. coli* ATP-PFK (*E.c*). **B**. Shifting the single insert yields a better, though erroneous alignment.

It is interesting to note that the group II PP_i-specific Tyr426 of the FG**Y**EAR motif is not involved in binding the sulfate ion at the F6P/F16bP 6-phosphate binding site in either of the *S. thermophilum* PP_i-PFK or *B. burgdorferi* PP_i-PFK structures. The homologous residue of *E. coli* ATP-PFK, His249, is involved in binding the 6-phosphate group of F6P. An *E. coli* ATP-PFK His249Ala mutant exhibited a 400-fold reduction in $K_{M[F6P]}$ whereas a His249Tyr mutant exhibited a mere 5-fold reduction, suggesting that their function is redundant (Deng *et al.* 2000). The exclusive conservation of a His residue in group I and group III PFKs and a Tyr residue in group II PFKs indicates that their function is not completely redundant, as illustrated by the reciprocal mutation, Tyr420His, in the group II, long clade PP_i-PFK of *E. histolytica*. The mutation resulted in a 250-fold reduction in $K_{M[F6P]}$. A Tyr420Phe mutation resulted in a 80-fold reduction, suggesting that the aromatic functional group of the Tyr residue is of primary importance (Deng *et al.* 2000). The function of the aromatic ring of His249 in the ATP-PFK of *E.*

coli is unclear due to the lack of site-directed mutagenesis studies and kinetic information. In all four subunits of *E. coli* ATP-PFK (1PFK) the terminal methyl group (Cε) of Met169 of the **M**GR motif points towards and forms van der Waals contacts with the imidazole ring of His249 of the F6P 6-phosphate binding LG**H**IQ**R**. The sulfur atom forms a weak hydrogen bond with the O2 hydroxyl group of F16bP. The peptide nitrogen atom of Gly170 of *E. coli* ATP-PFK (1PFK) forms a hydrogen bond with the O3 hydroxyl group of F16bP. An exclusively conserved Glu, Glu222 forms a hydrogen bond with the O4 hydroxyl group of F16bP (and presumably with F6P) and with the amino nitrogen atoms of Gly170 and Arg171 of the MGR motif (Shirakihara and Evans 1988 [see Figure 1.6d]). The conformation of the Met251 and Arg253 (**MGR**) side chains in both subunits of the *B. burgdorferi* PP_i-PFK atomic model approximate those of *E. coli* ATP-PFK (1PFK). Superposition of the MGR motifs illustrates that Met251 of *B. burgdorferi* PP_i-PFK is in an active conformation, the Cε forming a van der Waals contact with the aromatic ring of Tyr428 (FG**Y**EGR) (see Figure 4.17).

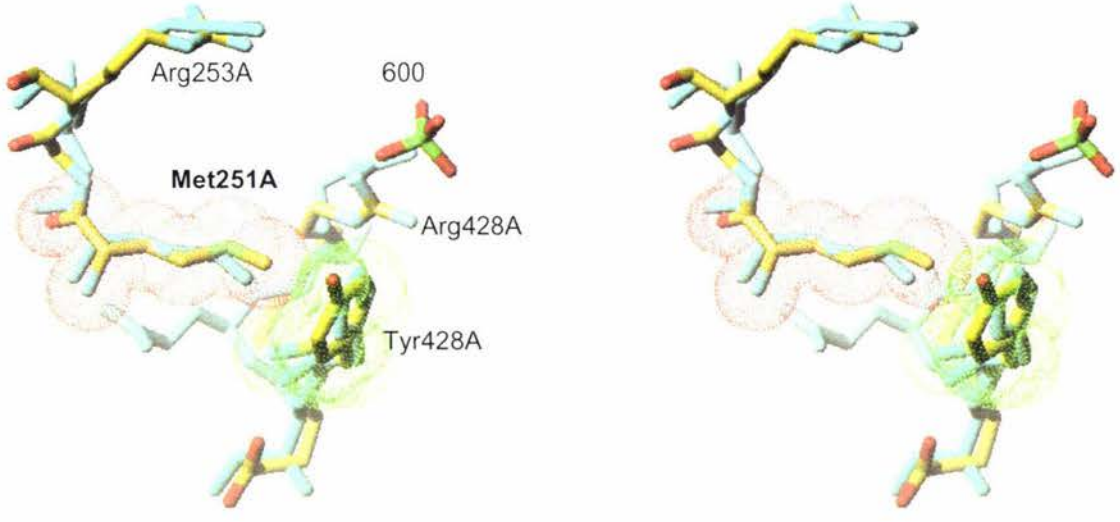


Figure 4.17: Ball-and-stick diagram of the 'active' van der Waals contact between Met251A (**M**GR) and Tyr428A (FG**Y**EGR) of the *B. burgdorferi* PP_i-PFK (1KZH [Moore *et al.* 2002]). Van der Waals surfaces are shown in red and green (carbon = yellow, nitrogen = blue, oxygen = red, sulfur = green). The MGR and LGHIQR motifs of subunit A of *E. coli* ATP-PFK (1PFK) is shown blue. Drawn with TURBO-FRODO (Roussel and Cambillau 1991).

The van der Waals contact is significantly stronger than that between Met169 and His429 of *E. coli* ATP-PFK (1PFK). The average distance of the six van der Waals contacts between the Cε of Met251 A and the aromatic carbon atoms of Tyr428 is 3.63 Å, compared to 3.8 Å for the three van der Waals contacts formed between the Cε of Met169 and His429.

The function of the group II PP_i-specific Tyr426 (FG**Y**EAR) in *S. thermophilum* PP_i-PFK is to correctly coordinate Met251 (**M**GR) into the active site, the orientation of which is stringently defined by the conformation of the C-terminal domain. This is best illustrated by a discussion of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model.

4.9.1 Conformational change: the MGR motif

Superposition of the MGR motif of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model with that of *E. coli* ATP-PFK (1PFK) indicates that the MGR motif in both subunits of *S. thermophilum* is in a conformation that is not conducive with F6P/F16bP binding (see Figure 4.18). The Cε of Met251 points away from the aromatic ring of Tyr428, into the active site and the side chain of Arg253 is oriented out of the active site (see Figure 4.18).

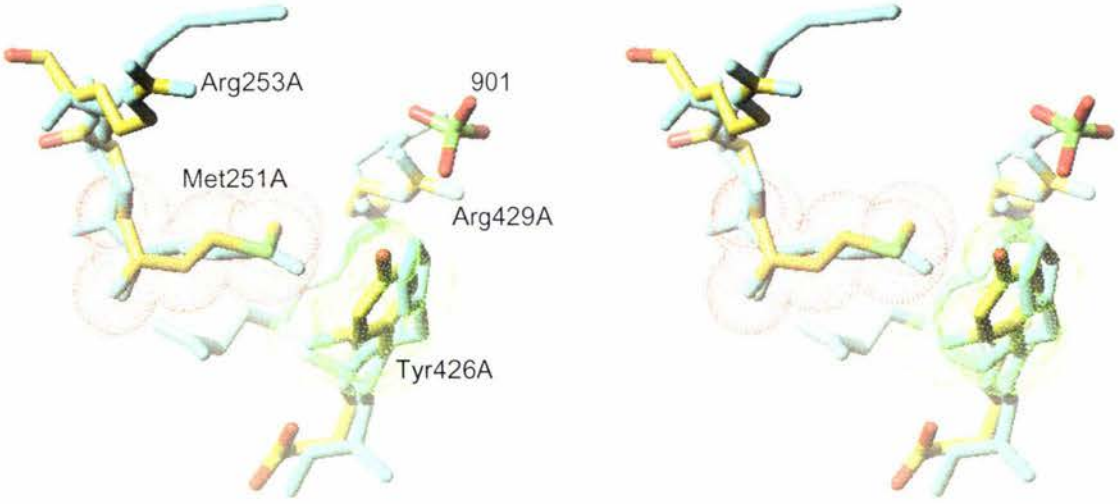


Figure 4.18: Ball-and-stick diagram of the subunit A **MGR** and FG**Y**EA**R** motifs of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model illustrating the inactive conformation of the active site (carbon = yellow, nitrogen = blue, oxygen = red, sulfur = green). Note that Cε of Met251 is rotated away from the aromatic ring of Tyr426 and Arg253 is oriented out of the active site. Van der Waals surfaces are shown for Met251 (red) and Tyr426 (green). The MGR and LGHIQR motifs of subunit A of *E. coli* ATP-PFK (1PFK) are shown blue. Drawn with TURBO-FRODO (Roussel and Cambillau 1991).

In this conformation it is likely that F6P/F16bP would not be able to bind due to unfavorable steric interactions with Cε of Met251. The dihedral angles of the conserved MGR motifs of *E. coli* ATP-PFK and *S. thermophilum* PP_i-PFK differ markedly (see Table 4.8).

Residue		*S. thermophilum* 2.2 Å		*S.thermophilum* 1.85 Å		*B. burgdorferi* 1KZH	
		A	B	A	B	A	B
Met251	phi	-67	-61	-62	-59	-56	-55
	psi	145	146	151	148	160	150
Gly252	phi	127	126	121	123	126	123
	psi	94	93	99	91	76	81
Arg253	phi	-81	-79	-94	-76	-63	-68
	psi	-42	-41	-36	-43	-46	-22
Ser254	phi	-49	-49	-74	-59	-122	-105
	psi	-41	-48	-34	-49	4	-33
Ala255	phi	-157	-150	-147	-132	-164	-143
	psi	-174	-173	179	-172	158	-178
Ser256	phi	-107	-107	-105	-113	-109	-125
	psi	20	17	22	20	19	16
His257	phi	-56	-53	-59	-55	-55	-51
	psi	-50	-48	-46	-43	-44	-46

Table 4.8: Comparison of the dihedral angles of the MGR motif of the *E. coli* ATP-PFK (1PFK), and *B. burgdorferi* (1KZH [Moore *et al.* 2002]) and *S. thermophilum* PP_i-PFK atomic models. The dihedral angles of Ser254 (shaded) deviate the greatest.

The resultant backbone conformation of the *S. thermophilum* PP_i-PFK 2.2 Å atomic models MGR motifs would prevent the peptide amino group of Gly252 (M**G**R) forming a hydrogen bond with the O3 hydroxyl group of F6P/F16bP and the exclusively conserved Glu312 (Glu222 in *E. coli* ATP-PFK) from forming a hydrogen bond with Arg253. Note that the dihedral angles of Arg171 of *E. coli* ATP-PFK (1PFK) are highly unfavorable, falling within a disallowed region of the Ramachandran plot. The dihedral angles of the MGR motif and the orientation of the Arg253 side chain out of the active site closely resembles that of the unliganded *E. coli* ATP-PFK structure (2PFK [Rypniewski and Evans 1989]).

Despite rotation about the Cβ-Sδ bond of Met251 in either subunit of the *S. thermophilum* PP_i-PFK 2.2Å atomic model, formation of van der Waals contacts between

the Met251 Cε and the aromatic ring of Tyr426 is impossible due to their distant spatial orientation (average distance between Cε of Met251 A and the six aromatic carbon atoms of Tyr428 is 4.01 Å). The conformational change associated with the *S. thermophilum* PP_i-PFK 1.85 Å atomic model, specifically the rotation of the C-terminal domain reorients the Sδ of Met251 relative to the aromatic ring of Tyr426 (FG**Y**EAR) (see Figure 4.20). The Cε rotates about the Cβ-Sδ bond into a putative 'active' conformation, forming a strong van der Waals contact with Tyr426 that mirrors that of Met251 of *B. burgdorferi* PP_i-PFK (1KZH) and Met169 of *E. coli* ATP-PFK (1PFK) (see Figure 4.20). The conformational change of the Cε is forced by the rotation of the C-terminal domain. Rotation of the C-terminal domain concomitant with maintenance of the inactive conformation of Met251 would result in unfavorable steric contacts between Cε of Met251 and Cδ of Ile205 (PKT**I**DND) and Cγ and Cδ of Arg429 (FGYEA**R**). Ile205 is highly conserved within the PKT**I**DGD motif in PFK A family members. Its function may involve the correct orientation of Met251. There is no crystallographic evidence suggesting that the conserved Met of the MGR motif and the C-terminal domain of prokaryotic ATP-PFKs undergo such a conformational change. A partial function of the conserved Ile (PGT**I**DND) in ATP-PFKs may be to simply maintain the 'active' conformation of Met251.

The MGR motif loop (β8-α10) of both subunits of the *S. thermophilum* PP_i-PFK 1.85Å model is disordered to a greater degree, relative to the *S. thermophilum* PP_i-PFK 2.2 Å atomic model suggesting movement of the loop within the crystal. The disorder is illustrated by comparison of the respective correlation coefficients and divergence from the average residue B factor for each subunit of the respective *S. thermophilum* PP_i-PFK atomic models (see Figure 4.19). The increased disorder is resultant from the conformational change of the helical domain. Group II long clade family members all exhibit a highly conserved loop, (V/I)P$\underline{E}^{312}$G(L/V)(V/I)($\underline{E/D}^{316}$) which immediately precedes the first α-helix of the helical domain, α13 and includes the exclusively conserved Glu312. In both subunits of the *S. thermophilum* PP_i-PFK 2.2Å model Glu316 forms a salt-bridge with Arg253. An acidic residue (Asp/Glu) is exclusively conserved at this position in Group II long clade family members (see Figure 1.11). Arg253A NH2 contacts OE1 of Glu316A (2.52 Å) and Arg253A Nε contacts OE2 of Glu316A (2.61 Å).

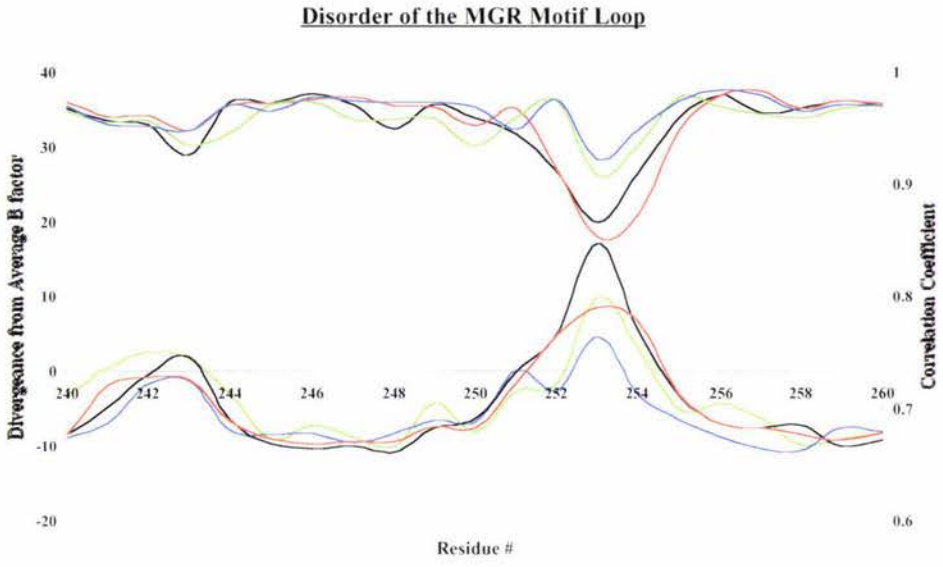


Figure 4.19: A plot of the correlation coefficients (top) and divergence from the average B factor (bottom), calculated for each subunit of the *S. thermophilum* PP_i-PFK 1.85 Å (subunit A = black, subunit B = red) and 2.2 Å (subunit A = green, subunit B = blue) atomic models.

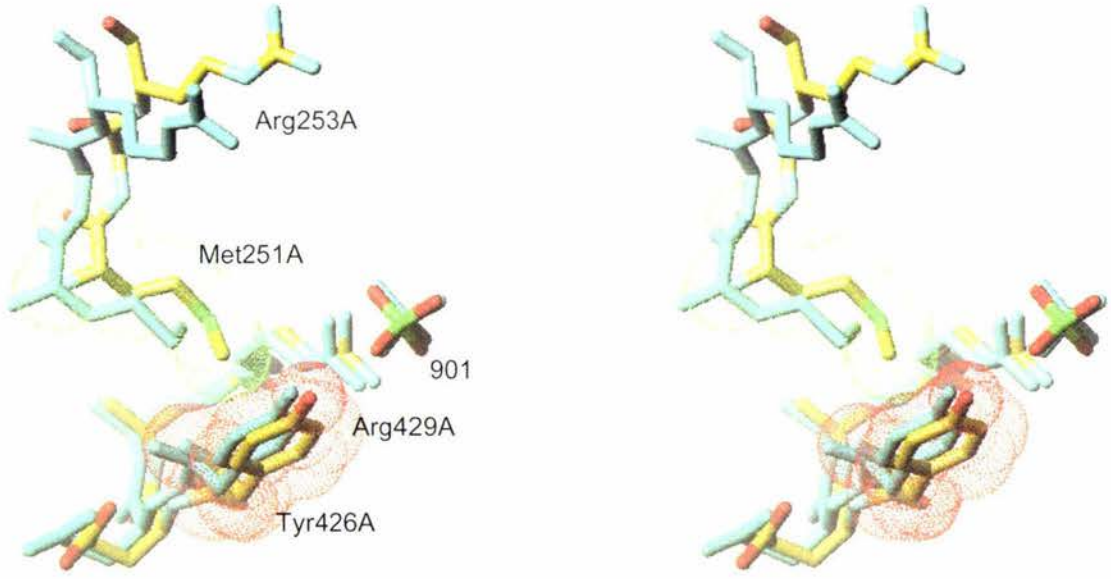


Figure 4.20: Stereo diagram of the superposition of the *S. thermophilum* PP_i-PFK atomic models (2.2 Å light blue, 1.85 Å carbon = yellow, nitrogen = blue, oxygen = red, sulfur = green) MGR and FGYEAR motifs based upon the large domain core. Resultant of the conformational change is the transition of the MGR motif into an active conformation. Drawn with TURBO-FRODO (Roussel and Cambillau 1991).

In the *S. thermophilum* PP_i-PFK 1.85 Å atomic model neither MGR motif forms this contact (see Figure 4.21a). The side chain of Arg253B of the *S. thermophilum* PP_i-PFK 1.85Å model is highly disordered. Electron density is not discernible. The side chain conformation of Arg253A is however discernible. Superposition of the subunit A C-terminal domain from both *S. thermophilum* PP_i-PFK atomic models illustrates the altered conformation of Arg253 (MG**R**) (see Figure 4.21a). The conformation of Arg253A is altered as a direct result of the conformational change of the helical domain. Arg253A of the *B. burgdorferi* PP_i-PFK best illustrates this. Despite comparative dihedral angles with the *S. thermophilum* PP_i-PFK 1.85 Å atomic model for the MGR motif (see Table 4.8) its side chain is oriented further into the active site. In *B. burgdorferi* PP_i-PFK Arg253A of the MGR motif forms a van der Waals contact with Asn386A of the 380-390 β-hairpin, which also forms a hydrogen bond with Asp177A of the GGDD motif. Ser390A OG forms a hydrogen bond with Glu312A OE2 and Val389A CG2 forms a van der Waals contact with Glu312A CG. Movement of the 380-390 β-hairpin into the active site 'folds' the side chain of Arg253 into the active site (see Figure 4.21b). The same effect causes the observed deviation of the side chain conformation of Arg153A between the two *S. thermophilum* PP_i-PFK atomic models. It can therefore be concluded that folding of the side chain of Arg253 into the active site breaks the salt bridge contact with Glu316, resulting in the increased disorder of the MGR motif observed for both subunits of the *S. thermophilum* PP_i-PFK 1.85 Å structure. The conformation of the **MGR** motif of *S. thermophilum* and *B. burgdorferi* PP_i-PFK is stringently defined by the orientation of the C-terminal and helical domains relative to the N-terminal domain.

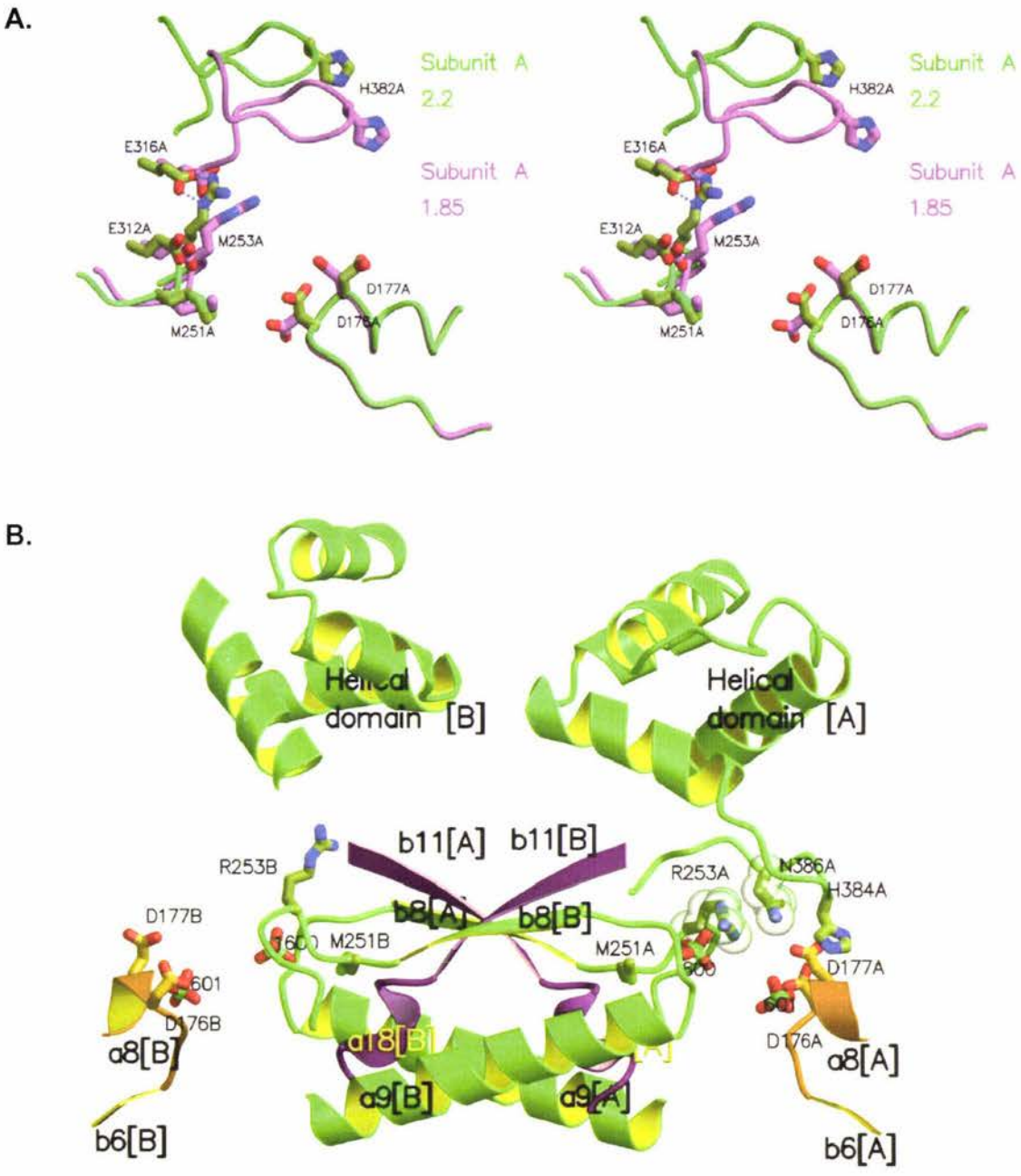


Figure 4.21: **A.** Superposition of the 2.2 Å (green) and 1.85 Å (purple) *S. thermopphilum* PP_i-PFK atomic models illustrating the distinct contacts formed by Arg253A. Drawn with TURBO-FRODO (Roussel and Cambillau 1991). B. Ribbon diagram of the F6P binding sites of the *B. burgdorferi* PP_i-PFK atomic model (Moore *et al.* 2002). The displacement of the subunit A 380-190 β-hairpin into the active site 'folds' the side chain of Arg253 into an active conformation. Note the absence of the subunit B 380-390 β-subunit iinto the active site results in the orientation of Arg253B out of the active site. Van der Waals surfaces are shown from Arg253A and Asn386A (green). Drawn with Molscript (Kraulis 1991) and Raster3D (Merritt and Bacon 1997).

4.9.2 Arg146:Transition state stabilization

The conformation of Arg146 differs between subunits of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model (see Figure 4.14). Recall that only subunit A exhibits sulfate ions bound at the putative F16bP 1-phosphate binding site and below the GGDD motif, and that both subunits exhibit sulfate ions at the F6P/F16bP binding site, bound by Arg146, the FGYEAR motif, and Lys243 and Tyr244 from the opposing subunit.

The guanidinium group NH1 and NH2 nitrogen atoms of Arg146A point towards, but do not contact the sulfate ion bound at the putative F16bP 1-phosphate binding site (902) (see Figure 4.14 and Table 4.7b). NH2 and Nε of the guanidinium group form contacts with the sulfate ion bound at the F6P/F16bP 6-phosphate binding site (901) (see Figure 4.14 and Table 4.7a). This conformation of Arg146A is observed in both subunits of the *S. thermophilum* PP_i-PFK 2.2 Å structure. The guanidinium group orientation of Arg146B differs; it is rotated about the Cδ-Nε bond by approximately 120°, relative to Arg146A and points down into the putative F6P/F16bP 6-phosphate binding site. NH1 and NH2 of the guanidinium group contact the sulfate ion (905) (see Figure 4.14 and Table4.7a). Superposition of the *B. burgdorferi* PP_i-PFK subunits based upon the N-terminal domain core illustrates a deviation of 0.62 Å between the sulfur atoms of the sulfate ions bound at the F6P/F16bP 6-phosphate binding sites. An identical superposition of the *S. thermophilum* PP_i-PFK 2.2 Å atomic models subunits illustrates a deviation between the sulfur atoms of the bound sulfate ions of only 0.1 Å. The sulfate ions bound at the F6P/F16bP 6-phosphate binding sites of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model deviate by 1.38 Å when the subunits are superimposed. The altered conformation of the guanidinium group between subunits of the 1.85 Å structure is a direct result of the difference in the orientation of the sulfate ions bound at the F6P/F16bP 6-phosphate binding site, which is stringently defined by the orientation of the C-terminal domain relative to the N-terminal domain. The conformation of the guanidinium group alters with the orientation of the sulfate ion to conserve contacts. Of the three PP_i-PFK structures the subunit asymmetry relative to the C-terminal domain is greatest for the *S. thermophilum* PP_i-PFK 1.85 Å structure (see Figure 4.22). The average main chain rms difference between the C-terminal domains of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model based upon superposition of the N-terminal domains of

subunit A and B is 1.8 Å, relative to 1.3 Å for the *B. burgdorferi* PP_i-PFK atomic model and 0.7 Å for the *S. thermophilum* PP_i-PFK 2.2 Å atomic model. A direct consequence is that the orientation of the F6P/F16bP 6-phosphate binding motif; FGYEAR (positioned upon a C-terminal domain loop) into the active site deviates between subunits A and B the greatest in the *S. thermophilum* PP_i-PFK 1.85 Å atomic model. Contacts formed between the motif and the sulfate ions bound at the F6P/F16bP binding site of both subunits is conserved (see Figure 4.14). Therefore deviation between subunits in the orientation of the FGYEAR motif correlates with a deviation in the orientation of the sulfate ions bound at the F6P/F16bP binding sites. In all three structures the dimer interface is relatively conserved i.e. the orientation of the conserved Lys and Tyr residues from the opposing subunit to the FGYEAR motif is conserved. Therefore the conformation of Arg146 deviates between subunits to conserve interactions with the bound sulfate ions. The distinct conformation of Arg146B is observed in the *B. stearothermophilum* ATP-PFK (4PFK), which exhibits F6P and ADP bound at the active site.

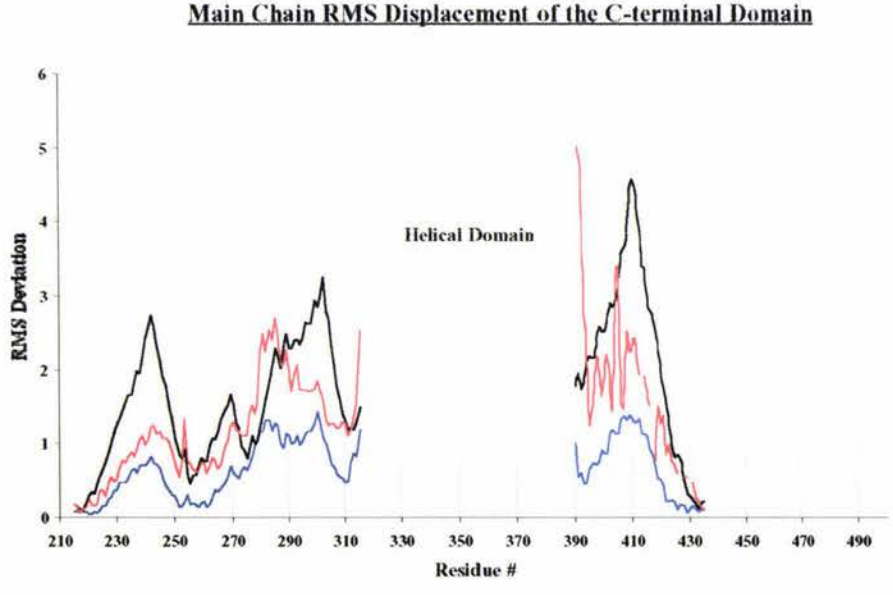


Figure 4.22: A plot of the C-terminal domain rms difference between subunits of *B. burgdorferi* PP_i-PFK (red), *S. thermophilum* PP_i-PFK 1.85 Å (black), and *S. thermophilum* PP_i-PFK 2.2 Å (blue) respectively, based upon superposition of the N-terminal domain core (see Table 4.2a). Note the difference is greatest for the *S. thermophilum* PP_i-PFK 1.85 Å structure.

4.10 Pyrophosphate Binding Site

The pyrophosphate-binding site of *S. thermophilum* PP_i-PFK is formed within the N-terminal domain, at the base of helix α8 and involves the motifs GG**DD** and P**K**TI**D**G**D** (see Figure 4.23). The GGDD motif forms the N-terminus of helix α8. The PKTIDGD motif is positioned at the C-terminus of strand β7 and the β7-α9 loop (see Figure 4.2 and Figure 4.24).

A.

H.s	^{161}CVI	GGDG	SLTGADTFRSEWS^{180}
E.c	^{98}VVI	GGDG	SYMGAMRLTEMG-^{116}
*B.b*1	^{176}FNI	GGDG	TQKGSLLIAEEIE^{195}
T.t	^{86}VAI	GGDD	TLGAAGEAGRRG-^{104}
P.f	^{116}HTI	GGDD	TNTTAADLAAYLA^{135}
N.f	^{117}VTI	GGDD	TAFSSMSVAKAAN^{136}
S.t	^{171}VII	GGDD	SNTNAALLAEYFV^{190}
B.b	^{171}III	GGDD	SNTNAAILAEYFK^{190}

B.

H.s	^{206}VGL	VGSIDND	FCG--TDMTIGTDS^{227}
E.c	^{120}IGL	PGTIDND	IKG--TDYTIGFFT^{141}
*B.b*1	^{204}VGI	PKTVDND	FMF--VQKSFGFET^{225}
T.t	^{119}VGI	PKTIDND	VYG--TDYTIGFDS^{140}
P.f	^{144}VGL	PKTIDND	IVP--IRQSLGAWT^{165}
N.f	^{142}CHV	PKTIDND	LPLPYGIPTFGYET^{165}
S.t	^{199}IGI	PKTIDGD	LKNEYIEASFGFDT^{222}
B.b	^{199}IGV	PKTIDAD	LRNDHIEISFGFDS^{222}

Figure 4.23: A. and B. Structure-based sequence alignment of the GGD(D/G) and P(K/G)TxDxD motifs of *S. thermophilum* PP_i-PFK (*S.t*), *B. burgdorferi* PP_i-PFK (*B.b*), *B. burgdorferi* (pfk1) (*B.b*1), *T. tenax* PP_i-PFK (*T.t*), *H. sapiens* ATP-PFK (liver-type), *N. fowleri* PP_i-PFK (*N.f*), *P. freudenreichii* PP_i-PFK (*P.f*), and *E. coli* ATP-PFK.

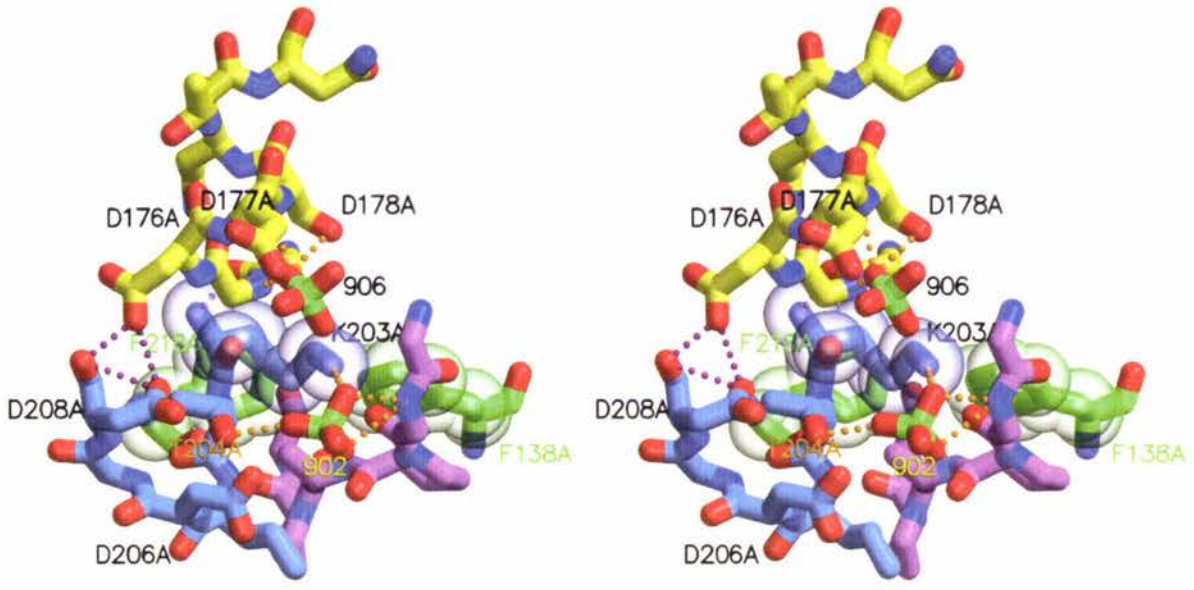


Figure 4.24: Ball-and-stick diagram of the PP_i-binding site of subunit A of the *S. thermophilum* PP_i-PFK 1.85 Å structure (oxygen = red, nitrogen = dark blue, sulfur = green). Sulfate ions are shown bound beneath the helix α8 (yellow) GGDD motif (906) and at the likely F16bP 1-phosphate binding site (902). The PKTIDGD motif (residues 202-208) is shown in blue and the GGPAPG loop in purple. Phe218A and Phe138A, which form van der Waals contacts with Lys203A are shown in green. Figured prepared with MOLSCRIPT (Kraulis 1991) and Raster3D (Merrit and Bacon 1997).

Sulfate	Atom	Residue	Atom	Distance Å
β-phosphate				
Subunit A				
906	O2	S178A	N	2.88
906	O2	S178A	OG	3.04
906	O2	G175A	N	3.07
906	O2	X508		3.03
906	O1	X508		2.51
906	O1	X531		3.23

Table 4.9: Contacts made by the sulfate ion (906) bound beneath the subunit A GGDD motif of a8 of the *S. thermophilum* PP_i-PFK 1.85 Å structure.

The general architecture of the ATP binding site of *E. coli* ATP-PFK is conserved in *S. thermophilum* PP_i-PFK. The orientation and binding of the pyrophosphate α- and β-phosphate groups at the base of the helix α8 GGDD motif is likely to be similar to that of the β- and γ-phosphates of ATP beneath the helix α5 GGDG motif of *E. coli* ATP-PFK. A sulfate ion is observed bound at the PP_i-binding site of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model in subunit A only (see Figure 4.24 and Table 4.9). Based upon superposition of the PP_i-binding site (helix α8) with the subunit D ATP-binding site (helix α5) of *E. coli* ATP-PFK (1PFK) the sulfate ion occupies the same position beneath the GGDD motif as the β-phosphate of ADP (see Figure 4.25a). A sulfate ion is observed in a similar orientation beneath the α8 GGDD motif of both subunits of the *B. burgdorferi* PP_i-PFK structure (Moore *et al.* 2002). The orientation and conformation of the GGDD motif of the *S. thermophilum* PP_i-PFK 1.85 Å and 2.2 Å structures is distinct relative to that of *B. burgdorferi* PP_i-PFK and *E. coli* ATP-PFK (see Figure 4.25b). Superposition of the subunit A N-terminal domain of *B. burgdoferi* PP_i-PFK (residues 70-147, 197-210, 432-462) and *E. coli* ATP-PFK illustrates that the GGD(D/G) and P(K/G)TID(G/N)D motifs superimpose (see Figure 4.25c). Recall that the subunit A GGDG motif of *E. coli* ATP-PFK is displaced into the active site, relative to subunit D via a shearing of helix α5 into the active site (Figure 1.5d). This was concluded by Shirakihara and Evans (1988) to be physiologically relevant due to the resultant closer proximity of the ADP β-phosphate group to the F16bP 1-phosphate group (Shirakihara and Evans 1988). It is interesting to note that asymmetry between the subunits of *B. burgdorferi* PP_i-PFK involves a similar displacement of the subunit A helix α8, which displaces the GGDD motif into the active site, relative to that of subunit B. Despite the fact that the motifs of the respective A subunits superimpose, the mechanism of displacement is distinct. Displacement of the subunit A GGDD motif of *B. burgdorferi* PP_i-PFK into the active site is stringently linked to the orientation of the β14-β15 β-hairpin insert, as discussed below. The α–helix of group II long clade PFKs (see Figure 1.12) homologous to helix α8 of *S. thermophilum* PP_i-PFK has the consensus sequence; GGD(G/D)(S/T)x(G/x)AxxLAE (see Figure 1.12). The underlined residues putatively contribute to phosphoryl donor specificity. The C-terminus of helix α8, which is

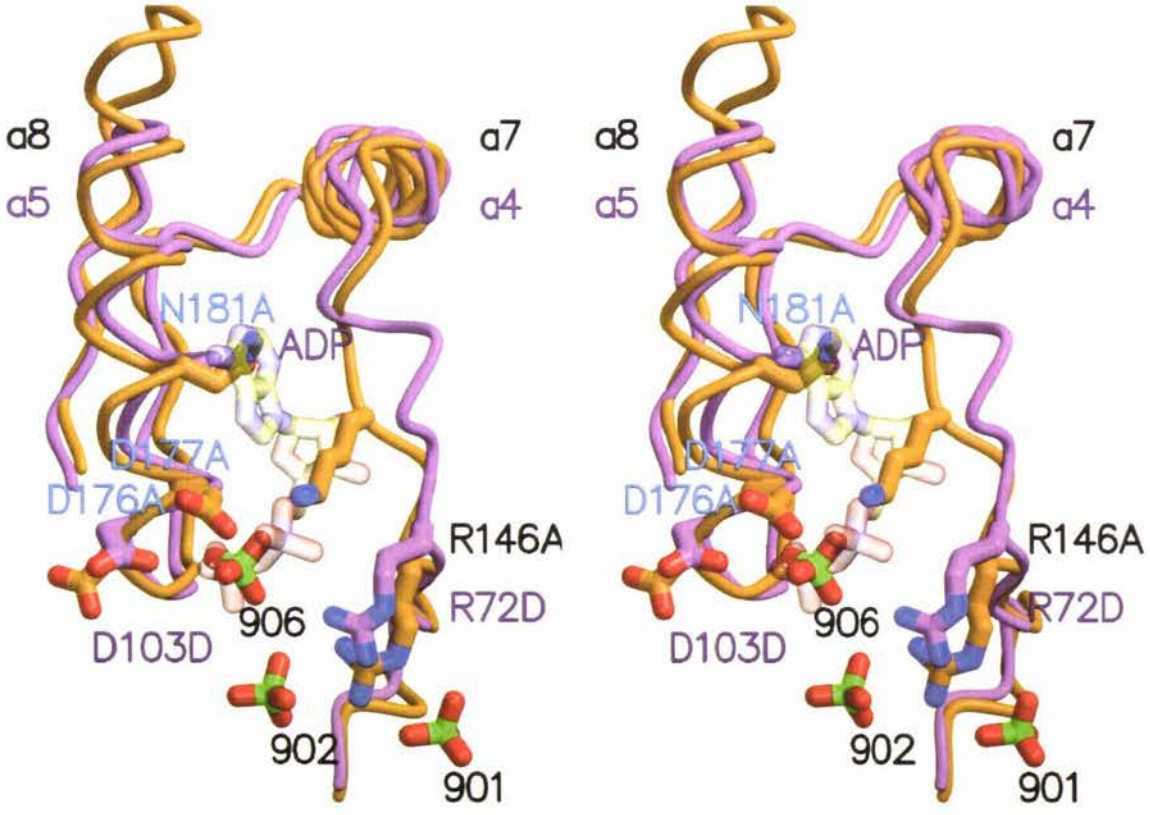
extended by a complete helical turn via a six-residue insert relative to the ATP-PFK of *E. coli* (see Figure 4.2) forms extensive contacts with helix α3 of the N-terminal insertion (see Figure 4.25d). In the *S. thermophilum* PP_i-PFK atomic models Glu187 OE1 (α8) forms hydrogen bonds with Thr53 N and OG (α3). Glu187 OE2 (α8) forms a hydrogen bond with Asn52 ND2 (α3) as does Gln191 OE1 (α8). Phe50 (α3) forms van der Waals contacts with Ala183 (α8) and Leu184 (α8). The N-terminus of helix α8 packs against the base of the β14-β15 β-hairpin. Thr180 (α8) forms hydrophobic van der Waals contacts with Pro495, Val496, and Ile497 of β15. Leu480 and Met484 form van der Waals contacts with Ala183 (α8). Ala183 is almost exclusively conserved in PFK A PFKs. β14 forms hydrophobic contacts with the N-terminus of helix α3 of the N-terminal insert. Leu46 (α3) forms van der Waals contacts with Met484, Asn485, and Met486 (see Figure 4.25d). The hydrophobic packing between helix α3 of the N-terminal insert, the β14-β15 β-hairpin, and helix α8 suggests that movement of helix α8 is coupled to conformational changes of the N-terminal insert, and the β14-β15 β-hairpin. Therefore it is suggested that the observed displacement of the subunit A β14-β15 β-hairpin of the *B. burgdorferi* PP_i-PFK (1KZH) structure, relative to subunit B, is coupled to the observed displacement of helix α8, which displaces the GGDD motif into the active site (see Figure 4.25c). Displacement of helix α3, the N-terminus of helix α8, and the β14-β15 β-hairpin is concomitant to conserve the contacts. Superposition of the A and B subunits from the respective *S. thermophilum* PP_i-PFK atomic models based upon N-terminal core (residues 70-147, 197-210, 432-462) illustrates that the conformation of the N-terminal insert (residues 1-70) differs significantly with the β14-β15 β-hairpin due to the conserved hydrophobic packing of helix α3 and the β14-β15 β-hairpin. The β-hairpin is not displaced into the active site in either subunit therefore neither is the N-terminus of helix α8. Superposition of the four subunits from the *S. thermophilum* PP_i-PFK 1.85 Å and 2.2 Å atomic models illustrates no structural deviation of the Cα backbone for the N-terminus of helix α8 (see Figure 4.26a).

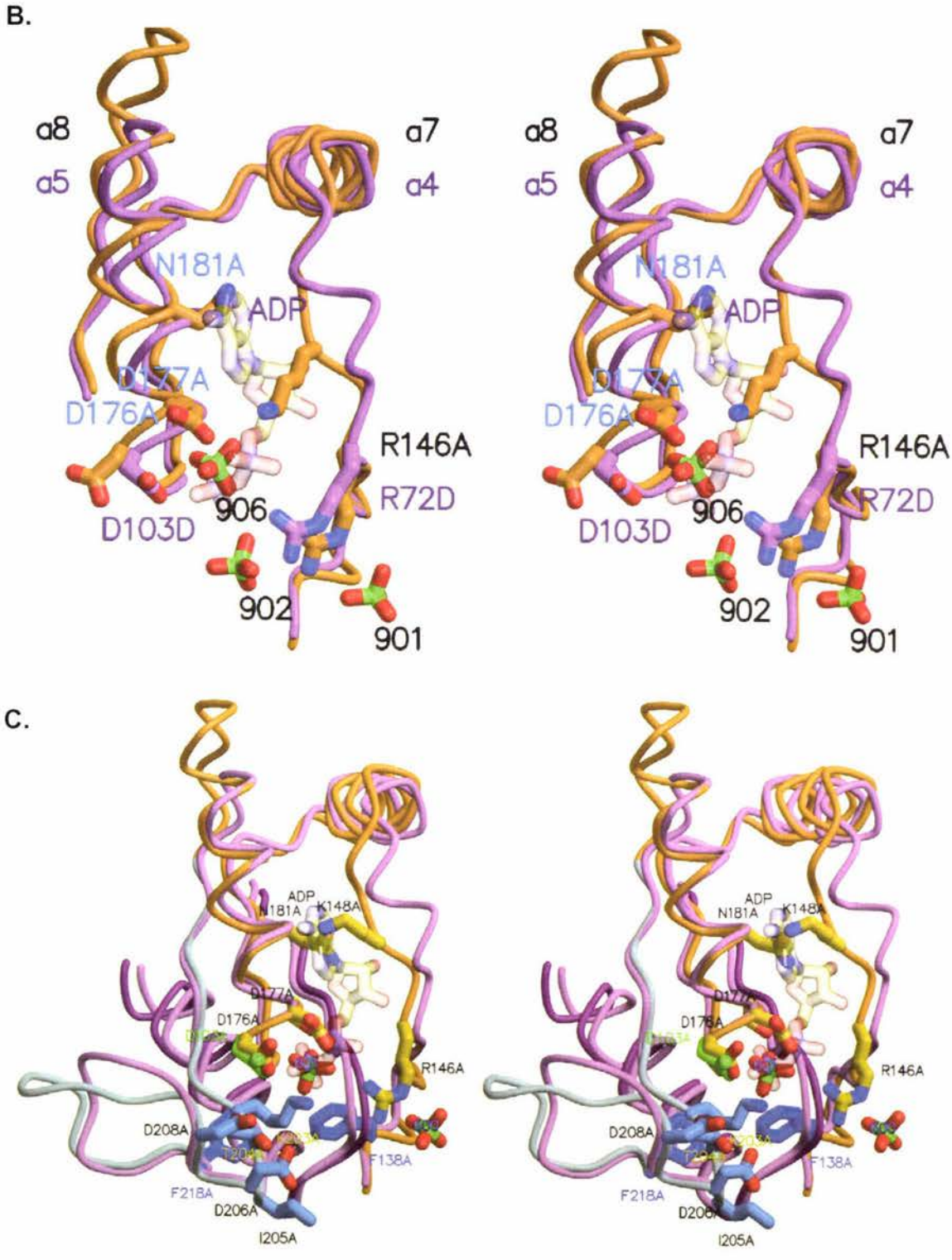
The displacement of the subunit A GGDD motif of *B. burgdorferi* PP_i-PFK into the active site, relative to subunit B probably results from the conserved hydrophobic van der Waals contacts between strand β15 and the N-terminus of helix α8. To conserve the

contacts and to prevent unfavourable steric interactions the GGDD motif is displaced concomitantly with the β14-β15 β-hairpin into the active site. The conserved hydrophobic packing of the N-terminal insert; helix α3 with the LAE tripeptide of helix α8 and strand β14 dictates that the conformational deviation is translated through to the N-terminal insert.

Up to this point the conformation of the C-terminal and helical domains, the N-terminal insert, and the β14-β15 β-hairpin insert, relative to the ATP-PFK of *E. coli* have been suggested to elicit a direct effect upon the conformation of essential motifs at the active site, conserved in all PFK A PFKs.

A.




B.
α8
α7
α5
α4
N181A
ADP
D177A
D176A
R146A
R72D
906
D103D
902
901
α8
α7
α5
α4
N181A
ADP
D177A
D176A
R146A
R72D
906
D103D
902
901
C.
ADP
N181A
K148A
D177A
D176A
R146A
D208A
F138A
F218A
D206A
I205A
ADP
N181A
K148A
D176A
R146A
D208A
F138A
F218A
I205A

D.

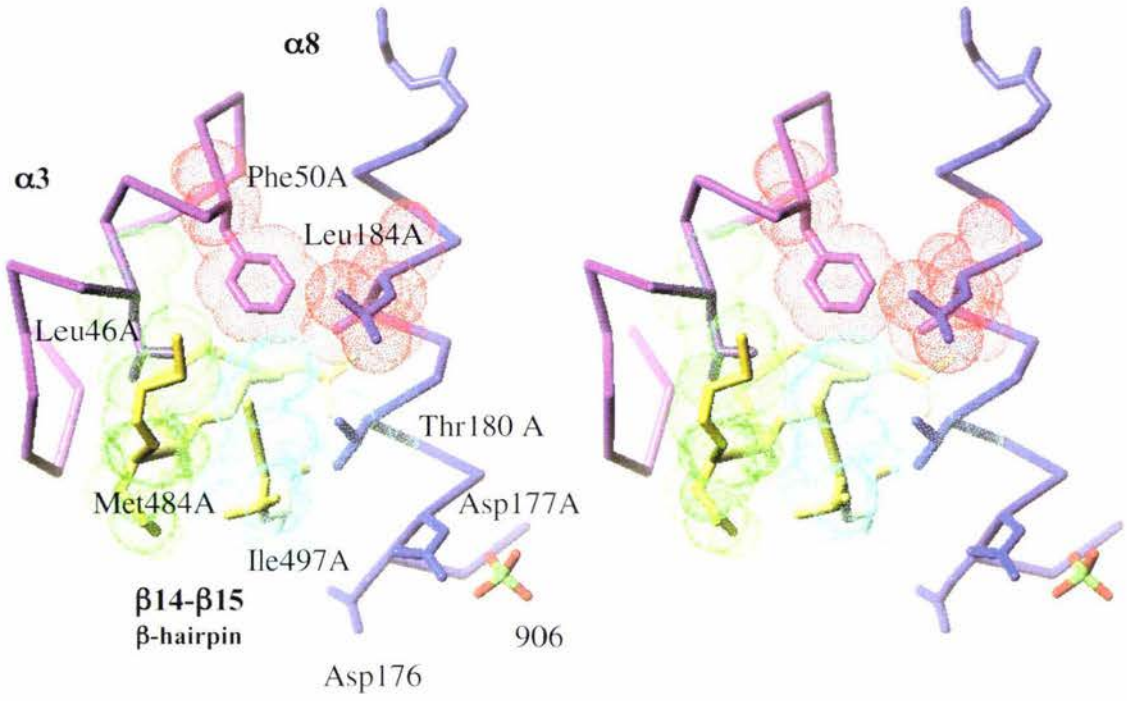


Figure 4.25:PP_i-binding site of *S. thermophilum* and *B. burgdorferi* PP_i-PFKs.

A. Structural overlay of the subunit A PP_i-binding site of the *S. thermophilum* PP_i-PFK 1.85 Å structure (orange) with the subunit D ATP-binding site of *E. coli* ATP-PFK (1PFK [purple]) based upon the N-terminal domain. Note that the sulfate ion bound beneath the helix α8 GGDD motif (906) superimposes with the ADP β-phosphate. Note that the protrusion of the α6-α7 loop and Asn181A into the hydrophobic groove constrict the ATP binding site. **B.** Structural overlay of the subunit A PP_i-binding site of the *S. thermophilum* PP_i-PFK 1.85 Å structure (orange) with the subunit A ATP-binding site of *E. coli* ATP-PFK (1PFK [purple]) based upon the N-terminal domain. Note that the sulfate ion bound beneath the helix α8 GGDD motif (906) does not superimpose with the ADP β-phosphate due to displacement of the subunit A ATP-PFK helix α5. **C.** Structural overlay of the subunit A PP_i-binding site of the *B. burgdorferi* PP_i-PFK 1.85 Å structure (orange and light blue) with the subunit A ATP-binding site of *E. coli* ATP-PFK (1PFK [purple]) based upon the N-terminal domain. Note that the sulfate ion bound beneath the helix α8 GGDD motif (601) superimposes with the ADP β-phosphate due to the displacement of the PP_i-PFK helix α8, analogous to that of the ATP-PFK helix α5. D. Cα diagram of the *S. thermophilum* PP_i-PFK 1.85 Å structure subunit A helices α3 and α8 and the β14-β15 β-hairpin illustrating van der Waals contacts between α8 and the α3 (red), α8 and the β14-β15 β-hairpin (light blue), and α3 and the β14-β15 β-hairpin (light blue). Note, only selected contacts are shown for clarity. Figure A., B. and C. prepared with MOLSCRIPT (Kraulis 1997) and Raster3D (Merrit and Bacon 1997). Figure D. prepared with TURBO-FRODO (Roussel and Cambillau 1991).

4.10.1 Phosphoryl donor Specificity

Despite the hydrophobic environment of the groove formed between the α6-α7 loop, helix α7, and helix α8 of *S. thermophilum* PP_i-PFK, ATP binds with very low affinity. Only background activity is observed with ATP as the phosphoryl donor (Ronimus *et al.* 2001). Modeling of ADP into the α7-α8 groove identifies possible unfavorable steric interactions between the side chains of Asn181 and Lys148 and the adenine and ribose moieties of ADP (see Figure 4.25a). Asn181 aligns with Gly108 of *E. coli* ATP-PFK (see Figure 4.2), which is highly conserved in group I and III PFKs to accommodate the adenine ring. Asn181 is not highly conserved in group II family members though homologues are exclusively larger than a Gly. Lys148 is highly conserved in group II, long clade family members (see Figure 1.12). The only exception is *C. trachomatis* PFK, which exhibits an Asn at this position. Interestingly *C. trachomatis* PFK exhibits the typically ATP-dependent GGDG motif indicating alternative phosphoryl donor specificity. The α6-α7 loop of *S. thermophilum* PP_i-PFK has an altered conformation relative to the homologous α(3)-α4 loop of *E. coli* ATP-PFK (see Figure 4.25a), which contains a single residue insertion (see Figure 4.2). The altered conformation of the loop constricts the hydrophobic groove, contributing to the discrimination against ATP (Moore et al. 2002). The Cα of Lys148 is 1.5 Å from the O2* hydroxyl group of the ribose moiety of ADP modeled into the active site based upon superposition of the N-terminal domain β-sheet and helices α7 and α8 (see Figure 4.24a). Closure of the hydrophobic groove concomitant with the presence of Asn181 and Lys148 may function to prescribe phosphoryl donor specificity. Recall from chapter 1 (section 1.9.5) that the group II, long clade PP_i-PFK of the amitochondriate protist *E. histolytica* exhibits Asn179 and Lys146 within the putative hydrophobic groove, homologous to Asn181 and Lys148, respectively and does not exhibit the insertion of group I and group III family members (see Figure 4.2). Mutation of the PP_i-dependent motif GGD**D** of *E. histolytica* PP_i-PFK by site-directed mutagenesis to the ATP-dependent motif (GGD**G** [Asp175Gly]) abrogated k_{cat} for pyrophosphate by 4-orders of magnitude concomitant with a 4000-fold increase in k_{cat} with ATP as the phosphoryl donor at low ATP concentrations (Chi and Kemp 2000). The single mutation altered the enzyme preference (defined as $[k_{cat}/K_M]_{PPi}/[k_{cat}/K_M]_{ATP}$) by 10^7! Structure-based sequence alignment of group II long

clade family members (see Figure 4.2) fails to identify significant variation to explain this seemingly anomalous result leading to the suggestion that Asn179 and Lys146 are not the major proponent's in the mitigation of ATP activity. Sequence identity (55.8% between *S. thermophilum* PP_i-PFK and *E. histolytica* PP_i-PFK) allows the tentative extrapolation of this result to all group II long clade family members. From Chi and Kemp's (2000) result it is suggested that Asp175 of the GGD**D** motif is of primary importance in phosphoryl donor discrimination. Recall from section 1.6.2 that the α-phosphate O1 oxygen atom of ADP bound at the active site of subunit A of *E. coli* ATP-PFK (1PFK [Shirakihara and Evans 1988]) coordinates a water molecule (X60) to the essential magnesium cation (see Figure 1.6b). Superposition of the subunit A active site of *B. burgdorferi* PP_i-PFK with that of *E. coli* ATP-PFK (1PFK) demonstrates that the OD2 oxygen atom of the pyrophosphate-specific Asp177 (GGD**D**) overlays with the α-phosphate O1 oxygen atom of ADP, indicating that Asp177 (GGD**D**) likely functions in the indirect coordination of the essential magnesium cation (Moore *et al.* 2002).

4.10.2 GGDD motif

Despite the conserved backbone conformation of helix α8, particularly the N-terminus between the *S. thermophilum* PP_i-PFK atomic models, the conformation of the side chains of Asp176 and Asp177 (GG**DD**) are distinct (see Figure 4.26a). In both subunits of the *S. thermophilum* PP_i-PFK 1.85 Å structure the carboxylate group of Asp176 (GG**D**D) is rotated away from the active site and contacts the peptide oxygen and carboxylate oxygen atoms of Asp208 (see Figure 4.24). The Asp176A carboxylate oxygen atom, OD2 is positioned 2.76 Å from the peptide oxygen atom of Asp208A (PKTIDG**D**) and 3.13 Å from the Asp208A carboxylate oxygen atom, OD2 . The van der Waals radii of oxygen is 1.52 Å, which suggests that the first contact (2.76 Å) is a hydrogen bond resultant from protonation of Asp176A's carboxylate oxygen atom, OD2. The crystallization pH of 4.6 is conducive with protonation (pK_a of an Asp residue is 3.5). If OD2 of Asp176A is protonated then it is permissible for the 3.13 Å contact with OD2 of Asp208A to be either a van der Waals contact or a hydrogen bond. The conformation of the carboxylate group of Asp177 (GGD**D**) is highly disordered in both subunits, particularly in subunit A. Positive 2σ difference electron density ($|F_o| - |F_c|$) peaks denote

that Asp177A may exist in alternate conformations; the carboxylate group rotated into and away from the PP_i-binding site. Despite modeling of the alternate conformations, both with occupancies of 0.5 there is sparse 1σ $2|F_o|$ - $|F_c|$ electron density (see Figure 4.26b). The hydroxyl group of Ser178A also exists in an alternate conformation. The carboxylate group of Asp177B is rotated into the PP_i-binding site.

Asp176 in both subunits of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model is oriented into the active site, adopting a conformation similar to its homologue *E. coli* ATP-PFK, Asp103 (GG**D**G). In both subunits OD2 of Asp176 forms a salt-bridge contact with NZ of Lys499 (see Figure 4.26a). Lys499 is positioned upon the $\beta15-\alpha21$ loop, which proceeds the $\beta14$-$\beta15$ β-hairpin (see Figure 4.2). Lys499 is highly conserved in group II, long clade family members with the exception of *C. trachomatis* PFK (see Figure 1.12). The side chain of Asp177 is orientated out of the PP_i-binding site, towards the hydrophobic groove formed by the $\alpha6$-$\alpha7$ loop and helices $\alpha7$ and $\alpha8$. Asp177 OD2 forms a salt-bridge contact with NZ of the aforementioned conserved Lys148 (see Figure 4.26a). The altered conformation of the side chains of Asp176 and Asp177 (GG**DD**) in the *S. thermophilum* PP_i-PFK 1.85 Å atomic model result in neither forming salt bridge's with Lys499 and Lys148, respectively (see Figure 4.26a). Relative to the *S. thermophilum* PP_i-PFK 2.2 Å structure the concomitant deviation of the N-terminal insert and $\beta14$-$\beta15$ β-hairpin of the 1.85 Å atomic model results in a slight displacement of Lys499.

Interestingly Asp103 (GG**D**G) in all four subunits of the T-state *B. stearothermophilus* ATP-PFK tetramer (6PFK [Schirmer and Evans 1990]) and Asp103 of the unliganded *E. coli* ATP-PFK (2PFK[Rypniewski and Evans 1989]) structure adopt a similar conformation to Asp176 of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model. The carboxylate oxygen atom OD2 of Asp103A (6PFK) forms a van der Waals contact (3.06 Å) with the peptide oxygen atom of Asp129A (PGTIDN**D**). The crystallization pH was 7.6 (Schirmer and Evans 1990).

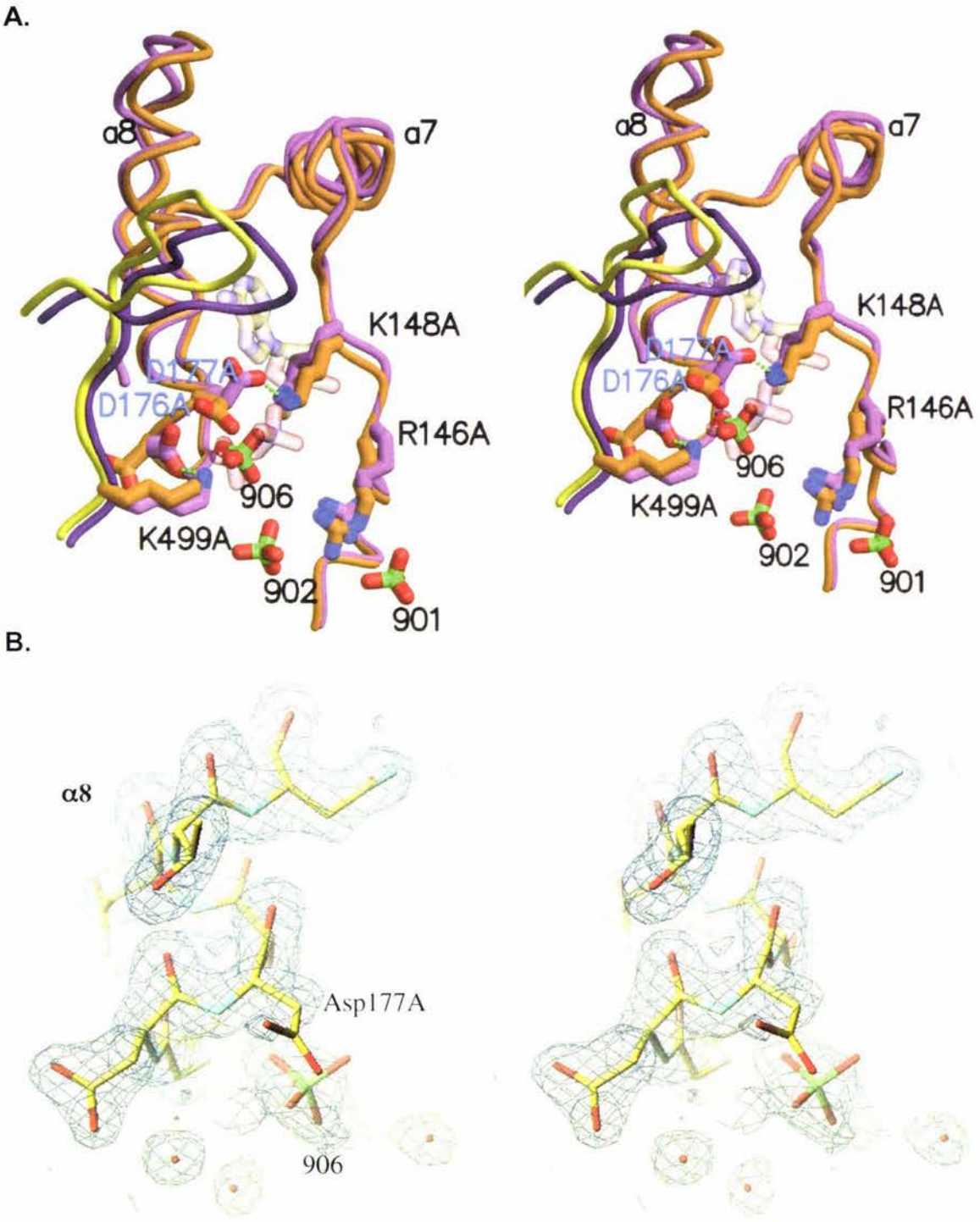


Figure 4.26: **A.** Superposition of subunit A from the *S. thermophilum* PP_i-PFK 1.85 Å (purple) and 2.2 Å (orange and yellow) structures. Note the distinct conformation of the GGDD motif (oxygen = red, nitrogen = blue, sulfur = green). Figure prepared with MOLSCRIPT (Kraulis 1991) and Raster3D (Merrit and Bacon 1997). **B.** Stereo view of the N-terminus of helix a8. Electron density ($|F_o| - |F_c|$) is shown contoured at 1.6s. Note that density is scarce for the side chain of Asp177A, denoting its disorder. Figure prepared with TURBO-FRODO (Roussel and Cambillau 1991).

4.10.3 *B. burgdorferi* PP_i-PFK: PP_i-binding site

Superposition of the N-terminal domain of subunits A and B (see Table 4.2a) identifies movement of the N-terminus of helix α8 of subunit A by approximately 1.0 Å relative to subunit B. As outlined earlier the shift in helix α8 of subunit A is concomitant with movement of the β14-β15 β-hairpin (Moore *et al.* 2002).

Asp 176 and Asp177 of both subunits are rotated into the active site (see Figure 4.27). Movement of the helical domain 380-390 β-hairpin within subunit A positions the imidazole ring of His384A into the base of the hydrophobic groove, adjacent to the N-terminus of helix α8, where it is sandwiched between and forms van der Waals contacts with Asp177A and the aliphatic chain of Lys148A; van der Waals contacts exist between the imidazole ring of His384A and Cβ of Asp177A and between the Cβ of His384A and Cγ of Asp177A. Asn386A ND2 of the 380-390 β-hairpin forms a hydrogen bond with OD2 of Asp177A (see Figure 4.27). The packing of the His384A against Asp177A causes minor deviations in the conformation of the Asp177A side chain, relative to subunit B. In both subunits the carboxylate groups of the two Asp residues of the GGDD motif point towards each other. The carboxylate group of Asp177B is rotated slightly higher than Asp176B. OD1 of Asp177B is 3.55 Å from OD2 of Asp176B. In subunit A the carboxylate group of Asp177A is in closer proximity to Asp176A. Asp177A OD1 forms a van der Waals contact with OD1 of Asp176A (3.27 Å). Asp177A OD2 is 4.38 Å from OD2 of Asp176A, both of which coordinate a water molecule, X794, which is further coordinated by the O2 and O4 oxygen atoms of the bound sulfate ion (601). *B. burgdorferi* PP_i-PFK was crystallized from 2M ammonium sulfate and 100 mM Tris, at pH 8 (Moore *et al.* 2002) therefore it is highly unlikely that the carboxylate groups are protonated. The electrostatic repulsion arising from the proximity of the Asp176 and Asp77 carboxylate groups may be offset by the close proximity of positively charged side chains; Lys501 (Lys499 in *S. thermophilum* PP_i-PFK), Lys148, and Arg491 (Arg489 in *S. thermophilum* PP_i-PFK) (see Figure 4.27). Arg491 is positioned upon the β14-β15 β-hairpin and is highly conserved in group II long clade family members (see Figure 1.12). Movement of the β-hairpin into the active site positions the guanidinium group of Arg491 directly above the respective OD1 atoms of Asp176 and Asp177 (see Figure 4.27).

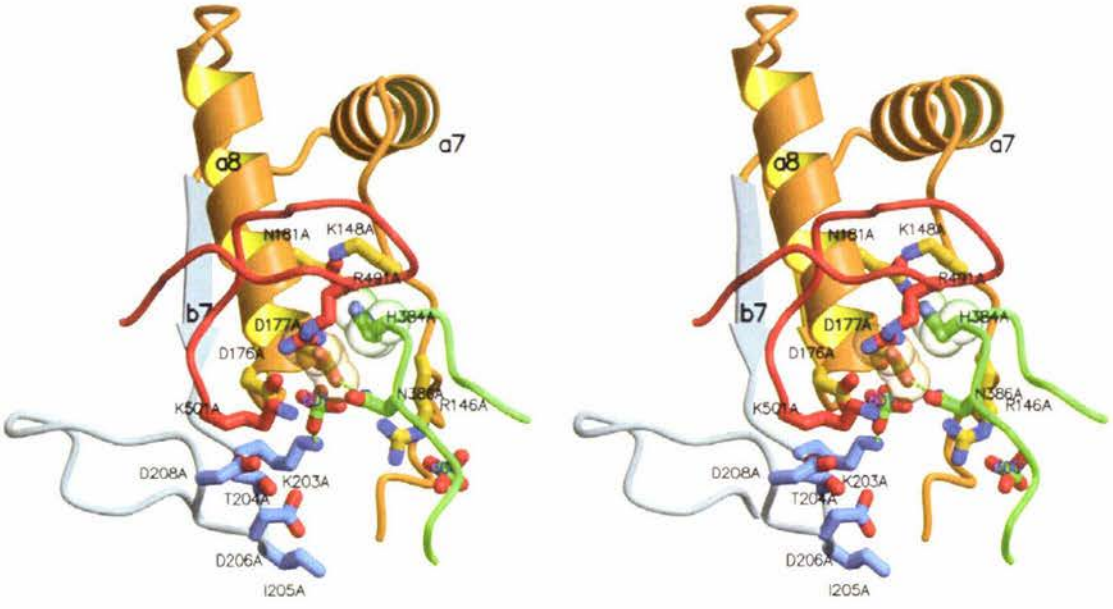


Figure 4.27: Ribbon diagram of the subunit A PP_i-binding site of *B. burgdorferi* PP_i-PFK (1KZH). Note the displacement of the helical domain 380-390 β-hairpin (green) into the active site positions the imidazole ring of His384A into the base of the hydrophobic groove where it forms van der Waals contacts with Asp177A and Lys148. Its orientation prevents Asp177 from rotating out of the PP_i-binding site. Note the displacement of the β14-β15 β-hairpin (red) positions Arg491A above the carboxylate groups of Asp176A and Asp177A. Note also that Lys203A forms a salt-bridge with the sulfate ion (601) bound at the N-terminus of α8, beneath the GGDD motif (oxygen = red, nitrogen = dark blue, sulfur = green). Figure prepared with MOLSCRIPT (Kraulis 1991) and Raster3D (Merrit and Bacon 1997).

The altered conformation of the GGDD motifs of the *S. thermophilum* PP_i-PFK atomic models is hypothetically resultant from the absence of the 380-390 β-hairpin at the active site (see Figure 4.26). Specifically, the absence of the van der Waals contact between His382 and Asp177 (GGD**D**) permits Asp177 to rotate away from Asp176, out of the PP_i-binding site. The conformation of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model GGDD motifs may represent the most thermodynamically stable conformation due to the salt-bridge contacts between Asp176, Asp177 and Lys499 and Lys148, respectively concomitant with the relieved electrostatic repulsion (see Figure 4.26a). If this is the case why is Asp177B of *B. burgdorferi* PP_i-PFK not rotated away from the PP_i-binding site if both are deprotonated? The side chains of Asp177B and Lys148B of *B. burgdorferi* PP_i-PFK are disordered relative to their subunit A counterparts (see Table 4.10). This indicates that they are relatively free to move within the crystal. The conformation of Lys148B prevents formation of the salt bridge observed in the *S. thermophilum* PP_i-PFK 2.2 Å model. The sulfate ion may also contribute a degree of electrostatic repulsion to the orientation of the carboxylate group into the hydrophobic groove. Asp177 of the GGD**D** motif of both subunits of the *S. thermophilum* PP_i-PFK 1.85 Å is also highly disordered, as previously mentioned. The subunit A conformation of the *B. burgdorferi* PP_i-PFK GGDD motif is likely to represent the 'active' conformation of PP_i-PFKs, coordinating a Mg^{+2} cation to PP_i-PFK bound below the N-terminus of helix α8.

Structure based sequence alignment identifies a three residue insertion within the 380-390 β-hairpin of the *B. burgdorferi* PP_i-PFK relative to that of *S. thermophilum* PP_i-PFK (see Figure 4.2) indicating that the helical domain, specifically the 380-390 β-hairpin of *S. thermophilum* PP_i-PFK will have to be displaced to a greater degree with respect to that of subunit A of the *B. burgdorferi* PP_i-PFK (1KZH [Moore *et al.* 2002]) if the conserved His382 is to make the same contacts with the GGDD motif.

	Main chain B average	Main chain average B rms	Side chain B average	Side chain average B rms	Whole chain
Chain A	25.903	1.088	32.008	2.216	30.929
Chain B	37.629	1.112	38.092	2.048	37.849

Table 4.10a: Average temperature factors for chain A and B of *B. burgdorferi* PP_i-PFK.

	Main chain B average	Main chain B rms	Side chain B average	Side chain B rms	Residue B average	Residue B rms
Asp177						
Subunit A	25	0.0305	31.8	0.0561	28.4	1.3696
Subunit B	34	0.0183	46.2	0.096	40.1	3.1706
Lys148						
Subunit A	37.2	0.0262	36.3	0.0451	36.7	1.3831
Subunit B	55.5	0.025	66.2	0.0361	63.2	2.0409

Table 4.10b: Average temperature factors for Asp177 (GGD**D**) and Lys148 of *B. burgdorferi* PP_i-PFK, indicating that they are disordered.

4.10.4 P**K**TIDGD motif: Lys203

The PKTIDGD motif of *S. thermophilum* PP_i-PFK is positioned at the C-terminus of strand β7 and the β7-α9 loop (see Figure 4.24). The function of Asp127 and Asp129 (PGTI**D**N**D**) of *E. coli* ATP-PFK (1pfk); general base and co-ordination of an essential Mg^{2+} cation, respectively is conserved at the sequence level in PP_i-PFKs. The motif in group I PFK family members exhibits an exclusively conserved Gly (P**G**TIDxD), whereas group II and III PFK family members exhibit an exclusively conserved Lys, P**K**TIDxD (see Figure 4.23b). Exclusive conservation of a Lys in group II PFK family members indicates that the side chain has a significant function. Exclusive conservation of a Gly in group I PFK family members is hypothetically selected for, either by unfavorable steric contributions of a larger residue i.e. the presence of a Cβ atom, or to allow a backbone conformation that would otherwise force unfavorable dihedral angles. It is important to note that a Lys residue is exclusively conserved in PP_i-PFKs, however a Lys residue is also observed in the group II x clade and group III ATP-dependent family members. Both of these clades exhibit an exclusively conserved Lys though they exhibit both ATP- and PP_i-PFKs.

The conformation of the PKTIDGD motif is conserved between *B. burgdorferi* PP_i-PFK and the *S. thermophilum* PP_i-PFK structures, therefore the following discussion will

primarily focus on subunit A of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model. The conformation of the PKTIDGD motif differs relative to that of *E. coli* ATP-PFK (1PFK). The difference arises to accommodate the side chain of Lys203A into the PP_i-binding site. The Cα of Lys203A is displaced by approximately 1.5 Å relative to the Cα of Gly124A. Thr204A (PK**T**IDGD) of *S. thermophilum* PP_i-PFK 1.85 Å atomic model has sterically strained dihedral angles (phi = 164, psi = 126), falling within a generously allowed region of a Ramachandran plot (see Figure 3.3 and 3.4). It forms similar contacts to Thr125A of *E. coli* ATP-PFK (1PFK) (PG**T**IDND [phi = -170, psi = 140]). The PKTIDGD motif loop contacts the N-terminus of helix α8 in *S. thermophilum* PP_i-PFK. The peptide nitrogen atom of Lys203A forms a hydrogen bond with the peptide oxygen atom of Gly174A (2.78 Å [see Figure 4.24]). Leu209A forms a van der Waals contact with Gly175A and Gly176A. Ile130A of *E. coli* ATP-PFK (1PFK) makes similar contacts. Asp208A (PKTIDG**D**) forms contacts with Asp176A in the *S. thermophilum* PP_i-PFK 1.85 Å atomic model, as previously stated (see Figure 4.24). Orientation of the Lys203A side chain into the PP_i-binding site is tightly constrained by a highly conserved loop; GGxxPG and two Phe residues; Phe218 and Phe139 (see Figure 4.24). Both are highly conserved in group II PFK long clade family members (see Figure 4.2). The aromatic ring of Phe218A packs against the Cα, Cβ, and Cγ of Lys203A, forming strong van der Waals contacts. The aromatic ring of Phe139 protrudes through the conserved GGPAPG loop, forming a van der Waals contact with Cγ of Lys203A, further constraining the orientation of its side chain (see Figure 4.24). NZ of Lys203A forms hydrogen bonds with the peptide oxygen atoms of Ser80A and Gly82A and a salt bridge with the O4 oxygen atom of a sulfate ion bound at the putative F16bP 1-phosphate binding site (see Figure 4.24).

Packing and orientation of Lys203 is conserved in the *B. burgdorferi* PP_i-PFK atomic model, however the interactions the Lys203A NZ group forms are not, due to the observed displacement of the *B. burgdorferi* PP_i-PFK subunit A GGDD motif. In the *B. burgdorferi* structure NZ of Lys203A forms a salt-bridge with the sulfate ion (601) bound beneath the GGDD motif (see Figure 4.27). Comparison of the orientation of sulfate ions bound at the active site of the *S. thermophilum* PP_i-PFK 1.85 Å and *B. burgdorferi* PP_i-PFK atomic models illustrates the putative function of Lys203 (P**K**TIDGD) in the

coordination and stabilization of the penta-coordinate transition state. It may also function in correct orientation of PP_i and/or a proton-shuttle. It has been suggested that the function of Lys203 in PP_i-PFKs and the requirement of ATP-PFKs for NH_4^+ or K^+ cations may be related (Moore *et al.* 2002). Ronimus *et al.*(1999) has demonstrated the enhancement of activity of *S. thermophilum* PP_i-PFK when either Na^+ or K^+ cations are present at 100mM (115% of control activity). Bertagnolli and Cook (1994) propose that PP_i-PFKs require two metal cations for catalysis, one for chelation of PP_i and the second as the essential activator. Such a requirement has not been documented for *B. burgdorferi* PP_i-PFK. The small enhancement in activation may be an artifact of *S. thermophilum* PP_i-PFKs thermophilic nature compared to the mesophilic PP_i-PFK of *B. burgdorferi*. Recall that neither subunit of the *B. burgdorferi* PP_i-PFK (1KZH) structure exhibits a sulfate ion bound at the putative F16bP 1-phosphate binding site, however both exhibit sulfate ions bound at the F6P/F16bP 6-phosphate binding site and below the GGDD motif, at the β-phosphate binding site (Moore et al. 2002). The displacement of helix α8 and therefore of the subunit A GGDD motif of the *B. burgdorferi* PP_i-PFK (1KZH) atomic model and concomitantly the sulfate ion bound beneath, resultant from the orientation of the β14-β15 β-hairpin into the active site, would prohibit a sulfate ion binding at the likely F16bP 1-phosphate binding site due to unfavorable electrostatic and steric interactions with the sulfate ion bound at the β-phosphate binding site. The relative displacement of the subunit B GGDD motif would accommodate a sulfate ion at the putative F16bP 1-phosphate binding site without unfavorable interactions, as seen in subunit A of the *S. thermophilum* PP_i-PFK 1.85 Å structure. The distinct conformations of the active sites of the *S. thermophilum* PP_i-PFK 1.85 Å and *B. burgdorferi* PP_i-PFK structures, with respect to the orientation of bound sulfate ions is perplexing. It is hypothesized that the active sites of the respective models represent different stages in the catalytic mechanism of phosphoryl transfer, the bound sulfate ions acting as substrate/product analogues. The spatial orientation of the three sulfate ions bound at the subunit A active site of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model is analogous to the spatial orientation of the β-phosphate and F16bP 1- and 6-phosphate moieties of the forward reaction products. The spatial orientation is distinct from the ligands bound at the active site of subunit A of *E. coli* ATP-PFK due to the displacement of helix α5 altering the orientation of the ADP

β-phosphate (see section 1.5, Figure 1.5d). The displacement of the GGDD motif, relative to subunit A of *B. burgdorferi* PP_i-PFK, the absence of the 380-390 and β14-β15 β-hairpins at the active site, and the superposition of the three sulfate ions with ADP and F16bP bound at the subunit D active site of *E. coli* ATP-PFK (1PFK) all support the hypothesis that the conformation of the subunit A active site of the *S. thermophilum* PP_i-PFK 1.85 Å atomic model is representative of the enzymatic state subsequent to phosphorylation of F6P i.e. products, F16bP and P_i are bound in a conformation associated with dissociation. The spatial orientation of the ADP β-phosphate and the F16bP 1-phosphate, bound at the subunit D active site of *E. coli* ATP-PFK (1PFK) do not superimpose with the analogous sulfate ions bound at either active site of the *B. burgdorferi* PP_i-PFK atomic model, further illustrating its distinct conformation, relative to the *S. thermophilum* PP_i-PFK 1.85 Å atomic model. The observed asymmetry between subunits in all three PP_i-PFK atomic models suggests that catalysis at both active sites is not simultaneous. Whilst one active site is 'closed' for phosphoryl transfer the other is 'open' for dissociation and association of products and reactants, respectively.

4.10.5 GGPAPG loop

The GGPAPG loop is highly conserved in PFK A family members, as is its conformation based upon the structures accessible through the PDB and those of *S. thermophilum* PP_i-PFK. The loop contributes to the orientation of two critical Arg residues into the active site; Arg146 and Arg429. In *S. thermophilum* PP_i-PFK the GGPAPG loop is positioned between strand β3 and helix α4. In the tertiary structure it packs beside the PKTIDGD motif loop, beneath the GGDD motif of the N-terminus of helix α8. Pro85 (GGPA**P**G) loop forms van der Waals contacts with the planar peptide group of Arg429 (FGYEA**R**) and Pro433, which immediately precedes the motif. A residue larger than Gly81 and Gly82 would interfere with the conformation of Arg146 (Arg72 in *E. coli* ATP-PFK) and the PP_i-binding site.

4.12 Summary: *S. thermophilum* PP_i-PFK

The group II, long clade β_2 PP_i-PFK of *S. thermophilum* is surprisingly complex with respect to the prokaryotic ATP-PFK of *E. coli*. Conformational changes required,

presumably for catalysis directly involve conformational changes of the three major insertions, relative to the ATP-PFK of *E. coli*, which directly effect the conformation of conserved motifs involved in catalysis and/or substrate binding. A synopsis of this discussion is given in the following chapter.

Chapter 5

Spirochaeta thermophilum PP_i-PFK: Synopsis

5.1 Synopsis: Data Quality and Atomic Model Statistics

The structure of a non-allosteric, PP_i-dependent group II, long clade family member from the thermophilic bacterium *S. thermophilum*, a homologue of the plant PP_i-PFK β-subunit has been determined in two distinct conformations at 1.85 Å and 2.2 Å resolution. The 1.85 Å data set was collected from a crystal grown from a mother liquor containing 2 M ammonium sulfate ($[NH_4]_2SO_4$) and 0.1 M sodium acetate trihydrate (pH = 4.6). The 2.2 Å data set was collected from a crystal grown from a mother liquor containing 2 M ammonium sulfate ($[NH_4]_2SO_4$), 0.1 M sodium chloride (NaCl), and 0.1 M HEPES (pH = 7.5). Molecular replacement was employed to determine initial phases for the calculation of $2m|F_o| - D|F_c|$ and $m|F_o| - D|F_c|$ electron density maps for both models. The structures were refined to R factors of 0.1923 (R_{free} = 0.2035) and 0.1991 (R_{free} = 0.2288), respectively. The *S. thermophilum* PP_i-PFK atomic models exhibited high stereochemical quality.

5.2 Synopsis: Structure and Function

The 554 residue (M_R 61080 $g.mol^{-1}$) PP_i-PFK assembles into a homodimer (see Figure 4.1). The subunit tertiary structure is comprised of a conserved N- and C-terminal domain similar to the allosterically regulated, prokaryotic, tetrameric ATP-PFK of *E. coli*, of which the N-terminal domain exhibits a Rossmann-like nucleotide-binding fold. Three major insertions are observed; an N-terminal insertion, a C-terminal insertion, and an insertion within the C-terminal domain, which forms an autonomous α-helical domain. A smaller insertion of note forms the β14-β15 β-hairpin within the N-terminal domain (see Figure 4.2). The 'open' and 'closed' subunit asymmetry of the *S. thermophilum* PP_i-PFK

1.85 Å atomic model mirrors that of *B. burgdorferi* PP_i-PFK with the exception that the 390-390 and β14-β15 β-hairpins of subunit A are not displaced into the active site. Relative to subunit B of the *B. burgdorferi* PP_i-PFK atomic model both subunits of the *S. thermophilum* PP_i-PFK 2.2 Å atomic model are in an 'open' inactive conformation. The conformational change between the 2.2 Å and 1.85 Å *S. thermophilum* PP_i-PFK atomic models involves partial active site closure of both subunits via independent rigid-body displacements of the small and helical domains. Despite the subunit asymmetry and the conformational change the dimer interface is symmetrical and conserved. The majority of contacts at the dimer interface are formed between the N-terminal domain and the C-terminal domain of the opposing subunit. Resultant of the conserved dimer interface is the conserved orientation of the N-terminal domain of both subunits relative to the C-terminal domain of the opposing subunit. Therefore the conformation and orientation of the C-terminal and helical domains of the homodimer are relatively independent. If the N-terminal domain and the C-terminal domain of the opposing subunit are thought of as rigid-bodies then displacement of the subunit A C-terminal domain relative to the subunit A N-terminal domain dictates that the N-terminal domain of subunit B is displaced with the subunit A C-terminal domain to conserve the dimer interface i.e. the subunit B N-terminal domain is effectively rotating towards the subunit B C-terminal domain. Therefore closure of one active site, in terms of the N- and C-terminal domains dictates closure of the other to conserve the dimer interface.

The active site is formed by a cleft at the interface of the large and C-terminal domains. F6P/F16bP is bound predominantly by the conserved motifs, MGR (β8-α10 loop) and FG**Y**EAR (α[18]) of the C-terminal domain and Lys243 and Tyr244 from the C-terminal domain of the opposing subunit. PP_i is bound predominantly by the conserved N-terminal domain motifs, GGD**D** at the N-terminus of helix a8 and P**K**TIDGD (β7-α9 loop). Sulfate ions are observed bound beneath the GGDD motif, at the putative F16bP 1-phosphate binding site, and the F6P/F16bP 6-phosphate binding site in subunit A of the *S. thermophilum* PP_i-PFK 1.85 Å structure. A sulfate ion is only observed bound at the

F6P/F16bP 6-phosphate binding site of subunit B and both subunits of the *S. thermophilum* 2.2 Å atomic model.

5.2.1 F6P/F16bP binding site

Conservation of the dimer interface despite the conformational change dictates that the orientation of Lys243 and Tyr244 into the active site of the opposing subunit is conserved. The conformation of the MGR motif is stringently defined by the conformation of the small and helical domains. The inactive MGR conformation of both subunits of the *S. thermophilum* 2.2 Å atomic model, the side chain of Arg253 oriented out of the active site and the rotation of the Cε of Met251 into the F6P/F16bP binding site is resultant of the 'open' subunit tertiary structure. The conformational transition from the 2.2 Å *S. thermophilum* PP_i-PFK atomic model to that of the 1.85 Å atomic model, specifically the rotation of the C-terminal domain, displaces the MGR motif into the active site. Cε of Met251 is subsequently forced to rotate about the Cβ-Sδ bond to form a van der Waals contact with Tyr426 (FG**Y**EAR), similar to that observed in both subunits of the *B. burgdorferi* PP_i-PFK atomic model (1KZH [Moore et al. 2002]) and *E. coli* ATP-PFK (1PFK [Shirakihara and Evans 1988]) between Met169 (**M**GR) and His249 (LG**H**IQR). Tyr426 is exclusively conserved in group II PFKs. In neither of the three PP_i-PFK atomic models does it form a contact with the 6-phosphate sulfate ion, unlike its conserved group I and III analogue, which directly binds the F16bP (*E. coli* ATP-PFK [1PFK]) 6-phosphate moiety. Based upon the crystallographic evidence presented here and by Moore *et al.* (2002), concomitant with the kinetic studies of Deng *et al.* (2000) it is concluded that the major function of Tyr426 (FG**Y**EAR), specific to group II PFKs is the correct orientation of Met251 (**M**GR) into the active site.

The side chain of Arg253 (MG**R**), putatively functioning in transition state stabilization, is partially displaced with the C-terminal domain but an active conformation is not observed, requiring displacement of the 380-390 β-hairpin into the active side. The displacement of the subunit A 380-390 β-hairpin of *B. burgdorferi* PP_i-PFK (1KZH [Moore et al. 2002]) into the active site 'folds' the side chain of Arg253 into an active conformation. The conformation of the F16bP/F6P 6-phosphate binding FG**Y**EA**R** motif, specifically the α(18) helix is relatively independent of the conformational change.

However, of the three PP_i-PFK atomic models the subunit asymmetry in terms of C-terminal domain displacement relative to superposition of the N-terminal domain core is greatest for the *S. thermophilum* 1.85 Å atomic model, which translates to the greatest deviation in the orientation of the sulfate ions bound at the F6P/F16bP 6-phosphate binding sites. Resultant of the displacement is the distinct conformation of Arg146 (α6-α7 loop), which putatively functions in stabilization of the penta-coordinate transition state. The conformation of the F6P/F16bP binding site, and therefore hypothetically $K_{M[F6P/F16bP]}$ and k_{cat} (concurring kinetic evidence is absent to date) is stringently defined by the conformation of the small and helical domains, a mechanism which is likely to be unique to long clade PP_i-PFKs (based upon limited crystallographic evidence). There is no crystallographic evidence to suggest that the C-terminal domain and MGR motif of prokaryotic ATP-PFKs undergo such a conformational transition.

5.2.2 Pyrophosphate specificity

Asp177 of the PP_i-specific GGD**D** motif putatively functions in the indirect coordination of an essential Mg^{2+} cation. The conformation of the PP_i-binding site is subject to that of the helical domain and the β14-β15 β-hairpin. The conformation of the Asp176 and Asp177 carboxylate groups (GG**DD** motif) is distinct between *S. thermophilum* PP_i-PFK atomic models, hypothetically resultant from electrostatic repulsion. The orientation of the subunit A α-helical domain 380-390 β-hairpin of *B. burgdorferi* PP_i-PFK (1KZH [Moore *et al.* 2002]) into the active site positions His384A into the base of the α7/α6 hydrophobic groove, preventing Asp177 from rotating up, out of the PP_i-binding site. Displacement of the subunit A β14-β15 β-hairpin of *B. burgdorferi* PP_i-PFK into the active site is concomitant with displacement of the GGDD motif into the active site relative to the *S. thermophilum* PP_i-PFK atomic models. A similar displacement of the subunit A GGDG motif of *E. coli* ATP-PFK is observed, in fact the motifs superimpose based upon superposition of the N-terminal domain, yet the mechanism is distinct. Displacement prohibits a sulfate ion binding at the F16bP 1-phosphate binding site of the *B. burgdorferi* PP_i-PFK atomic model (1KZH) and results in distinct binding of the sulfate ion bound beneath the GGDD motif, relative to the *S. thermophilum* PP_i-PFK 1.85 Å atomic model. Displacement reorients the GGDD motif with respect to the

P<u>K</u>^{203}TIDGD motif, allowing Lys203 to form a salt bridge with the GGDD sulfate ion in both subunits of the *B. burgdorferi* PP_i-PFK atomic model (1KZH). The relative distant orientation of the GGDD and PKTIDND interdicts such a contact in the *S. thermophilum* PP_i-PFK 1.85 Å atomic model; rather Lys203 forms a salt bridge with the F16bP 1-phosphate sulfate ion. It is therefore concluded that Lys203 functions in the coordination and hence stabilization of the penta-coordinate transition state (associative mechanism).

5.2.3 GGDD

The GGDD motif is pivotal in determining phosphoryl donor specificity. Mutation of the PP_i-dependent motif GGD**D** of *E. histolytica* PP_i-PFK by site-directed mutagenesis to the ATP-dependent motif (GGD**G** [Asp175Gly]) altered the enzyme preference by 10^7. The converse mutation of the ATP-dependent GGD**G** (Gly104Asp) motif of *E. coli* ATP-PFK, abrogated ATP-dependent activity but did not cause an increase in PP_i-dependent activity (Chi and Kemp 2000). It is concluded that the lack of reciprocity is due to the absence of the 380-390 β-hairpin. The Gly104Asp mutation in *E. coli* abrogated ATP-dependent activity due to unfavorable electrostatic and steric interactions between Asp104 and the ATP α-phosphate. PP_i-PFK activity was not concomitantly gained, hypothetically due to conformational flexibility of the Asp103 and Asp104 (GG**DD**) carboxylate groups preventing the correct coordination of the Mg^{+2} cation and orientation of the PP_i. The correct conformation of the GGDD motif requires displacement of the 380-390 β-hairpin, specifically the imidazole ring of His382 into the hydrophobic groove. The Asp175Gly mutation in *E. histolytica* PP_i-PFK abrogated PP_i-PFK activity due to the incorrect coordination of PP_i. The mutation abrogated the requirement for the 380-390 β-hairpin to be displaced into the active site.

5.2.4 Conformational change: substrate analogues

It is suggested that the distinct conformations of the three PP_i-PFK atomic models is induced by the distinct orientation of sulfate ions bound at the active sites. The comparative structures of the *B. burgdorferi* PP_i-PFK and *S. thermophilum* PP_i-PFK 1.85 Å atomic models, crystallized at distinct pH (8.0 and 4.5, respectively) indicate that the conformational change is not pH dependent. It is noted that pH will effect sulfate ion, and

therefore substrate binding via dictation of the ionization state of specific residues and that distinct pH optima for *S. thermophilum* PP_i-PFK activity exist (glycolytic pH 5.0-6.0, gluconeogenic pH 7.0-7.5 [Ronimus *et al.* 1999]).

There is no kinetic evidence to suggest that the 'inactive' *S. thermophilum* PP_i-PFK 2.2 Å atomic model is representative of a physiologically relevant inactive enzymatic form, though the failure of extensive crystallization trials to co-crystallize *S. thermophilum* PP_i-PFK complexed with its glycolytic reactants/products supports this conclusion. The causative agent of the 'inactivation' is yet to be ascertained, the subject of further work.

Based upon the crystallographic evidence presented here and by Moore *et al.* (2002) the following model is proposed for the conformational mechanism of the glycolytic reaction. The transition from an 'open' inactive to an 'open' active enzymatic form involves partial closure of the active site, rotating the C-terminal domain MGR motif into the active site facilitating a van der Waals contact between Met251 (**M**GR) and Tyr426 (FG**Y**EAR). The partial closure of one active site dictates partial closure of the other to maintain the dimer interface. F6P and PP_i-PFK subsequently bind. The displacement of the 380-390 β-hairpin into the active site positions His384 into the hydrophobic groove, preventing electrostatic repulsion from forcing the carboxylate group of Asp177 to rotate out of the active site and folds Arg253 into an active conformation. The displacement of the β14-β15 β-hairpin into the active site concomitantly displaces the GGDD and hence PP_i. The Mg^{2+} is coordinated directly by Asp176 and indirectly by Asp177 (GG**DD**). Asp208 (PKTI**D**G**D**) partially ionizes the phosphoanhydride bond of PP_i. Displacement of the GGDD motif reorients the PP_i α-phosphate, bringing it into closer proximity with the nucleophilic 1-hydroxyl group, which is deprotonated by Asp206. Phosphoryl transfer proceeds via a penta-coordinate transition state (associative mechanism [S_{N2}]), which is stabilized by Arg253, Arg146, and Lys203. Once phosphoryl transfer is complete the 380-390 and β14-β15 β-hairpins are displaced out of the active site, concomitant with the reversion of the GGDD motif. Electrostatic repulsion between Asp206 and the F16bP 1-phosphate group facilitate dissociation. Both active sites are not closed at the same time. Whilst one is closed, facilitating phosphoryl transfer and preventing wasteful hydrolysis of PP_i, the other is open allowing products, F16bP and P_i to dissociate and F6P and PP_i to bind.

5.3 Future Directions

The conformational flexibility of the group II, long clade PP_i-PFK from the thermophilic bacterium, *S. thermophilum* is very complex, much more so than that of prokaryotic ATP-PFKs. The insertions, relative to the prokaryotic ATP-PFK of *E. coli* directly affect the conformation of conserved motifs involved in substrate binding and catalysis. The insertions therefore function to modulate the enzymatic mechanism.

5.3.1 Site directed mutagenesis

Determination of the structure of the *S. thermophilum* PP_i-PFK in two distinct conformations, active and inactive, respectively in concert with that of *B. burgdorferi* PP_i-PFK (1KZH [Moore et al. 2002]) identifies targets within the insertions for further site-directed mutagenesis studies.

- A His382Ala (380-390 β-hairpin) mutant will validate or dismiss the aforementioned conclusion that its displacement into the hydrophobic groove is essential for PP_i-specific activity.
- Independent mutation of the conserved Lys499 and Arg489 will elucidate what role, if any they play.
- A Glu316Gln/Glu316Ala mutation to investigate its function in the orientation of the MGR loop.

The supposition that the *S. thermophilum* PP_i-PFK 2.2 Å atomic model is in an inactive conformation needs to be verified by *in vitro* kinetic studies.

5.3.2 Substrate/Product binding

A future direction of this research is to fully elucidate the conformational mechanism associated with substrate binding and catalysis. This requires resolution of the *S. thermophilum* PP_i-PFK structure complexed with reactants, products, and substrate analogues. Based upon the hypothesis that prior crystallization trials to co-crystallize *S. thermophilum* PP_i-PFK complexed with its reactants/products failed due to an induced inactive conformation it is hypothesized that *S. thermophilum* PP_i-PFK crystals, already

grown, at a pH of ~4.5 and in the absence of sulfate or phosphate ions are complexed with F16bP and Mg^{2+}.P_i and await data collection.

5.4.3 PP_i-PFKs: Inhibition

The presence of a PP_i-PFK in many pathogenic anaerobic microorganisms targets the enzyme for the development of anti-microbial agents. This research has identified that the architecture of the active site and the mode of substrate binding are similar to that of prokaryotic ATP-PFKs. This hinders the cause for the development of a PP_i-PFK specific inhibitor. A prerequisite for drug development is no adverse *in vivo* kinetic effect upon mammalian ATP-PFKs. The most likely target therefore is the PP_i-binding site, which involves Lys203 and Asp177, unique to group II PP_i-PFKs. Therefore a future direction of this research will entail crystallization of the *S. thermophilum* PP_i-PFK complexed with PP_i analogues and kinetic analyses.

5.3.4 Plant PP_i-PFKs

The literature review addresses the complexity of the allosteric regulation of prokaryotic ATP-PFKs. The only PP_i-PFK to exhibit allosteric regulation is the plant $\alpha_2\beta_2$ PP_i-PFK. The structural resolution of a group II, long clade PP_i-PFK, from *S. thermophilum*, a homologue of the plant β-subunit allows the future modeling of the binding of tripolyphosphate at the PP_i-binding site of *S. thermophilum* PP_i-PFK, which exhibits 95% activity with tripolyphosphate as the phosphoryl donor relative to PP_i (Ronimus *et al.* 1999). The structural resolution of a group II, long clade PP_i-PFK, from *S. thermophilum* allows the structure of a plant PP_i-PFK to be solved by molecular replacement.

5.4 Conclusion

The structure of a group II, long clade PP_i-PFK from the thermophilic bacterium *S. thermophilum* has been solved in two distinct conformations at 2.2 Å and 1.85 Å. The conformation of the insertions, relative to the prokaryotic ATP-PFK of *E. coli* are altered upon the conformational change observed and directly affects the conformation of

conserved motifs at the active site involved in substrate binding and/or catalysis. Therefore the major insertions function to modulate the activity of the PP_i-PFK. This research has demonstrated structural conservation between PFK A ATP- and PP_i-dependent family members. Surprisingly, despite their relatively simple regulation the conformational changes involved in, and required for substrate binding is very complex. Further research is required to resolve the structure complexed with reactants, products, and substrate analogues to fully elucidate the conformational mechanism. This will be further aided by kinetic studies of mutants suggested in this report. The determination of the structures of PP_i-PFKs from the spirochetes *B. burgdorferi* (Moore *et al.* 2002) and *S. thermophilum* has contributed an enormous advancement to this field of study.

References

Aleshin, A. E., Kirby, C., Liu, X., Bourenkov, G. P., Bartunik, H. D., Fromm, H. J., Honzatko, R. B. (2000). Crystal structures of mutant monomeric hexokinase I reveal multiple ADP binding sites and conformational changes relevant to allosteric regulation. ***Journal of Molecular Biology*** **296**, 10001-10015.

Alves, A. M. C. R., Meijer, W. G., Vrijbloed, J. W., and Dijkhuizen, L. (1996). Characterization and phylogeny of the pfp gene of *Amycolaptosis methanolica* encoding PP_i-dependent phosphofructokinase. *Journal of Bacteriology* **178**, 149-155.

Alves, A. M. C. R., Euverink, G. J. W., Bibb, M. J., Dijhuizen, L. (1997). Identification of ATP-dependent phosphofructokinase as a regulatory step in the glycolytic pathway of the actinomycete *Streptomyces coelicor* A3(2). *Applied and Environmental Microbiology* **63**, 956-961.

Alves, A. M. C. R., Euverrink, G. J. W., Santos, H., and Dijkhuizen, L. (2001). Different physiological roles of ATP- and PP_i-dependent phosphofructokinase isozymes in the methylotrophic actinomycete *Amycolaptosis methanolica*. *Journal of Bacteriology* **183**, 7231-7240.

Arvanitidis, A. and Heinisch, J. J. (1994). Studies of the function of yeast phosphofructokinase subunits by in vivo mutagenesis. *Journal of Biological Chemistry* **269**, 8911-8918.

Auzat, I., Le Bras, G., and Garel, J. (1994). The cooperativity and allosteric inhibition of *Escherichia coli* phosphofructokinase depend on the interaction between threonine-125 and ATP. *Proceedings of the National Academy of Sciences USA* **91**, 5242-5246.

Auzat, I., Le Bras, Branny, P., De La Torre, F., Theunissen, B., and Garel, J. (1994). The role of Glu187 in the regulation of phosphofructokinase by phosphoenolpyruvate. *Journal of Molecular Biology* **235**, 68-72.

Auzat, I., Le Bras, G., and Garel, J. (1995). Hypercooperativity induced by interface mutations in the phosphofructokinase from *Escherichia coli*. *Journal of Molecular Biology* **246**, 248-253.

Auzat, I., Le Bras, G., Branny, P., De La Torre, F., Theunissen, B., and Garel, J. (1995). Role of residue 161 in the allosteric transitions of two bacterial phosphofructokinases. *Biochemistry* **34**, 7062-7068.

Barton, G. (1993). ALSCRIPT: a toll to format multiple sequence alignments. *Protein Engineering*. 6, 37-40.

Benach, J., Atrian, S., Gonzalez-Duarte, R., and Ladenstein, R. (1998). The refined crystal structure of *Drosophila lebanonensis* alcohol dehydrogenase at 1.9 Å resolution. *Journal of Molecular Biology* **282**, 383-99

Berger, S. A. and Evans, P. R. (1992). Site-directed mutagenesis identifies catalytic residues in the active site of Escherichia coli phosphofructokinase. *Biochemistry* **31**, 9237-9242.

Berger, S. A. and Evans, P. R. (1990). Active-site mutants altering the cooperativity of *E. coli* phosphofructokinase. *Nature* **343**, 575-576.

Bertagnolli, B. and Cook, P. (1994). Lanthanide pyrophosphates as substrates for the pyrophosphate-dependent phosphofructokinase from *Propionibacterium freudenreichii* and *Phaseolus aureus*: evidence for a second metal ion required for reaction. *Biochemistry* **33**, 1663-1667.

Bork, P., Sander, C., and Valencia, A. (1992). An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proceedings of the National Academy of Sciences USA* **87**, 7290-7294.

Bossemeyer, D., Engh, R. A., Kinzel, V., Ponstingl, H., and Huber, R. (1993). Phosphotransferase and substrate binding mechanism of the cAMP-dependent protein kinase catalytic subunit from porcine heart as deduced from the 2.0A structure of the complex with Mn2+ adenylyl imidodiphosphate and inhibitor peptide. *EMBO Journal* **12**, 849-859.

Bruchhaus, I., Jacobs, T., Denart, M., and Tannich, E. (1996). Pyrophosphate-dependent phosphofructokianse of *Entamoeba histolytica*: molecular cloning, recombinant expression and inhibition by pyrophosphate analogues. *Biochemical Journal* **316**, 57-63.

Brunger, A. (1992). Free R-value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*. 355: 472-475.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R.J., Rice, L. M., Simonson, T., and Warren, G. L. (1998). Crystallography and NMR System: a new software suite for macromolecular structure determination. *Acta Crystallographica*. **D54**, 905-921.

Byrnes, W. M., Zhu, X., Younathan, E. S., and Chang, S. H. (1994). Kinetic characteristics of phosphofructokinase from *Bacillus stearothermophilus*: MgATP nonallosterically inhibits the enzyme. *Biochemistry* **33**, 3424-3431.

Byrnes, W. M., Hu, W., Younathan, E. S., and Chang, S. H. (1995). A chimeric bacterial phosphofructokinase exhibits cooperativity in the absence of heterotropic regulation. *Journal of Biological Chemistry* **270**, 3828-3835.

Cardenas, M. L., Cornish-Bowden, A., Ureta, T. (1998). Evolution and regulatory role of hexokinases. *Biochimica Biophysicas Acta* **1401**, 242-264.

Carlisle, S., Blakeley, S. D., Hemmingsen, S. M., Trevanion, S. J., Hiyoshi, T., Kruger, N. J., and Dennis, D. T. (1990). Pyrophosphate-dependent phosphofructokinase. *Journal of Biological Chemistry* **265**, 18366-18371.

CCP4 (Collaborative Computational Project Number 4) (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallographica.* D50:760-763.

Cheek, S., Zhang, H., and Grishin, N. V. (2002). Sequence and structure classification of kinases. *Journal of Molecular Biology* **320**, 855-881.

Chen, R., Greer, A., and Dean, A. M. (1995). A highly active decarboxylating dehydrogenase with rationally inverted coenzyme specificity. *Proceedings of the National Academy of Sciences USA* **92**, 11666-11670.

Chi, A., and Kemp, R. G. (2000). The primordial high-energy compound: ATP or inorganic pyrophosphate. *Journal of Biological Chemistry* **275**, 35677-35679.

Chi, A. S., Deng, Z., Albach, R. A., and Kemp, R. G. (2001). The two phosphofructokinase gene products of *Entamoeba histolytica*. *Journal of Biological Chemistry* **276**, 19974-19981.

Cho, Y., Matsunaga, T. O., Kenyon, G. L., Bertagnolli, B. L., and Cook, P. F. (1988). Isotope exchange as a probe of the kinetic mechanism of pyrophosphate-dependent phosphofructokinase. *Biochemistry* **27**, 3320-3325.

Cho, Y. and Cook, P. F. (1989). pH dependence of the kinetic parameters for the pyrophosphate-dependent phosphofructokinase reaction supports a proton-shuttle mechanism. *Biochemistry* **28**, 4155-4160.

Claustre, S., Denier, C., Lakhdar-Ghazal, F., Lougare, A., Lopez, C., Chevalier, N., *et al.* (2002). Exploring the active site of *Trypanosoma brucei* phosphofructokinase by inhibition studies: specific irreversible inhibition. *Biochemistry* **41**, 10183-10193.

Cook, A., Lowe, E. D., Chrysina, E. D., Skamnaki, V. T., Oikonomakos, N. G., and Johnson, L. N. (2002). Structural studies on Phospho-CDK2/cyclin A bound to nitrate, a transition state analogue: implications for protein kinase mechanism. *Biochemistry* **41**, 7301-7311.

Deng, Z., Huang, M., Singh, K., Albach, R. A., Latshaw, S. P., Chang, K., and Kemp, R. G. (1998). Cloning and expression of the gene for the active PP_i-dependent phosphofructokinase of *Entamoeba histolytica*. *Biochemical Journal* **329**, 659-664.

Deng, Z., Roberts, D., Wang, X., and Kemp, R. G. (1999). Expression, characterization, and crystallization of the pyrophosphate-dependent phosphofructokinase of *Borrelia burgdorferi*. *Archives of Biochemistry and Biophysics* **371**, 326-331.

Deng, Z., Wang, X., and Kemp, R. G. (2000). Site directed mutagenesis of the fructose-6-phosphate binding site of the pyrophosphate-dependent phosphofructokinase of *Entamoeba histolytica*. *Archives of Biochemistry and Biophysics* **380**, 56-62.

Deville-Bonne, D., Bourgain, F., and Garel, J. (1991). pH dependence of the kinetic properties of allosteric phosphofructokinase from *Escherichia coli*. *Biochemistry* **30**, 5750-5754.

Ding, Y. R., Ronimus, R. S., and Morgan, H. W. (2000). Sequencing, cloning, and high-level expression of the *pfp* gene, encoding a PP_i-dependent phosphofructokinase from the extremely thermophilic eubacterium *Dictyoglomus thermophilum*. *Journal of Bacteriology* **182**, 4661-4666.

Ding, Y. R., Ronimus, R. S., and Morgan, H. W. (2001). *Thermotoga maritima* phosphofructokinases: expression and characterization of two unique enzymes. *Journal of Bacteriology* **183**, 791-794.

Ding, Y. R., Ronimus, R. S., Morgan, H. W. (1999). Purification and properties of the pyrophosphate-dependent phosphofructokinase from *Dictyloglomus thermophilum* Rt46 B.1. *Extremophiles* **3**, 131-137.

Egan, S. E., Fliege, R., Tong, S., Shibata, A., Wolf, R. E. Jr, Conway, T. (1992). Molecular characterization of the Entner-Doudoroff pathway in *E. coli*: sequence analysis and localization of promoters for the *edd-eda* operon. *Journal of Bacteriology* **174**, 4638.

Evans, P. R. and Hudson, P. J, (1979). Structure and control of phosphofructokinase from *Bacillus stearothermophilus*. *Nature* **279**, 500-504.

Fenton, A. W. and Reinhart, G. D. (2002). Isolation of a single activating allosteric interaction in phosphofructokinase from *Escherichia coli*. *Biochemistry* **41**, 13410-13416.

Fothergill-Gilmore, L.A. and Michels, P. (1993). Evolution of glycolysis. *Progress in Biophysics and Molecular Biology* 59: 105-235.

Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., et al. (1995). The minimal gene complement of *Mycoplasma genitalium*. *Science* **270**, 397-403.

Fraser, C. M., Casjens, S., Huang, W. M., Sutton, G. G. Clayton, R., Lathigra, R., et al. (1997). Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*. *Nature* **390**, 580-586.

Fraser, C. M., Norris, S. J., Weinstock, G. M., White, O., Sutton, G. G., Dodson, S., et al. (1998). Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science* **281**, 375-388.

French, S. and Wilson, K. (1978). On the treatment of negative intensity observations. *Acta Crystallographica*. A34: 517-525.

Green, P. C., Latshaw, S. P., Ladror, U. S., and Kemp, R. G. (1992). Identification of critical lysyl residues in the pyrophosphate-dependent phosphofructo-1-kinase of *Propionibacterium freudenreichii*. *Biochemisty* **31**, 4815-4821.

Green, P. C., Tripathi, R. L., and Kemp, R. G. (1993). Identification of active site residues in pyrophosphate-dependent phosphofructo-1-kinase by site-directed mutagenesis. *Journal of Biological Chemistry* **268**, 5085-5088.

Halkides, C. J., Lithcap, E. S., and Frey, P. A. (1991). The substrate reactivity of μ-monothiopyrophosphate with pyrophosphate-dependent phosphofructokinase: evidence for a dissociative transition state in enzymatic phosphoryl group transfer. *Biochemistry* **30**, 10313-10322.

Hampton Research. 5225 Canyon Crest Drive #71-336 Riverside, CA 92507 U.S.A (909) 789-8932

Hansen, T. and Schonheit P. (2001). Sequence, expression, and characterization of the first archaeal ATP-dependent 6-phosphofructokinase, a non-allosteric enzyme related to the phosphofructokinase-B sugar kinase family, from the hyperthermophilic crenarchaeote *Aeropyrum pernix*. *Archives in Microbiology* **177**, 62-69.

Hansen, T., Musfeldt, M., and Schonheit, P. (2002). ATP-dependent 6-phosphofructokinase from the hypethermophilic bacterium Thermotoga maritima: characterization of an extremely thermophilic, allosterically regulated enzyme. *Archives of Microbiology* **177**, 410-409.

Hart, J. C. (1998). An alternative role for the conserved Asp residue in phosphoryl transfer reactions. *Journal of the American Chemical Society* **120**, 13535-13536.

Heinsch, J., Ritzel, R. G., von Borstel, R. C., Aguilera, A., Rodicio, R., and Zimmermann, F. K. (1989). The phosphofructokinase genes of yeast evolved from two gene duplication events. *Gene* **78**, 309-321.

Hellinga, H. W. and Evans, P. R. (1987). Mutations in the active site of *Escherichia coli* phosphofructokinase. *Nature* **327**, 437-439.

Herzberg, O., Chen, C. C., Kapadia, G., McGuire, M., Carroll, L. J., Noh, S. J., Dunaway-Mariano, D. (1996). Swiveling-domain mechanism for enzymatic phosphotransfer between remote reaction sites. *Proceedings of the National Academy of Sciences USA* **93**, 2652-2657.

Hinds, R. M., Xu, J., Walters, E. D., and Kemp, R. G. (1998). The active site of pyrophosphate-dependent phosphofructo-1-kinase based on site-directed mutagenesis and molecular modeling. *Archives of Biochemistry and Biophysics* **349**, 47-52.

Ito, S., Fushinobu, S., Yoshioka, I., Koga, S., Matsuzawa, H., and Wakagi, T. (2001). Structural basis for the ADP-specificity of a novel glucokinase from a hyperthermophilic archaeon. *Structure* **9**, 205-214.

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D.G., Gibson, T.J. (1998). Multiple sequence alignment with Clustal X. *Trends in Biochemical Science***23**, 403-5

Johnson, J. L. and Reinhart, G. D. (1994). Influence of MgADP on phosphofructokinase from *Escherichia coli*. Elucidation of coupling interactions with both substrates. *Biochemistry* **33**, 2635-2643.

Johnson, J. L. and Reinhart, G. D. (1997). Failure of a two-state model to describe the influence of phospho(enol)pyruvate on phosphofructokinase from *Escherichia coli*. *Biochemistry* **36**, 12814-12822.

Kemp, R. and Tripathi, R. (1993). Pyrophosphate-dependent phosphofructo-1-kinase complements fructose 1,6-bisphosphatase but not phosphofructokinase deficiency in *Escheichia coli*. *Journal of Bacteriology* **175**, 5723-5724.

Kemp, R. G. and Gunasekera, D. (2002). Evolution of the allosteric ligand sites of mammalian phosphofructo-1-kinase. *Biochemistry* **41**, 9426-30.

Kimmel, J. L., and Reinhart, G. D. (2001). Isolation of an individual interaction in tetrameric phosphofructokinase from *Bacillus stearothermophilus*. *Biochemistry* **40**, 11623-11629.

Kimmel, J. L., and Reinhart, G. D. (1999). Reevaluation of the accepted allosteric mechanism of phosphofructokinase from *Bacillus stearothermophilus*. *Proceedings of the National Academy of Sciences USA* **97**, 3844-3849.

Kraulis, P. (1991). Molscript: a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography*. 24, 946-950.

Krishna, S. S., Zhou, T., Daugherty, M., Osterman, A., Zhang, H. (2001). Structural basis for the catalysis and substrate specificity of homoserine kinase. *Biochemistry* **40**, 10810-10818.

Kundrot, C. E. and Evans, P. R. (1991). Designing an allosterically locked phosphofructokinase. *Biochemistry* **30**, 1478-1484.

Ladror, U. S., Gollapudi, L., Tripathi, R. L., Latshaw, S. P., and Kemp, R. G. (1991). Cloning, sequencing, and expression of pyrophosphate-dependent phosphofructokinase from *Propionibacterium freudenreichii*. *Journal of Biological Chemistry* **266**, 16550-16555.

Larsen, T. M., Benning, M. M., Rayment, I., Reed, G. H. (1998). Structure of the bis(Mg2+)-ATP-oxalate complex of the rabbit pyruvate kinase at 2.0 Å resolution: ATP binding over a barrel. *Biochemistry* **37**, 6247-6255.

Laskowski, R., MacArthur, M., Moss, D., and Thornton, J. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **26**, 283-291.

Lau, F. T., Fersht, A. R., Hellinga, H. W., Evans P. R. (1987). Site-directed mutagenesis in the effector site of *Escherichia coli* phosphofructokinase. *Biochemistry* **26**, 4143-4148.

Lau, F. T., and Fersht, A. R. (1989). Dissection of the effector-binding site and complementation studies of *Escherichia coli* phosphofructokinase using site-directed mutagenesis. *Biochemistry* **28**, 6841-6847.

Li, C., Kappock, T. J., Stubbe, J., Weaver, T. M., Ealick, S. E. (1999). X-ray crystal structure of aminoimidazole ribonucleotide synthetase (PurM), from the *Escherichia coli* purine biosynthetic pathway at 2.5 Å resolution. *Structure* **7**, 1155-1166.

Li, Y., Rivera, D., Ru, W., Gunasekera, D., and Kemp, R. G. (1999). Identification of allosteric sites in rabbit phosphofructo-1-kinase. *Biochemistry* **38**, 16407-16412.

Lipman, F. (1965). The origin of prebiological systems. (Fox, S.W. ed) 259-310, Academic Press, New York.

Merrit, E. and Bacon, D. (1997). Raster3D: photo realistic molecular graphics. *Methods in Enzymology*. 277, 505-524.

Mertens, E. (1991). Pyrophosphate-dependent phosphofructokinase, an anaerobic glycolytic enzyme? *Federation of European Biochemical Societies* **285**, 1-5.

Mertens, E. (1993). ATP versus pyrophosphate: glycolysis revisited in parasitic protists. *Parasitology Today* **9**, 122-127.

Mertens, E., Ladror, U. S., Lee, J. A., Miretsky, A., Morris, A., Rozario, C., Kemp, R. G., and Muller, M. (1998). The pyrophosphate-dependent phosphofructokinase of the portist, *Trichomonas vaginalis*, and the evolutionary relationships of protist phosphofructokinases. *Journal of Molecular Evolution* **47**, 739-750.

Michels, P. A. M., Chevalier, N., Opperdoes, F. R., Rider, M. H., and Rigden, D. J. (1997). The glycosomal ATP-dependent phosphofructokinase of *Trypanosoma brucei* must have evolved from an ancestral pyrophosphate-dependent enzyme. *European Journal of Biochemistry* **250**, 698-704.

Mildvan, A. S. (1997). Mechanisms of signaling and related enzymes. *Proteins* **29**, 401-416.

Moore, S., Ronimus, R. S., Roberson, R. S., and Morgan, H. W. (2002). The structure of a pyrophosphate-dependent phosphofructokinase from the Lyme disease spirochete *Borrelia burgdorferi. Structure* **10**, 659-671.

Mukund, S., Adams, M. (1995). Glyceraldehyde-3-phosphate ferrodoxin oxidoreductase, a novel tungsten-containing enzyme with a potential glycolytic role in the hyperthermophilic archaeon *Pyrococcus furiosus. Journal of Biological Chemistry.* 270, 8389-8392

Muller, M., Lee, J., Gordon, P., Gaasterland, T., and Sensen, C. W. (2001). Presence of prokaryotic and eukaryotic species in all subgroups pf the PP_i-dependent group II phosphofructokinase protein family. *Journal of Bacteriology* **183**, 6714-6716.

Navaza, J. (1994). AMORE: an automated package for molecular replacement. *Acta Crystallographica.* A50:157-163.

van der Oost, J., Schutt, G., Kengen, S. W. M., Hagen, W. R., Thomm, M, and de Vos, W. M. (1998). The ferrodoxin-dependent conversion of glyceraldehyde-3-phosphate in the hyperthermophilic archaeon *Pyrococcus furiosus* represents a novel site of glycoltyic regulation. *Journal of Biological Chemistry* **273**, 28149-28154.

Otwinowski, Z., Minor W. (1997). Processing X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*. 276, 307-326.

Peekhaus, N. and Conway, T. (1998). What's for Dinner?: Entner-Doudoroff metabolism in *Esherichia coli*. *Journal of Bacteriology* **180**, 3495-3502.

Pham, A. S. and Reinhart, G. D. (2001). MgATP activation by phosphoenolpyruvate of the E187A mutant of *Escherichia coli* phosphofructokinase. *Biochemistry* **40**, 4150-4158.

Pham, A. S., Janiak-Spens, F., and Reinhart, G. D. (2001). Persistent binding of MgADP to the E187A mutant of *Escherichia coli* phosphofructokinase in the absence of allosteric effectors. *Biochemistry* **40**, 4140-4149.

Poorman, R. A., Randolph, A., Kemp, R. G., and Heinrikson, R. L. (1984). Evolution of phosphofructokinase-gene duplication and creation of new effector sites. *Nature* **309**, 467-469.

Reeves, R., South, D. J., Blytt, H. J., and Warren, L. G. (1974). Pyrophosphate: D-fructoose 6-phosphate 1-phosphotransferase. *Journal of Biological Chemistry* **24**, 7737-7741.

Riley-lovingshimer, M. R., Ronning, D. R., Sacchettini, J. C., and Reinhart, G. D. (2002). Reversible ligand-induced dissociation of a tryptophan-shift mutant of phosphofructokinase from *Bacillus stearothermophilus*. *Biochemistry* **41**, 12967-12974.

Romano, A. and Conway, T. (1996). Evolution of carbohydrate metabolic pathways. *14th Forum in Microbiology* 448-455.

Ronimus, R. S and Morgan, H. W. (2002). Distribution and phylogenies of enzymes of the Embden-Meyerhof-Parnas pathway from archaea and hypethermophilic bacteria support a gluconeogenic origin of metabolism. *Archaea*. (article in press, page numbers yet to be assigned).

Ronimus, R. S. and Morgan, H. W. (2001). The biochemical properties and phylogenies of phosphofructokinases from extremophiles. *Extremophiles* **5**, 357-373.

Ronimus, R. S., Morgan, H. W., and Ding, Y. R. (1999). Phosphofructokinase activity within the order Spirochaetales and the characterization of the pyrophosphate-dependent phosphofructokinase from *Spirochaeta thermophilum*. *Archives of Microbiology* **172**, 401-406.

Ronimus, R. S., de Heus, E., Ruckert, A., ands Morgan, H. W. (2001). Sequencing, high-level expression and phylogeny of the pyrophosphate-dependent phosphofructokinase from the thermophilic spirochete *Spirochaeta thermophilum*. *Archives of Microbiology* **175**, 308-312.

Roussel, A. and Cambillau, C. (1991). TURBO-FRODO, Silicon graphics geometry partners directory, pg 77-78. Silicon Graphics, Mountain View, CA.

Rypniewski, W. R. and Evans, P. R. (1989). Crystal structure of unliganded phosphofructokinase from *Escherichia coli*. *Journal of Molecular Biochemistry* **207**, 805-821.

Sakuraba, H., Yoshioka, I., Koga, S., Takahashi, M., Kitahama, K., Satomura, T., *et al.* (2002). ADP-dependent glucokinase/phosphofructokinase, a novel bifunctional enzyme from the hyperthermophilic archaeon Methanococcus *jannaschii*. *Journal of Biological Chemistry* **277**, 12495-12498.

Shirakihara, Y. and Evans, P. R. (1988). Crystal structure of the complex of phosphofructokinase from *Escherichia coli*. *Journal of Molecular Biochemistry* **204**, 973-994.

Schirmer, T. and Evans, P. R. (1990). Structural basis of the allosteric behavior of phosphofructokinase. *Nature* **343**, 140-145.

Schlichting, I. and Reinstein, J. (1997). Structures of active conformations of UMP kinase from *Dictyostelium discoideum* suggest phosphoryl transfer is associative. *Biochemistry* **36**, 9290-9296.

Schlichting, I. and Reinstein, J. (1999). pH influences fluoride coordination number of the AlFx phosphoryl transfer transition state analog. *Nature Structure Biology* **6**, 721-723.

Selig, M., Xavier, K. B., Santos, H., and Schonheit, P. (1997). Comparative analysis of Embden-Meyerhof and Entner-Doudoroff glycolytic pathways in the hyperthermophilic archaea and the bacterium *Thermotoga maritima*. *Archives of Microbiology* **167**, 217-232.

Siebers, B., Klenk, H., and Hensel, R. (1998). PP_i-dependent phosphofructokinase from *Thermoproteus tenax*, an archaeal descendant of an ancient line in phosphofructokinase evolution. *Journal of Bacteriology* **180**, 2137-2143.

Sigrell, J. A., Cameron, A. D., Alwyn Jones, T., and Mowbray, S. L. (1997). Purification, characterization, and crystallization of *Escherichia coli* ribokinase. *Protein Science* **6**, 2474-2476.

Subramanian, G., Koonin, E. V., and Aravind, L. (2000). Comparative genome analysis of the pathogenic spirochetes *Borrelia burgdorferi* and *Treponema pallidum*. *Infection and Immunity* **68**, 1633-1648.

Tuininga, J. E., Verhees, C. H., van der Oost, J., Kengen, S. W. M., Stams, A. J. M., and de Vos W. M. (1999). Molecular and biochemical characterization of the ADP-dependent phosphofructokinase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Journal of Biological Chemistry* **274**, 21023-21028.

Valdez, B. C., French, B. A., Younathan, E. S., and Chang, S. H. (1989). Site-directed mutagenesis in *Bacillus stearothermophilus* fructose-6-phosphate 1-kinase. *Journal of Biological Chemistry* **264**, 131-135.

Verhees, C. H., Tuininga, J. E., Kengen, S. W., Stams, A. J., van der Oost, J., de Vos, W. M. (2001). ADP-dependent phosphofructokinases in mesophilic and thermophilic methanogenic archaea. *Journal of Bacteriology* **183**, 7145-7153.

Verhees, C. H., Koot, D. G. M., Ettema, T. J. G., Dijkema, C., de Vos, W. M., and van der Oost, J. (2002). Biochemical adaptations of two sugar kinases from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Biochemical Journal* **366**, 121-127.

Verhees, C. H., Tuiniga, J. E., Kengen, S. W. M., Stams, A. J. M., van der Oost, J., and de Vos, W. M. (2001). ADP-dependent phosphofructokinases in mesophilic and thermophilic methanogenic archaea. *Journal of Bacteriology* **183**, 7145-7153.

Vlahos, C.J. and Dekker, E.E. (1988). The complete amino acid sequence and identification of the active-site arginine peptide of *E. coli* 2-keto-4-hydroxyglutarate aldolase. *Journal of Biological Chemistry*. 24: 11683-11691.

Voet, D., Voet, J., and Pratt, C. (1999). Fundamentals of biochemistry. John Wiley and Sons, Inc, 382-416.

Wachtershauser, G. (1994). Life in a ligand sphere. *Proceedings of the National Academy of Science*. 91, 516-519.

Wang, X. and Kemp, R. G. (1999). Identification of residues of *Escherichia coli* phosphofructokinase that contribute to nucleotide binding and specificity. *Biochemistry* **38**, 4313-4318.

Wang, X. and Kemp, R. G. (2001). Reaction path of phosphofructo-1-kinase is altered by mutagenesis and alternative substrates. *Biochemistry* **40**, 3938-3942.

Wang, X., Deng, Z., and Kemp, R. G. (1998). An essential methionine residue involved in substrate binding by phosphofructokinases. *Biochemical and Biophysical Research Communications* **250**, 466-468.

Wessberg, K., Sklonick, S., Xu, J., Marciano-Carbal, F., and Kemp, R. G. (1995). Cloning, sequencing and expression of the pyrophosphate-dependent phosphofructo-1-kinase from *Naegleria fowleri. Biochemical Journal* **307**, 143-149.

Wistow, G. and Piatigorsky, J. (1987). Recruitment of enzymes as lens structural proteins. *Science*. 236, 1554-1556.

Woese, C. R., Kandler, O., and Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and eukarya. *Proceedings of the National Academy of Sciences USA* **87**, 4576-4579.

Wu, L., Reizer, A., Reizer, J., Cai, B., Tomich, J. M., and Saier (Jr), M. H. (1991). Nucleotide sequence of the *Rhodobacter capsulatus fuK* gene, which encodes fructose-1-phosphate kinase: evidence for a kinase superfamily including both phosphofructokinases of *Escherichia coli. Journal of Bacteriology* **173**, 3117

Xu, J., Green, P. C., and Kemp, R. G. (1994). Identification of basic residues involved in substrate binding and catalysis by pyrophosphate-dependent phosphofructokinase from *Propionibacterium freudenreichii. Journal of Biological Chemistry* **269**, 15553-15557.

Xu, Y., Morera, S., Janin, J., and Cherfils, J. (1997). AlF3 mimics the transition state of protein phosphorylation in the crystal structure of nucleoside diphosphate kinase and ADP. *Proceedings of the National Academy of Sciences* **94**, 3579-83.

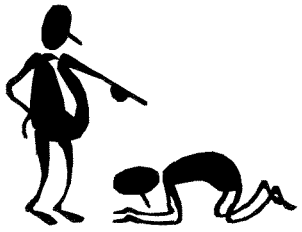
Zhao, G., Pease, A. J., Bharani, N., Winkler, M. E. (1995). Biochemical characterisation of *gapB*-encoded erythrose 4-phosphate dehydrogenase of E. coli K-12 and its possible role in pyridoxal 5'-phosphate biosynthesis. *Journal of Bacteriology* **177**, 2804-2812.

Zheng, R. and Kemp, R. G. (1992). The mechanism of ATP inhibition of wild type and mutant phosphofructo-1-kinase from *Escherichia coli*. *Journal of Biological Chemistry* **267**, 23640-23645.

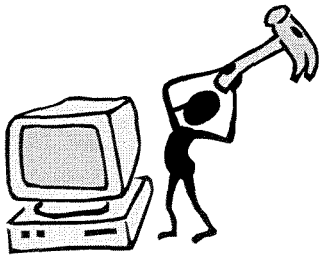
Zheng, R. and Kemp, R. G. (1994). Identification of interactions that stabilize the transition state in *Escherichia coli* phosphofructo-1-kinase. *Journal of Biological Chemistry* **269**, 18475-18479.

Zhu, X., Byrnes, M., Nelson, J. W., and Chang, S. H. (1995). Role of glycine 212 in the allosteric behavior of phosphofructokinase from *Bacillus stearothermophilus*. *Biochemistry* **34**, 2560-2565.

Appendix 1: An insight into my Masters:



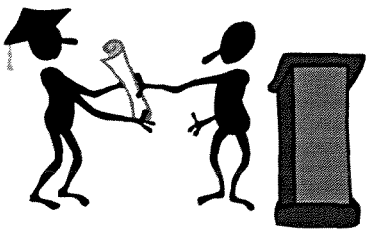
Micro-management by my supervisor.



I love my computer!

Nope, it won't work!

There is only one way to fix it!



I hope