The conformational landscape of β-phosphoglucomutase and its role in enzyme activity

Francisco Aaron Cruz Navarrete

Supervisor:

Professor Jonathan P. Waltho

Dept. of Molecular Biology and Biotechnology

Thesis submitted to the University of Sheffield

for the degree of Doctor of Philosophy

August 2021

Science endeavour is not about getting things right the first time, but about progressively getting things less wrong.

Table of Contents

List of figures

List of tables

Abbreviations

ΔG_{eq}	Equilibrium free energy
$\Delta G^{\ddagger}$	Activation free energy
k	First order rate constant
h	Plank's constant (6.63×10^{-34} J·s)
k_B	Boltzmann's constant (1.38×10^{-23} J K^{-1})
T	Temperature in Kelvin (298 K)
R	Gas constant (1.99×10^{-3} kcal K^{-1} mol^{-1})
βPGM	β-Phosphoglucomutase
βPGM^P	Phosphorylated β-phosphoglucomutase
A	Substrate-free β-phosphoglucomutase conformer A
B	Substrate-free β-phosphoglucomutase conformer B
A^P	Phosphorylated substrate-free β-phosphoglucomutase conformer A
B^P	Phosphorylated substrate-free β-phosphoglucomutase conformer B
βG16BP	β-Glucose 1,6-bisphosphate
αG16BP	α-Glucose 1,6-bisphosphate
βG1P	β-Glucose 1-phosphate
G6P	Glucose 6-phosphate
F16BP	Fructose 1,6-bisphosphate
AcP	Acetyl-phosphate
G6PDH	Glucose 6-phosphate dehydrogenase
NADP/H	Nicotinamide adenine dinucleotide phosphate
NMR	Nuclear magnetic resonance
HAD	Haloacid dehalogenase
S	Substrate
P	Product
E	Enzyme
ES	Enzyme complex
ETS	Enzyme-transition state complex
EP	Enzyme-product complex
E$_C$:S	Closed active Michaelis complex
E$_O$:S	Open inactive Michaelis complex
TS	Transition state
TSA	Transition state analogue
GSA	Ground state analogue
NAC	Near-attack conformer
ENAC	Enzyme-NAC complex
k_{cat}	Catalytic rate
K_m	Michaelis-Menten constant
K_i	Inhibition constant
EDTA	Ethylenediaminetetraacetic acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
TSP	Trimethylsilylpropanoic acid
HSQC	Heteronuclear single quantum coherence
TROSY	Transverse relaxation-optimised spectroscopy
Mg$_{cat}^{2+}$	Catalytic magnesium ion

Acknowledgements

accomplished in this thesis is enough to make you forget about fibronectin. Thanks also to Mike and Rosie for their advice and for asking interesting questions during my lab talks.

Thanks to Henry, Wai Ching and Gabe for reminding me that lab work is not the only thing that exists during a PhD. I hope our friendship lasts beyond this part of our lives.

Finally, thanks to Falafel King. Your wraps kept me well fed, happy and helped me power through the busy days at the lab.

Author contributions

Cruz-Navarrete, F.A., Baxter, N.J., Wood, H.P., Hounslow, A.M. and Waltho, J.P. (2019). H-1, N-15 and C-13 backbone resonance assignments of the P146A variant of beta-phosphoglucomutase from Lactococcus lactis in its substrate-free form. *Biomolecular Nmr Assignments.* **13(2), pp.349-356.**

F.A.C.N., N.J.B., H.P.W., J.P.W. designed research; F.A.C.N. produced isotopically enriched protein; F.A.C.N., A.M.H. acquired NMR experiments; F.A.C.N., N.J.B. performed backbone resonance assignments; F.A.C.N., N.J.B., J.P.W. wrote the paper with help from all authors.

Wood, H.P., Cruz-Navarrete, F.A., Baxter, N.J., Trevitt, C.R., Robertson, A.J., Dix, S.R., Hounslow, A.M., Cliff, M.J. and Waltho, J.P. (2020). Allomorphy as a mechanism of post-translational control of enzyme activity. *Nature communications.* **11(1), pp.5538-5538.**

H.P.W., F.A.C.N., N.J.B., C.R.T., A.J.R., J.P.W. designed research; H.P.W., F.A.C.N., C.R.T., A.J.R. produced isotopically enriched protein; H.P.W. performed coupled assay kinetic experiments; H.P.W., C.R.T. produced and purified βG1P and βG16BP; H.P.W., F.A.C.N., C.R.T., A.M.H., M.J.C. acquired NMR experiments; H.P.W., F.A.C.N., N.J.B., C.R.T. analysed NMR data; F.A.C.N., N.J.B., C.R.T. performed backbone resonance assignments; H.P.W., A.J.R., S.R.D. performed and analysed X-ray crystallography experiments; F.A.C.N. conceived and developed the allomorphy mechanism; H.P.W., F.A.C.N., N.J.B., C.R.T., J.P.W. wrote the paper with help from all authors.

Robertson, A.J., Cruz-Navarrete, F.A., Wood, H.P., Vekaria, N., Hounslow, A.M., Bisson, C., Cliff, M.J., Baxter, N.J. and Waltho, J.P. (2021). Enzymes that catalyse intrinsically difficult reactions harvest substrate binding energy for stabilisation of the closed state but at the risk of introducing substrate inhibition. Submitted.

A.J.R., F.A.C.N., H.P.W., N.J.B., J.P.W. designed research; A.J.R., H.P.W., N.V. produced isotopically enriched protein; H.P.W. performed coupled assay kinetic experiments; H.P.W. produced and purified βG1P; A.J.R., F.A.C.N., H.P.W., N.V., A.M.H., M.J.C. acquired NMR experiments; A.J.R., F.A.C.N., N.J.B. analysed NMR data; A.J.R., H.P.W., C.B. performed and analysed X-ray crystallography experiments; F.A.C.N. interpreted the data; F.A.C.N., H.P.W., N.J.B., J.P.W. wrote the paper with help from all authors.

Abstract

The research delineated in this thesis concerns the conformational landscape of the phosphoryl transfer enzyme β-phosphoglucomutase (βPGM) from *Lactococcus lactis* and its impact on its activity. βPGM catalyses the isomerisation of β-glucose 1-phosphate to glucose 6-phosphate via a β-glucose 1,6-bisphosphate intermediate. βPGM is the gatekeeper between the metabolism of maltose and trehalose and glycolysis. Here we show that βPGM exchanges between two conformations with different activities on a multi-second time scale through *cis-trans* proline isomerisation. This exchange process is dependent on the phosphorylating agent present and only β-glucose 1,6-bisphosphate couples the phosphorylation step with the conformational switch, allowing βPGM to reach maximal activity. We termed this control mechanism allomorphy, to emphasise its relationship to and distinction from both allostery and allokairy mechanisms. Allomorphy allows βPGM to differentially modulate its activity depending on the carbohydrate concentrations inside the cell. Additionally, βPGM can achieve high catalytic proficiency by efficiently utilising the substrate binding energy to facilitate the adoption of a closed active conformation instead of using that energy to stabilise the transition state of the chemical step. Furthermore, our results uncover a trend whereby enzymes that catalyse intrinsically difficult chemistry exhibit a greater requirement to stabilise the closed active form. However, this catalytic proficiency mechanism comes at the expense of introducing substrate inhibition to catalysis. This thesis highlights the importance of understanding of the conformational landscape of enzymes and offers novel tools in the design of synthetic enzymes that catalyse difficult reactions not found in Nature.

1 Introduction

The central dogma of molecular biology explains the flow of information in living beings (Crick, 1970). An analogous hypothesis can be constructed for proteins. First, the 1D amino acid sequence defines the 3D protein structure. This process is known as protein folding, and it has been an area of active research since around 1960 (Dill et al., 2008). Second, the protein structure defines protein function. This concept is known as the protein structure-function paradigm. Since the advent of the structural biology field, more emphasis has been directed toward understanding the structure-function relationships in proteins (Thornton et al., 2000; Moult and Melamud, 2000; Redfern, Dessailly and Orengo, 2008). Moreover, given the constant increase in protein structures solved and deposited in the Protein Data Bank (Berman et al., 2000) since its creation, an effort has been made in generating bioinformatic tools to predict the biochemical and biological properties of proteins from their structures (Orengo, Todd and Thornton, 1999; Yang et al., 2014).

However, experimentally determined protein structures (mainly by X-ray crystallography or Cryo-electron microscopy techniques) are static pictures. Since 1975, proteins have been described as "screaming and kicking" in solution (Weber, 1975). Proteins sample a large ensemble of conformations around an energy minimum due to thermal fluctuations and changes in the chemical environment (Frauenfelder, Sligar and Wolynes, 1991). In fact, instead of referring to the lowest-energy state of a protein in solution as the native state, suggesting a single dominant conformation, it could be referred to as the native ensemble, which highlights the existence of a collection of low-energy conformations in exchange (Vila, 2021). A complete understanding of protein structure-function relationships requires the description of the protein conformational landscape (Henzler-Wildman and Kern, 2007). The conformational landscape contains information about the relative populations of conformations (thermodynamics) and the energy barriers between them (kinetics). These concepts are particularly relevant in the study of enzymes. Enzymes must balance a stable structure with flexibility to facilitate substrate binding, catalysis and product release (Wolf-Watz et al., 2004; Malabanan, Amyes and Richard, 2010). The knowledge of the conformational landscape of enzymes will accelerate the design of synthetic enzymes with novel functions or broaden the scope of targets suitable for drug design.

This thesis investigates the conformational landscape of a phosphoryl transfer enzyme and the impact on its activity. β-Phosphoglucomutase (βPGM) from *Lactococcus lactis* (*L. lactis*) exchanges between two conformations with different activities on a multi-second time scale. This exchange process allows βPGM to differentially modulate its activity depending on the carbohydrate concentrations inside the cell. Additionally, βPGM can achieve high catalytic proficiency by efficiently utilising the substrate binding energy to facilitate the adoption of a closed active conformation. However, this catalytic proficiency mechanism risks introducing substrate inhibition to catalysis.

This thesis is structured as an alternative thesis format. Chapter 3 and chapter 4 are manuscripts that have been published in peer-reviewed journals (Cruz-Navarrete et al., 2019; Wood et al., 2020) and chapter 6 is a manuscript currently under preparation.

1.1 Chemical equilibrium

For a chemical reaction, chemical equilibrium is the state where the relative concentrations of reactants and products do not change over time. From a kinetic perspective, equilibrium is reached when the forward and backward rates are equal. A chemical reaction is said to be spontaneous when the equilibrium concentrations of the products are larger than the equilibrium concentrations of the reactants. Thermodynamically, this means that the free energy of the products is lower than the free energy of the reactants, and the equilibrium free energy (ΔG_{eq}) is therefore negative (Figure 1.1a). For a unimolecular reaction A ↔ B, the product and reactant concentrations and the ΔG_{eq} are related by the Gibbs–Helmholtz equation:

$$\Delta G_{eq} = -RT \ln K_{eq} \qquad\qquad 1.1$$

$$K_{eq} = \frac{[B]}{[A]} \qquad\qquad 1.2$$

where K_{eq} is the equilibrium constant, R is the gas constant and T is the temperature in Kelvin.

1.2 Transition state theory

The transition state theory relates rates of chemical reactions to the thermodynamic properties of a high-energy state known as the transition state (TS) (Eyring, 1935). It conceives a chemical reaction as a progression through a pseudo dimensional

reaction coordinate, where a reactant molecule must pass through the TS to become a product molecule (Figure 1.1a). The TS is the maximum energy state of the reaction, i.e., is at the top of a high-energy barrier. The height of the barrier (also known as the activation free energy ($\Delta G^{\ddagger}$)) is related to rate of the reaction by the Eyring–Polanyi equation:

$$\Delta G^{\ddagger} = -RT \ln \frac{hk}{\kappa k_B T} \qquad 1.3$$

where k is the first order rate constant, h is the Plank's constant, k_B is the Boltzmann's constant and κ is the transmission coefficient (assumed to be 1 in this thesis). The transition state theory offers a theoretical framework to rationalise catalysis in terms of thermodynamic properties (Truhlar, Garrett and Klippenstein, 1996). The relationship between the equilibrium and activation free energies is given by:

$$\Delta G_{eq} = \Delta G^{\ddagger}_{forward} - \Delta G^{\ddagger}_{backward} \qquad 1.4$$
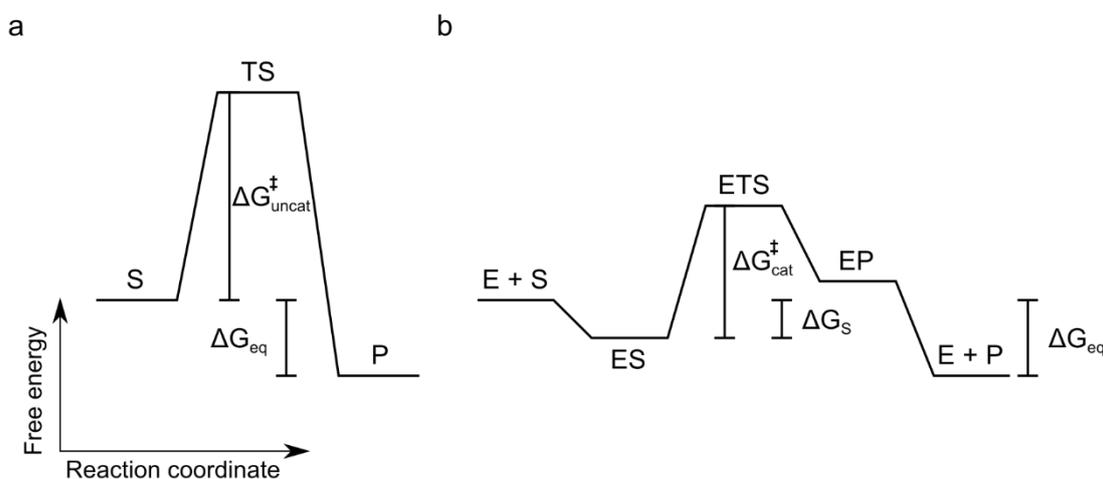
Figure 1.1 Energy level diagrams for a unimolecular reaction. (a) For a chemical reaction, the transition state theory establishes that a substrate (S) must pass through a maximum energy state, named the transition state (TS), to become a product (P). The rate of the reaction depends on the height of the energy barrier ($\Delta G^{\ddagger}_{uncat}$). The relative equilibrium concentrations of S and P are determined by the equilibrium free energy (ΔG_{eq}). (b) In the presence of an enzyme (E), the height of the energy barrier decreases ($\Delta G^{\ddagger}_{cat}$), that is, the rate of the reaction becomes faster. This is achieved by the formation of an enzyme-substrate complex (ES), which is more stable than its individual components by the ground state free energy (ΔG_S). After the reaction through the enzyme-transition state complex (ETC) occurs, the relatively high-energy enzyme-product complex (EP) rapidly dissociates into E and P. However, the ΔG_{eq} of the reaction remains unchanged.

1.3 Enzyme catalysis

Enzymes are biological catalysts, that is, biomolecules that increases the rate of a reaction without being consumed at the end of it. Enzymes, and catalysts in general, do not shift the equilibrium of a reaction towards the products (ΔG_{eq} is not affected). Instead, enzymes lower the free energy barriers between reactants and products ($\Delta G^{\ddagger}$ become less positive) (Cornish-Bowden, 1995). Enzymes achieve this acceleration by forming enzyme-substrate complexes (Figure 1.1b). To determine the activity of an enzyme, initial reaction rates (v) at different substrate concentrations are measured and fitted to the Michaelis-Menten equation:

$$\frac{v}{[E]} = \frac{k_{cat}[S]}{K_m + [S]}$$

1.5

where k_{cat} is the catalytic rate constant, [E] is the enzyme concentration, [S] is the substrate concentration, K_m is the Michaelis-Menten constant (the substrate concentration where $v = k_{cat}/2$). The catalytic prowess of an enzyme can be quantified by calculating its catalytic proficiency (Miller and Wolfenden, 2002):

$$\text{Catalytic proficiency} = \frac{k_{cat}/K_m}{k_{non}}$$

1.6

where k_{non} is the first-order rate constant of the uncatalysed reaction.

Unlike non-natural catalysts, enzymes can accelerate intrinsically difficult chemistry to biologically relevant timescales at physiological temperatures. Hence, enzymes represent an emerging tool with great potential to aid in sustainable development initiatives (Sheldon and Brady, 2021). For example, enzymes already have a significant impact in biofuel production, plastic degradation and "green chemistry" (Wen, Nair and Zhao, 2009; Yoshida et al., 2016; Scherer et al., 2021). Therefore, studies that aim to further the understanding of the mechanism of enzyme catalysis is of great interest. Below is an overview of some of the mechanisms that have been proposed to explain the high catalytic proficiency of enzymes.

1.3.1 Transition state stabilisation by electrostatics

A proposed mechanism of enzyme catalysis is focused on favourable electrostatic interactions between the enzyme active site residues and the TS (Figure 1.2a)

(Warshel, 1998; Garcia-Viloca et al., 2004; Warshel, A. et al., 2006). In solution, the charges that accumulate upon formation of the TS need to be stabilised by the surrounding solvent molecules. However, this organisation comes with a high entropic cost and therefore the net stabilisation is negligible. In contrast, enzymes provide a preorganised polar environment that stabilises the TS architecture more than water. The entropic cost of a preorganised active site comes with a reduction in structural stability. Moreover, further stabilisation is delivered by lowering the dielectric constant upon desolvation and enzyme closure (Kukic et al., 2013), which increases the strength of the enzyme-TS electrostatic interactions. Enhanced catalysis by TS stabilisation has been suggested to explain chorismite mutase catalysis (Burschowsky et al., 2014). Using a series of high-resolution X-ray crystallography structures of a mutant enzyme that could effectively bind the substrate, it was observed that electrostatic complementarity between the active site and the TS is the most important contributor to catalysis.

1.3.2 Ground state destabilisation

An alternative hypothesis to explain the high catalytic proficiency of enzymes is that substrate binding energy is used directly to destabilise the bound substrate relative to the transition state (Figure 1.2b). In this mechanism, destabilisation can be achieved by desolvation, hydrogen-bond breaking, electrostatic repulsion, entropic loss and geometric destabilisation (Jencks, 1975; Carey, 2006). The enzyme active site provides an optimal fit for the TS, but it coordinates the substrate less effectively. Thus, the bound substrate is in a high-energy conformation, which forces its transformation towards the TS. Thermodynamically, this means that the reduction of the activation free energy comes at the expense of an suboptimal ground state free energy (Figure 1.1b, Figure 1.2b). This mechanism has been used to explain the high catalytic proficiency of alkaline phosphatase (AP), which has wide substrate specificity (Andrews, Fenn and Herschlag, 2013). AP leverages electrostatic repulsion from anionic nucleophiles to destabilise the ground state phosphodianion group of the substrate. However, this mechanism has been questioned on the basis that the mutation of active site residues in several proteins to hinder their catalytic proficiencies does not result in ground state stabilisation but in TS destabilisation (Warshel, 1998).

1.3.3 Near-attack conformers

Enzyme catalytic proficiency can also be explained as an entropic effect originating from the stabilisation of a configuration that pushes the reactants to a close reaction distance (Figure 1.2c) (Bruice, 1976). In the reaction coordinate, this configuration occurs prior to TS formation, and therefore is termed a near-attack conformer (NAC) (Hur and Bruice, 2002). The NAC model partitions the activation free energy into chemical and nonchemical transformation steps; the latter corresponding to NAC formation. In water, NAC formation is unfavourable due to electrostatic and geometric constrains. Enzymes resolve this issue by establishing electrostatic interactions with charged substrate groups and by sequestering the substrate molecule in a confined space (Hur and Bruice, 2003). That is, enzymes lower the activation free energy of the nonchemical step. It has been suggested that the activation free energy of the chemical step remains constant in reactions with similar mechanisms (Lightstone and Bruice, 1996). Evidence of NAC formation in enzymes has been reported in chorismate mutase (Hur and Bruice, 2003) and β-phosphoglucomutase (Griffin et al., 2012; Johnson et al., 2018). However, it has been argued that NAC stabilisation is just a consequence of TS stabilisation by electrostatic effects (Strajbl et al., 2003; Warshel, A. et al., 2006).

1.3.4 Conformational dynamics

The conformational dynamics present in the folded enzyme could also have a role in lowering the reaction activation free energy. Enzymes, and proteins in general, exhibit motions operating on different timescales (Figure 1.3). In particular, conformational dynamics in the µs–ms timescale have been suggested to be directly coupled to the chemical step (Benkovic and Hammes-Schiffer, 2003). For example, molecular simulations and theoretical studies of the human cyclophilin A (CypA), a peptidyl-prolyl *cis-trans* isomerase, indicate that conserved residues in the active and distal sites show correlated motions on the same timescale of the reaction (Agarwal, Geist and Gorin, 2004; Agarwal, 2005; Doshi et al., 2012). These conformational fluctuations, together with energy transfer from the first hydration shell, alter the activation free energy. Coupling between dynamics and catalysis has been proposed to be present in the thermophilic alcohol dehydrogenase to explain temperature-dependent hydrogen-deuterium exchange experiments (Liang et al., 2004). However, a direct link

between conformational dynamics and catalysis is still under debate (Warshel, A. et al., 2006; Pisliakov et al., 2009; Kamerlin and Warshel, 2010).
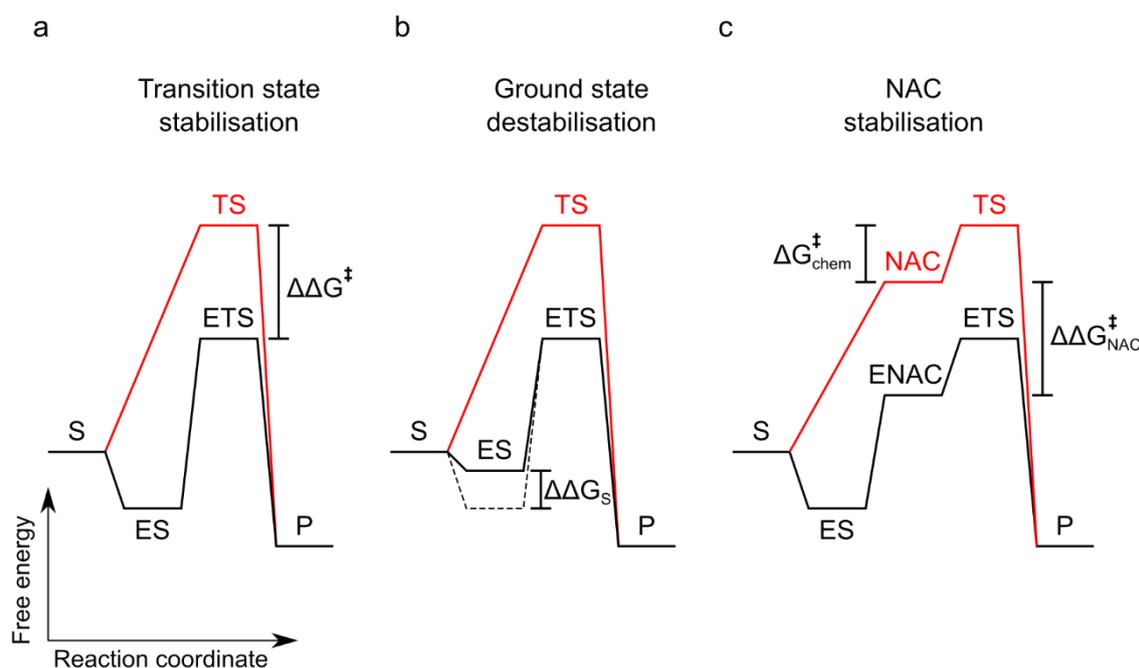
Figure 1.1   Proposed mechanisms of enzyme catalysis. Energy level profiles representing three mechanisms of enzyme catalysis (a) Transition state stabilisation: the enzyme establishes favourable electrostatic interactions between the TS and active site residues, reducing the activation free energy ($\Delta\Delta G^{\ddagger}$) of the reaction. (b) Ground state destabilisation: the enzyme invests the substrate binding energy ($\Delta\Delta G_S$) to impose geometric and electrostatic constrains on to the structure of the substrate, forcing it to adopt a TS-like conformation. A feature of this mechanism is the observation of a marginally stable ES complex. (c) NAC stabilisation: the enzyme lowers the energy of near-attack conformers ($\Delta\Delta G^{\ddagger}_{NAC}$) prior to the chemical step. However, the activation free energy of the chemical step ($\Delta G^{\ddagger}_{chem}$) remains constant.

Alternatively, enzyme conformational dynamics could have a role in the events preceding and following the chemical step (Nashine, Hammes-Schiffer and Benkovic, 2010). Due to the chemical similarity between substrates, products and the TS, enzymes must have mechanisms to avoid getting trapped in non-productive complexes. In the HIV-1 protease, the preorganisation of the active site is suggested to be the rate-limiting step of its catalytic cycle (Torbeev et al., 2011). In contrast, conformational changes in glutathione transferase allow for efficient product release (Codreanu et al., 2002). Finally, a comprehensive study of the catalytic cycle of dihydrofolate reductase using ground state analogues demonstrated that its conformational landscape efficiently biases its conformations through a catalytically

competent kinetic pathway (Boehr et al., 2006). Therefore, conformational dynamics in enzymes have an important role in catalysis, albeit not directly related to the chemical step.

The importance of conformational dynamics has been acknowledged in enzyme evolution and enzyme design. Enzymes with a rugged conformational landscape sometimes display functional promiscuity and moonlighting, which could be a route for the generation of new functionalities by gene duplication and natural selection (James and Tawfik, 2003). Additionally, synthetic enzymes display increased active site flexibility that hampers their catalytic potential (Bhabha, Biel and Fraser, 2015). Rigidifying these active sites without compromising substrate binding and product realise could aid in the design of enzymes that catalyse intrinsic difficult chemistry of biotechnological importance (Ruscio et al., 2009; Preiswerk et al., 2014; Crean, Gardner and Kamerlin, 2020). Recently, the importance of approaching conformational dynamics under a protein evolution framework has been emphasised in the context of enzyme design (Trudeau and Tawfik, 2019; Gardner et al., 2020).
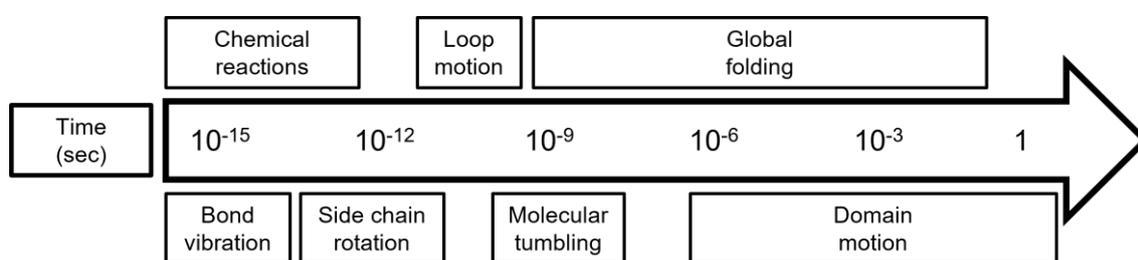
Figure 1.2 Protein dynamics at different timescales. Different structural conformations are hierarchically sampled within characteristic timescales. The study of conformational dynamics is an all-inclusive scientific framework that brings together structural biology, thermodynamics and kinetics.

1.3.5 The role of regions distant from the active site in catalysis

The effect of mutations in sites distal to the active site has highlighted the importance of conformational dynamics in enzyme catalytic proficiency (Lee and Goodey, 2011). A distal site residue is characterised by not being within van der Waals distance to portions of the substrate. Additionally, residues that make contacts with the substrate but are far away from the active site are also termed as distal. This thesis does not make a distinction between these two kinds of distal site residues. Distal site perturbations can affect several aspects of enzyme catalysis: substrate binding and

product release (Miller, Wahnon and Benkovic, 2001; Chalissery et al., 2007), enzyme catalytic activity (Huang and Breitwieser, 2007; Klyuyeva, Tuganova and Popov, 2008) and enzyme structure, conformational dynamics and quaternary assembly (Chaptal et al., 2007; Hong et al., 2007).

In particular, the contribution of distal site interactions to the catalytic proficiency of enzymes that bind substrates via their non-reacting phosphodianion groups has been studied comprehensively (Amyes, O'Donoghue and Richard, 2001; Amyes, Richard and Tait, 2005; Tsang, Amyes and Richard, 2008; Amyes and Richard, 2013). Distal site phosphodianion group binding contributes a consistent 11–13 kcal·mol^{-1} to the reduction of the activation energy barrier for the reactions of glycerol 3-phosphate dehydrogenase (GPDH), orotidine 5'-monophosphate decarboxylase (OMPDC) and triose phosphate isomerase (TIM) (Amyes, O'Donoghue and Richard, 2001; Amyes, Richard and Tait, 2005; Tsang, Amyes and Richard, 2008; Morrow, Amyes and Richard, 2008). It has been proposed that the energy derived from the phosphodianion group binding is used to stabilise a closed active form rather than to induce TS stabilisation. Structural modifications of sidechains within the distal site or absence of the substrate phosphodianion group both result in the impairment of enzyme closure. Hence, the adoption of a catalytically active state becomes rate-limiting (Reyes et al., 2015a).

1.4 Enzyme regulation

Living beings can perform their metabolic functions due to the action of enzymes. Due to their importance, enzyme activity is tightly regulated so as to avoid catastrophic waste of energy and metabolites (Cornish-Bowden, 1995). Physiologically, failure of regulation often results in metabolic disorders and disease (Pardee, 2006; Iommarini et al., 2017). Moreover, enzyme regulation allows an organism to effectively respond to environmental stimuli or changes in food source. Enzymes involved in regulatory mechanisms operate at different levels and at distinct timescales. These mechanisms can be broadly categorised into two groups: coarse control and fine control mechanisms. These groups will be discussed briefly below.

1.4.1 Coarse control mechanisms

Coarse control mechanisms act upon the concentration of enzymes and operate on a timescale of hours to days. The regulatory mechanisms of gene transcription and

protein translation are part of this category. An example of gene regulation is the *lac* operon in *Escherichia coli* (*E. coli*) (Jacob and Monod, 1961; Lewis, 2005). The *lac* operon encodes the enzymes that are responsible for lactose metabolism. Transcription of the *lac* operon is stimulated by the presence of allolactose. In the absence of lactose, the lacI repressor prevents the transcription of the *lac* operon. Hence, *E. coli* only synthesises these enzymes in response to a change in the environment. Enzyme compartmentalisation and co-clustering are also coarse control mechanisms. Enzyme activity can be localised by sequestering the enzyme in a particular compartment (de Graffenried and Bertozzi, 2004; Boonstra and Verkleij, 2004). Enzyme co-clustering increases the effective concentration of the enzyme and avoids undesirable diversion of metabolites by substrate channelling (Cornish-Bowden, 1995; Huang, Holden and Raushel, 2001; Castellana et al., 2014). Another mechanism of coarse control is protein degradation (Schimke and Doyle, 1970). The constant targeting of enzymes for degradation ensures that their activity does not continue for longer than the cell requires it (Ciechanover, 2005). It also allows for the repurposing of amino acids.

1.4.2 Fine control mechanisms

Fine control mechanisms modulate enzyme activity by directly acting upon the enzyme structure and operate on a timescale of sub-second to minutes. This category includes post-translational covalent modifications and binding of regulatory (effector) molecules (Krebs and Beavo, 1979; Cornish-Bowden, 1995; Zhao et al., 2010). Finer control mechanisms often work through allostery, whereby an effector molecule binds in a site separate from the active site (Monod, Changeux and Jacob, 1963). Allostery requires the existence of at least two different enzyme conformations with different activities that can exchange with each other. Binding of the effector molecule promotes a conformational change that stabilises a form of the enzyme with reduced or enhanced activity. In monomeric enzymes, allosteric regulation can be observed as a switch between an inactive and an active conformation (Figure 1.4). The structural mechanism of allostery is the subject of substantial interest (Changeux, 2013). A proposed mechanism is that allostery acts through the high structural heterogeneity of the conformational landscape of enzymes (Motlagh et al., 2014). Allosteric sites have been successfully exploited as drug targets, and a large number of allosteric drugs

acting as inhibitors, regulators and activators have been commercialised (Liu et al., 2020).

Another (but less well-recognised) fine control mechanism is allokairy (Hilser, Anderson and Motlagh, 2015). Allokairy, also known as kinetic cooperativity, is a kinetic control mechanism where the enzyme activity is modulated by the near-equivalence of the catalytic rate and the exchange rate of a conformational switch between an active conformation and an inactive conformation. Enzyme activity under allokairy regulation has a sigmoidal response to substrate concentration (Figure 1.4). Allokairy has been thoroughly studied in human glucokinase, which catalyses the ATP-dependent phosphorylation of glucose (Kamata et al., 2004; Larion and Miller, 2012; Larion et al., 2012; Larion et al., 2015; Whittington et al., 2015). Glucokinase spontaneously exchanges between two conformations with different glucose affinities. At low glucose concentrations, the enzyme has time to relax towards the low affinity state between glucose binding events in the active site. In contrast, at high glucose concentrations, the rate of glucose binding is faster than the exchange rate of the conformational switch and glucokinase activity is maintained at the catalytic rate.

1.5 β-Phosphoglucomutase

β-Phosphoglucomutase (βPGM, EC 5.4.2.6, 24 kDa) from *L. lactis* is a monomeric enzyme of the haloacid dehalogenase superfamily (HAD) (Lahiri et al., 2002a; Zhang et al., 2005; Baxter et al., 2006; Baxter et al., 2010; Johnson et al., 2018). βPGM catalyses the magnesium-dependent isomerisation between β-glucose 1-phosphate (βG1P) and glucose 6-phosphate (G6P) via phosphoryl transfer steps (Figure 1.5). During its catalytic cycle, phosphorylated βPGM (βPGM^P, phosphorylated on D8) binds βG1P and generates β-glucose 1,6-bisphosphate (βG16BP). Next, βG16BP is released to solution and rebinds in the alternative orientation (Dai et al., 2006). Finally, βPGM catalyses phosphoryl transfer from βG16BP, consequently forming G6P and regenerating βPGM^P. βPGM has been studied extensively using enzyme activity experiments (Zhang et al., 2005; Dai et al., 2006; Dai et al., 2009; Golicnik et al., 2009; Wood et al., 2020), Nuclear magnetic resonance (NMR) spectroscopy (Baxter et al., 2006; Baxter et al., 2008; Baxter et al., 2009; Baxter et al., 2010; Griffin et al., 2012; Jin et al., 2014; Johnson et al., 2018), X-ray crystallography (Lahiri et al., 2002a; Lahiri et al., 2002b; Lahiri et al., 2003; Tremblay et al., 2005; Baxter et al., 2010; Griffin et

al., 2012; Jin et al., 2014; Johnson et al., 2018), and density functional theory approaches (Webster, 2004; Marcos, Field and Crehuet, 2010; Elsasser, Dohmeier-Fischer and Fels, 2012; Barrozo et al., 2018).
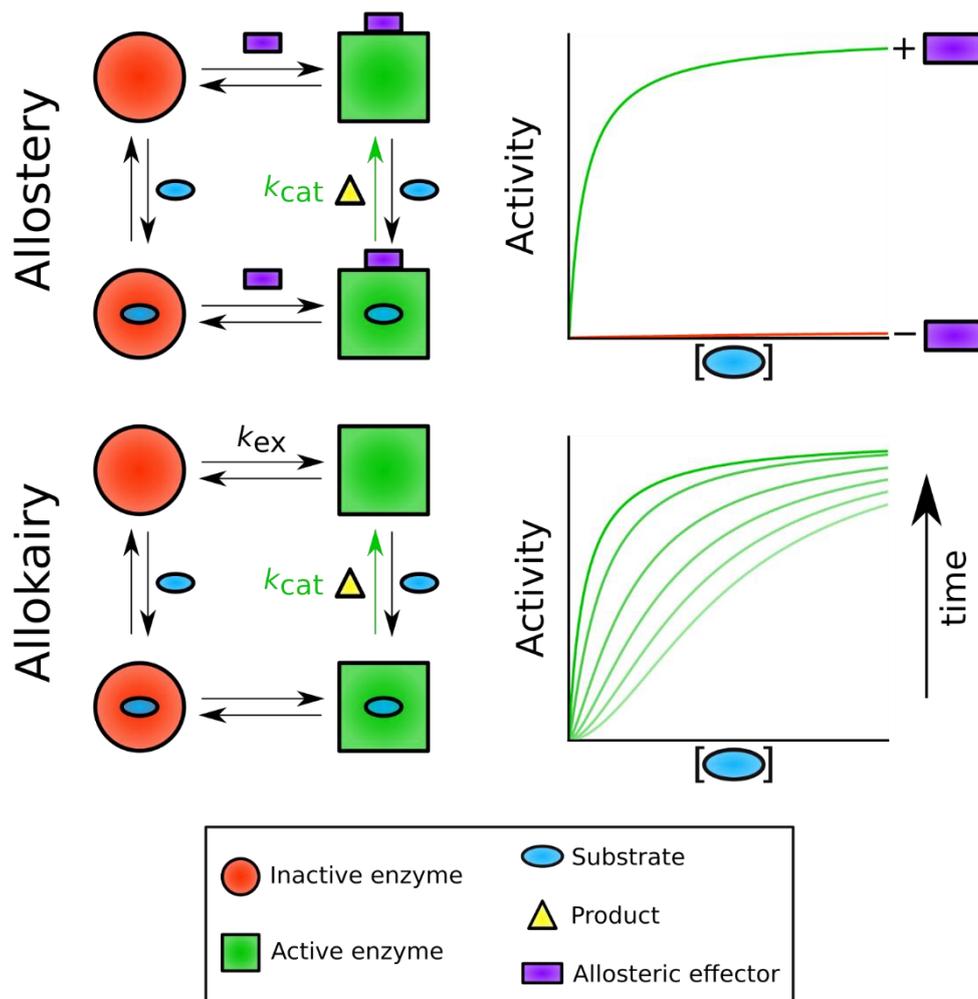
Figure 1.3 Allostery and allokairy mechanisms of regulation and activity profiles in monomeric enzymes. In allostery, binding (or reaction) of an allosteric effector (purple rectangle) on a site away from the active site facilitates the conformational change from an inactive form (red circle and red profile) to an active form (green square and green profile). In consequence, the substrate (blue oval) is transformed into the product (yellow triangle) at the catalytic rate (k_{cat}, green arrow). In allokairy, binding of the substrate in the active site shifts the enzyme conformation from an inactive form to an active form, at an exchange rate (k_{ex}) that is similar to k_{cat}, resulting in time-dependent activity profiles (gradient of light green to dark green profiles). Upon substrate depletion, the enzyme returns to the original equilibrium position.

1.5.1 Structure

Substrate-free βPGM exhibits an open conformation where the active site is located in the cleft formed between the α/β core domain (M1–D15, S88–K216) and the α-helical cap domain (T16–V87) (Figure 1.6a). The catalytic Mg^{2+} ion is coordinated by the backbone carbonyl of D10 together with the sidechains of D8, E169 and D170 (Figure 1.6b). Upon substrate binding, the cap domain rotates about 33–36⁰ at the interdomain hinge (V12–T16), leading to a closed transition state conformation (Baxter et al., 2010; Jin et al., 2014; Johnson et al., 2018). Phosphoryl transfer occurs from the substrate to the carboxylate nucleophile of residue D8 forming an aspartylphosphate group in the proximal site, adjacent to the catalytic Mg^{2+} ion. The carboxylate sidechain of residue D10 serves as a general acid-base, facilitating proton transfer reactions between the substrate and enzyme synchronous with phosphoryl transfer. The enzyme-bound phosphate group is coordinated by the backbone amide groups of L9, D10 and A115 and the sidechains of S114 and K145 (Figure 1.6b). The non-reacting phosphodianion group of the substrates binds in the distal site ~8 Å away in the closed enzyme (Baxter et al., 2010), which is coordinated by the backbone amide group of K117, the guanidinium group of residue R49 and the sidechains of residues S116 and N118 (Figure 1.6c).
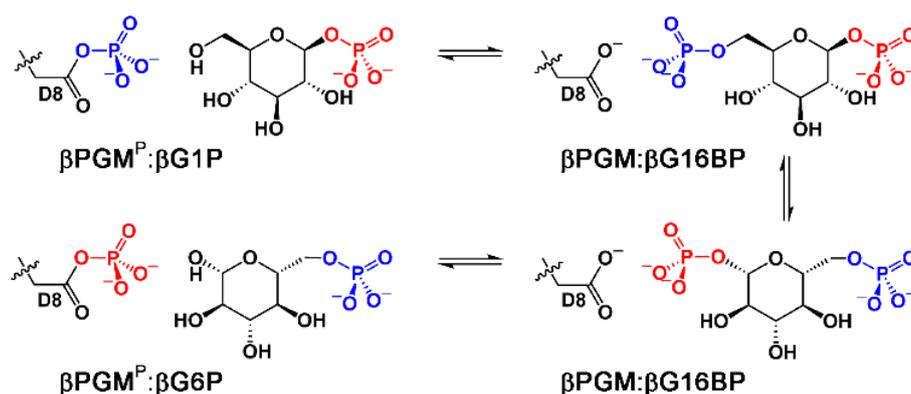
Figure 1.4 Reaction scheme for the βPGM catalytic cycle. βPGM catalyses the isomerisation between βG1P and G6P via a βG16BP intermediate. First, βPGM^P transfers its phosphate (blue) to the 6-OH of βG1P. The 1-phosphate (red) of βG1P is coordinated in the distal site. After releasing and binding in the alternative orientation, βPGM catalyses phosphoryl transfer of the 1-phosphate of βG16BP (red) to D8, producing G6P and regenerating βPGM^P. In this last step, the 6-phosphate (blue) of βG16BP is coordinated in the distal site.

1.5.2 Biological context

βPGM allows for the complete catabolism of maltose and trehalose in *L. lactis*. Trehalose is transported into the cell as trehalose 6-phosphate and then phosphorolysed by P_i-dependent trehalose 6-phosphate phosphorylase to βG1P and G6P (Andersson, Levander and Radstrom, 2001). In comparison, maltose is phosphorolysed to βG1P and glucose by the action of P_i-dependent maltose phosphorylase (Sjoberg and Hahnhagerdal, 1989). βPGM isomerises βG1P to G6P which then enters glycolysis via fructose 1,6-bisphosphate (F16BP). βPGM deficient *L. lactis* has impaired growth when the sole carbon source is trehalose or maltose (Levander, Andersson and Radstrom, 2001). With maltose, βG1P accumulates intracellularly and is excreted into the growth medium. Therefore, in trehalose and maltose catabolism, βPGM acts as the gatekeeper to and from glycolysis. βPGM is subjected to coarse control. Transcription of the βPGM gene (*pgmB*) (located in the *tre* operon) is subject to negative control by glucose and lactose (Qian et al., 1997). Transferring *L. lactis* from a glucose to a maltose-rich medium rises the specific activity of βPGM over a period of several hours. However, no fine control mechanism has yet been identified at basal levels of βPGM.

1.5.3 βPGM in vitro kinetics

In vitro βPGM activity experiments require the addition of a phosphorylating agent to prime the reaction since the half-life of βPGM^P is ~30 s (Golicnik et al., 2009; Johnson et al., 2018). When βG16BP is used as the phosphorylation agent, βPGM displays linear initial rates that can be successfully described by a conventional ping pong mechanism with substrate inhibition (Figure 1.7) (Golicnik et al., 2009). In contrast, when F16BP, βG1P, G6P, α-glucose 1,6-bisphosphate (αG16BP), or acetyl phosphate (AcP) are used, a significant lag phase is observed in its kinetic profile (Figure 1.7) (Golicnik et al., 2009; Johnson et al., 2018; Wood et al., 2020). Slow phosphorylation rates and competitive inhibition have been suggested as the source of the lag phase (Golicnik et al., 2009). Indeed, F16BP, βG1P, G6P and αG16BP are similar in structure to βG16BP (Zhang et al., 2005). However, a similar lag phase is observed when AcP is present, despite it being a highly reactive phosphate donor and structurally different to the other phosphorylating agents. Hence, other factors may be contributing to the lag phase of βPGM.
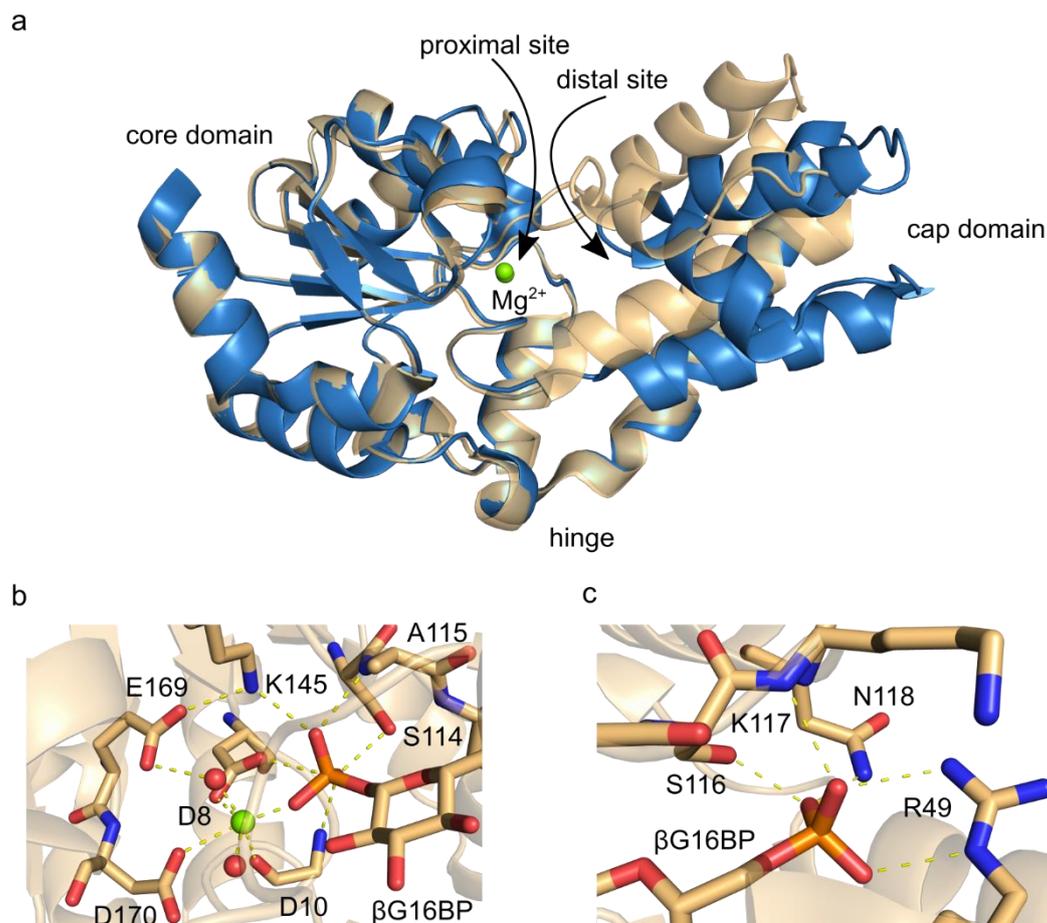
Figure 1.5 Global and local structure of βPGM. (a) Cartoon representation of the open substrate-free βPGM$_{WT}$ crystal structure (blue; PDB 6YDL) and the closed βPGM$_{D10N}$:βG16BP complex (orange; PDB 5OK1) highlighting the rotation of the cap domain due to substrate binding. (b) Reacting phosphodianion group coordination in the proximal site prior to phosphoryl transfer to the sidechain of D8. (c) Non-reacting phosphodianion group coordination in the distal site. The catalytic Mg^{2+} ion is coloured green. Water molecules are indicated as red spheres. Hydrogen bonds are presented as yellow dashes.

1.5.4 Metallofluorides as transition state analogues in βPGM

βPGM performs its function by catalysing phosphoryl transfer reactions between the carboxylate of D8 and the phosphodianion group of the substrate (Figure 1.5). Phosphoryl transfer is one of the most difficult reactions to work with in synthetic chemistry, due to the high stability of the phosphodianion group and the slow rates of reaction observed in solution (k_{non} = 2 × 10^{-20} s^{-1}) (Lad, Williams and Wolfenden, 2003; Bowler et al., 2010). In contrast, βPGM catalyses this reaction with a catalytic proficiency of 4 × 10^{26} M^{-1}, one of the highest proficiencies reported so far (Wolfenden and Snider, 2001). Hence, insights into the catalytic mechanism of βPGM, and

phosphoryl transfer enzymes in general, are of great interest, both from an enzyme catalysis and biosynthesis perspective.

Metallofluorides have been used to study phosphoryl transfer in enzymes due to their relative high stability in solution and their ability to function either as ground state analogues (GSA) or as transition state analogues (TSA) of the chemical step (Jin, Molt and Blackburn, 2017). In βPGM, NMR and X-ray crystallography studies of both GSA and TSA complexes have been performed using BeF_3^-, MgF_3^- and AlF_4^- moieties (Baxter et al., 2006; Baxter et al., 2008; Baxter et al., 2010; Griffin et al., 2012; Johnson et al., 2018). BeF_3^- displays a tetrahedral geometry with near-identical bond lengths and angles to a phosphodianion group. Thus, $βPGM:BeF_3$ and $βPGM:BeF_3$:hexose monophosphate complexes display GSA and NAC conformations (Figure 1.8a, b) (Griffin et al., 2012; Johnson et al., 2018). MgF_3^- is isoelectronic and isosteric with the TS of phosphoryl transfer (Jin, Molt and Blackburn, 2017) and hence $βPGM:MgF_3$:hexose monophosphate complexes adopt TSA conformations (Figure 1.8c) (Baxter et al., 2006; Baxter et al., 2009; Baxter et al., 2010). AlF_4^- is also isoelectric with the TS of phosphoryl transfer, but has a square-planar geometry. AlF_4^- anions free in solution are more stable than their MgF_3^- counterparts (Bodor et al., 2000). $βPGM:AlF_4$:hexose monophosphate complexes show TSA conformations (Figure 1.8d) (Baxter et al., 2008; Johnson et al., 2018).
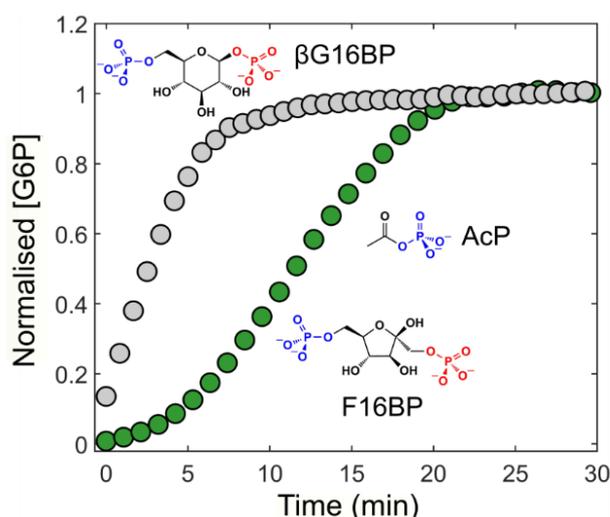
Figure 1.6 Activity of βPGM in the presence of different phosphorylating agents. During an in vitro kinetic experiment using a glucose 6-phosphate dehydrogenase coupled assay, linear initial rates are observed when βG16BP is used as a phosphorylating agent (grey circles). Conversely, in the presence of F16BP or AcP, βPGM activity displays a lag phase (green circles) and never reaches full activity.
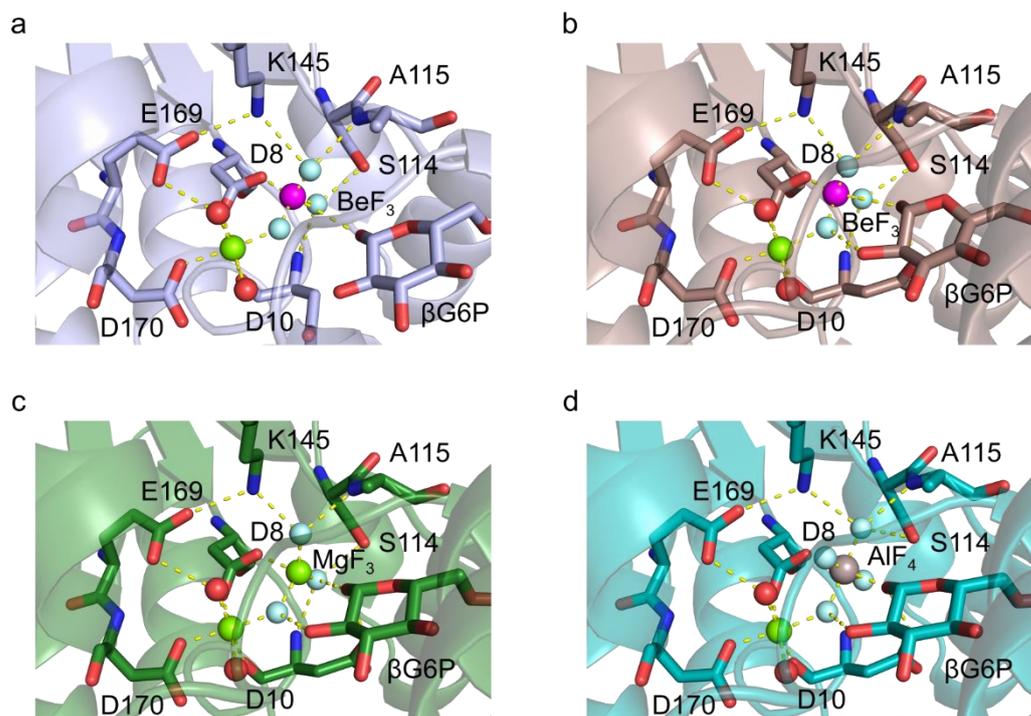
Figure 1.7 Metallofluorides as ground state and transition state analogues in βPGM. (a–b) The βPGM:BeF$_3$:G6P complex can adopt a (a) GSA (PDB:2WF9) or (b) NAC (PDB:2WF8) conformation. Notice the different position of the sidechain of D10 and the distance between the BeF$_3^-$ moiety and the 1-OH group of G6P between panels. (c) The βPGM:MgF$_3$:G6P complex (PDB:2WF5) displays a TSA conformation. (d) The βPGM:AlF$_4$:G6P complex (PDB:2WF6) also displays a TSA conformation. The Be^{2+}, Mg^{2+} and Al^{3+} cations are coloured magenta, green and grey respectively. F$^-$ anions are coloured cyan. Water molecules are illustrated as red spheres.

In conclusion, the structural, dynamic and unusual kinetic properties of βPGM, together with the amenability of substrate-free βPGM and βPGM metallofluoride complexes to high-resolution X-ray crystallography and NMR structural approaches makes this enzyme well-placed to investigate some of the key questions of how the conformational landscape of an enzyme controls its activity.

2 General Methods

An overview of the general methods used in this thesis is included. Each of the following results chapters contains a specific methods section for clarity.

2.1 Materials

Unless stated otherwise, all reagents were analytical grade and were purchased from Sigma-Aldrich (UK), Fischer Scientific, Alfa Aesar, Melford Laboratories and VWR. The isotopically-labelled compounds $^{15}NH_4Cl$ (99%), ^{13}C, $^2H_7$-D-Glucose (U–$^{13}C_6$, 99%; 1,2,3,4,5,6,6-d$_7$ 97–98%) and ^2H_2O (99.8%), which were purchased from CortecNet (France) and were used as received.

β-glucose 1-phosphate (βG1P) was synthesised enzymatically from maltose using maltose phosphorylase (EC 2.4.1.8) (Johnson et al., 2018). Briefly, a solution of maltose (600 mM) was incubated overnight with 1.2 units mL^{-1} maltose phosphorylase in 0.5 M sodium phosphate buffer (pH 7.0) at 30 °C. βG1P production was confirmed using ^{31}P NMR spectroscopy. Maltose phosphorylase was removed by filtration using a Vivaspin (5 kDa MWCO, Sartorius) at 4,500 rpm (Thermo Scientific Heraeus Labofuge 400 R) and the flow-through was used without further purification. The final yield was 24%.

β-glucose 1,6-biphosphate (βG16BP) was synthesised enzymatically using βPGM$_{D170N}$ (Wood et al., 2021). Briefly, βPGM$_{D170N}$ was incubated with 20 mM βG1P and 40 mM acetyl-phosphate (AcP) in 50 mM HEPES (pH 7.2) supplemented with 100 mM MgCl$_2$ at 25 °C for 265 min. The reaction was quenched by heat-denaturation of βPGM$_{D170N}$ at 90 °C for 10 min, and the protein pellet was separated by centrifugation at 4,500 rpm (Thermo Scientific Heraeus Labofuge 400 R). The resulting enzyme-free solution was passed through a 20/100 column packed with IR120 (H$^+$) ion-exchange resin, previously washed with 15 mL of ultrapure water. Then, the acidified solution was kept on ice and neutralised with 0.2 M barium hydroxide, resulting in significant precipitation of predominately the βG16BP barium salt. The precipitate was collected by centrifugation at 4,500 rpm (Thermo Scientific Heraeus Labofuge 400 R) at 4 °C. The pellet was solubilised in approximately 1 L of cold ultrapure water and passed through a 20/100 column packed with IR120 (Na$^+$) ion-exchange resin. The flow-through was lyophilised to leave a fine white powder as the final βG16BP sodium salt

product. Purity was assessed using ^{31}P, ^{1}H and natural abundance ^{13}C NMR spectroscopy. The final yield was 34%.

2.2 Buffers and solutions

All buffers and solutions were prepared using fresh ultrapure water (ddH$_2$O) produced by a Milli-Q® system (Merck Biosciences) and were sterilised by ultrafiltration through a 0.2 μm filter, unless stated otherwise. Table 2.1 contains the compositions of the buffers, media and solutions frequently used.

Table 2.1 Buffer and media compositions.

Solution	Use	pH	Components
Lysogeny broth	Protein expression	7.2	For 1 L: ▪ 10 g Tryptone ▪ 5 g Yeast extract ▪ 10 g NaCl Sterilised by autoclaving.
M9 salts	Isotopically labelled protein expression	7.4	For 1 L: ▪ 6 g Na$_2$HPO$_4$ ▪ 3 g KH$_2$PO$_4$ ▪ 0.5 g NaCl Sterilised by autoclaving. For $^{2}H^{13}C^{15}N$-labeled protein expression, $^{2}H_2O$ was used instead of ddH$_2$O, and the solution was sterilised by ultrafiltration through a 0.2 μm filter.
M9 minimal media	Isotopically labelled protein expression	7.4	To 1 L of M9 salts, the following was added: ▪ 1 mM MgSO$_4$ ▪ 0.1 mM CaCl$_2$ ▪ 1 mg thiamine ▪ 100 mg Ampicillin (Amp) ▪ 650 μL Trace elements (see below) ▪ 3 g glucose* ▪ 1 g NH$_4$Cl* *Isotopic labelled glucose and NH$_4$Cl were used when needed. 1 M MgSO$_4$ and CaCl$_2$ stock solutions were sterilised by autoclaving. All other solutions were sterilised by ultrafiltration through a 0.2 μm filter. For $^{2}H^{13}C^{15}N$-labeled protein expression, all solution were prepared on $^{2}H_2O$.
Trace elements	Isotopically labelled protein expression	8.0	To 100 mL of ddH$_2$O, the following was added: ▪ 550 mg CaCl$_2$·2H$_2$O ▪ 140 mg MnSO$_4$·H$_2$O ▪ 40 mg CuSO$_4$·5H$_2$O ▪ 220 mg ZnSO$_4$·7H$_2$O ▪ 45 mg CoCl$_2$·6H$_2$O ▪ 26 mg Na$_2$MoO$_4$·2H$_2$O ▪ 40 mg H$_3$BO$_3$ ▪ 26 mg KI ▪ 500 mg ethylenediaminetetraacetic acid (EDTA) ▪ 375 mg FeSO$_4$·7H$_2$O Sterilised by autoclaving.

Standard NMR buffer	βPGM purification and experiments	7.2	▪ 50 mM HEPES, adjusted with KOH ▪ 5 mM MgCl$_2$ ▪ 2 mM NaN$_3$
Standard kinetic buffer	βPGM experiments	7.2	▪ 200 mM HEPES, adjusted with KOH ▪ 5 mM MgCl$_2$ ▪ 2 mM NaN$_3$

2.3 Site-directed mutagenesis

Synthetic oligonucleotide primers for site-directed mutagenesis were synthesised by Eurofins Genomics (Germany). A QuickChange II Site-Directed Mutagenesis Kit was purchased from Agilent Technologies and a QIAprep miniprep kit was bought from Qiagen (Manchester, UK). The standard kit protocols were followed.

The primers used for βPGM$_{P146A}$ mutagenesis were:

Forward: 5'-TATCTGGTGCTGCTTTTGATGCTGCAACTTCAGCC-3'
Reverse: 5'-GGCTGAAGTTGCAGCATCAAAAGCAGCACCAGATA-3'

2.4 Production of isotopically labelled recombinant protein

2.4.1 ^{15}N–labelled recombinant protein overexpression

The genes of the relevant βPGM variants studied in this thesis were cloned into the pET22b(+) expression vector for recombinant protein production. The plasmid for each variant was transformed into *Escherichia coli* BL21(DE3) cells and cultures were grown at 37 °C with shaking to an OD$_{600}$ of 0.6 using 1 L ^{15}N–labelled M9 minimal media (Reed et al., 2003). The cell cultures were subsequently cooled to 25 °C and overexpression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. The cultures were then incubated with shaking for 18 hours and harvested by centrifugation at 10,000 rpm for 10 min (Beckman Coulter Avanti centrifuge, Rotor: F500). Dry cell pellets were stored at -80 °C prior to purification.

2.4.2 ^{2}H^{13}C^{15}N–labelled recombinant protein overexpression

For the production of ^{2}H^{13}C^{15}N–labelled βPGM$_{P146A}$, cells transformed with the relevant plasmid were adapted to 100% ^{2}H$_2$O M9 minimal media by stepwise increase of ^{2}H$_2$O content prior to protein overexpression. First, a 25 mL culture of 100% H$_2$O M9 minimal media was incubated at 37 °C with shaking overnight. Then, 100 μL of this culture was used to inoculated 25 mL of 90% ^{2}H$_2$O ^{15}N–labelled M9 minimal media, followed by incubation at 37 °C with shaking overnight. The last adaptation

step was achieved by adding 100 μL of the latter culture to 25 mL of 100% ^{2}H$_2$O ^{15}N–labelled M9 minimal media, which was then incubated at 37 °C with shaking overnight. The overnight cultures from the final step were used to inoculate 1 L of 100% ^{2}H$_2$O ^{13}C^{15}N–labelled M9 minimal media, and cells were grown, induced, harvested and stored as described above.

2.4.3 Protein purification

The cell pellets were resuspended in ice-cold standard NMR buffer supplemented with 1 mM EDTA and cOmplete^{TM} protease inhibitor cocktail (Roche). Cell lysis was achieved by sonication on ice for 6 × 20 s pulses separated by 60 s intervals. The cell lysate was separated from the insoluble cell debris by centrifugation at 20,000 rpm for 40 min at 4 °C (Beckman Coulter Avanti centrifuge, Rotor: JA-20). The cleared soluble fraction was loaded onto a DEAE-Sepharose fast flow anion-exchange column connected to an ÄKTA Prime purification system (GE Healthcare) that had been previously washed with 1 column volume of 6 M guanidine hydrochloride, 1 column volume of 1 M NaOH and equilibrated with 3–5 column volumes of standard NMR buffer supplemented with 1 mM EDTA. Elution of proteins bound to the column was achieved with a gradient of 0 to 50% standard NMR buffer containing 1 mM EDTA and 1 M NaCl over 300 mL. Fractions were checked for the presence of βPGM by SDS-PAGE, pooled together and concentrated to 3–5 mL using centrifugation in a Vivaspin (10 kDa MWCO, Sartorius) at 4,500 rpm and 4 °C (Thermo Scientific Heraeus Labofuge 400 R). Further purification was achieved by loading the protein sample onto a pre-packed packed Hiload 26/600 Superdex 75 size-exclusion column connected to an ÄKTA Prime purification system, previously washed with degassed 1 M NaOH and equilibrated with 1.5 column volumes of standard NMR buffer containing 1 mM EDTA and 1 M NaCl. The presence of a high concentration of NaCl during elution minimises the contamination of recombinant βPGM with *E. coli* enzymes (e.g., phosphoglucose isomerase and βPGM) (Johnson et al., 2018). Fractions containing βPGM were checked for purity by SDS-PAGE, pooled together and buffer exchanged into standard NMR buffer and concentrated to 1–1.6 mM by Vivaspin (10 kDa MWCO, Satorius) and storage as 0.5 mL aliquots at -20 °C.

No additional procedure was required to promote back exchange to amide protium atoms in ^{2}H^{13}C^{15}N–labelled βPGM$_{P146A}$.

Protein concentration was measured by absorbance at 280 nm (βPGM molecular weight = 24.2 kDa, ε_{280} = 19,940 M^{-1} cm^{-1}) using a Nanodrop OneC (Thermo Scientific).

2.5 Nuclear Magnetic Resonance Spectroscopy

2.5.1 Background

Nuclear magnetic resonance (NMR) spectroscopy is a powerful method to study the structure and the dynamics of molecules, and in particular proteins, in solution. It can provide information about the chemical environment of each of the atoms of a given protein with exquisite resolution and without perturbing the sample. The development over the past 50 years of more sensitive NMR spectrometers and increasing magnetic field strengths (Cavanagh, 2007), together with a robust theoretical background and advances in experimental and computational tools and resources (Mittermaier and Kay, 2009; Felli and Brutscher, 2009), have made NMR spectroscopy a first-choice technique in the fields of biochemistry, structural biology and biomolecular dynamics (Kleckner and Foster, 2011). Additionally, advances in isotopic labelling have increased the range of protein and protein complex sizes accessible to NMR techniques (Gardner and Kay, 1998). The following section summarises the key NMR concepts and experiments that are relevant to this thesis.

2.5.2 Theory

NMR spectroscopy is based in the interaction between strong magnetic fields (between 6–24 T, compared to 3×10^{-5} T from the Earth's magnetic field) (Finlay et al., 2010), generated by superconducting coils submerged in liquid helium at -269 °C (4.2 K), and the spin magnetic moment of the magnetically active nuclei in the sample (Levitt, 2008). In biochemistry, the commonly used magnetically active nuclei are ^{1}H, ^{13}C, ^{15}N, ^{19}F and ^{31}P (Kwan et al., 2011). The concept of spin is highly abstract, and its complete understanding requires the ability to manipulate the quantum mechanical equations. Instead, it is possible to approximate the behaviour of the spin, and its interaction with an external magnetic field, by describing it as an angular momentum (Levitt, 2008).

The spin angular momentum and the spin magnetic moment are proportional to each other by the gyromagnetic ratio (γ):

$$\hat{\mu} = \gamma \hat{S} \hspace{6cm} 2.1$$

and hereafter they will be used interchangeably.

The angular momentum of the spin is a vector. Its direction indicates the axis of rotation. For a given free nucleus in solution, (all the sample conditions used in this thesis allow for free diffusion of molecules), the angular momentum vector may point to any possible direction in space. In the absence of a magnetic field, the distribution of magnetic moments is isotropic. Hence, the total magnetic moment of the sample is zero. When the sample is loaded into the NMR spectrometer, the angular momentum vector moves on a cone, keeping a constant angle between the vector and the static magnetic field (conventionally parallel to the z-axis). This motion is called spin precession. The frequency of precession is called the Larmor frequency (ω), which is proportional to the external magnetic field (**B**$_0$):

$$\omega = -\gamma \mathbf{B}_0 \hspace{6cm} 2.2$$

The Larmor frequency of the ^{1}H nucleus (also referred simply as proton) is commonly used to represent the strength of an NMR spectrometer. For example, in a 500 MHz spectrometer (magnetic field strength of 11.74 T), an individual proton does 5×10^8 rotations per second. For comparison, at the same magnetic field strength, a ^{13}C nucleus has a Larmor frequency of 125 MHz, and a ^{15}N nucleus has a frequency of 50 MHz (Levitt, 2008).

2.5.3 Longitudinal Relaxation (T$_1$)

Let us consider the proton nuclei in a protein sample. In the absence of an external magnetic field, all the magnetic moments are distributed isotropically. When the same sample is inserted in an NMR spectrometer, the magnetic moments start to precess around the field. However, if one could immediately measure the total magnetic moment of the sample, it would be zero, as the angular momentum vectors are still distributed isotropically. In the sample, the individual protein molecules are freely diffusing, i.e., their position and relative orientations are changing constantly. Free diffusion generates a time-dependent local magnetic field due to two main reasons: (i) transient interactions mediated by electrons within a particular protein molecule and

between protein molecules are formed and destroyed and (ii) magnetic nuclei within the protein molecules interact with each other. The observed Larmor frequency (ω_{obs}) then becomes:

$$\omega_{obs} = -\gamma \mathbf{B}_{eff} = -\gamma(\mathbf{B}_0 + \mathbf{B}_{local}(t)) \qquad 2.3$$

where the effective magnetic field ($\mathbf{B}_{eff}$) is the sum of the static magnetic field ($\mathbf{B}_0$) and the local field ($\mathbf{B}_{local}(t)$). Moreover, the molecular diffusion is biased towards spin orientations parallel to $\mathbf{B}_{eff}$, resulting in an anisotropic distribution of magnetic moment. As a result, the sample total magnetic moment on the z-axis ($M_z(t)$) will start to build up exponentially:

$$M_z(t) = M_{eq}(1 - e^{-\frac{t}{T_1}}) \qquad 2.4$$

where M_{eq} is the magnetisation at equilibrium and the time constant T_1 represents the longitudinal relaxation, i.e., the average time the sample takes to reach equilibrium (Figure 2.1).
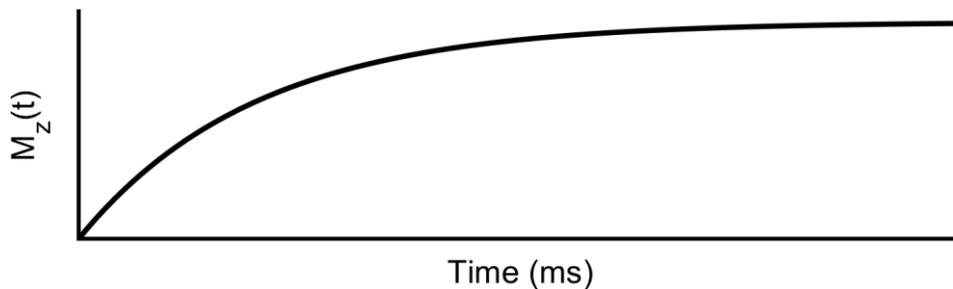
Figure 2.1 Longitudinal magnetisation. The magnetisation along the z-axis starts to build up as soon as the sample is loaded inside the magnetic field $\mathbf{B}_0$.

The longitudinal relaxation time determines how frequently data can be acquired. In proteins, proton T_1 values typically range between 0.2–2 seconds (Kleckner and Foster, 2011), and are dependent on the size of the protein, the temperature, the viscosity of the solvent and the magnitude of $\mathbf{B}_0$.

2.5.4 Transverse Magnetisation and Transverse Relaxation (T_2)

At equilibrium, the sample M_{eq} is orders of magnitude smaller than the static magnetic field $\mathbf{B}_0$, which are both aligned parallel to the z-axis. Hence, M_{eq} is technically difficult to measure. Instead, it is easier to measure the total magnetic moment perpendicular

to $\mathbf{B}_0$. To achieve this, a radiofrequency pulse (an oscillating magnetic field, the frequency of which depends on the nucleus of interest) is applied to the sample. The resulting transient magnetic field rotates the sample total magnetic moment to the xy-plane. This magnetic moment perpendicular to $\mathbf{B}_0$ is called transverse magnetisation. These pulses have a duration between a few microseconds to several milliseconds. Immediately after the pulse is turned off, the individual spins start to precess on their cones at their respective ω_{obs}. Since every spin rotates, the total magnetic moment also rotates, now on the xy-plane. The transverse magnetisation has an x and y-component with the form:

$$M_x(t) = M_{eq} \sin(\omega_{obs}t)\, e^{-\frac{t}{T_2}}$$
$$M_y(t) = -M_{eq} \cos(\omega_{obs}t)\, e^{-\frac{t}{T_2}}$$

2.5

where the time constant T_2 represents the transverse relaxation. Since the $\mathbf{B}_{eff}$ of each individual spin is slightly (but measurably) different, the precessing magnetic moments gradually become out of phase until their distribution becomes isotropic (Figure 2.2).
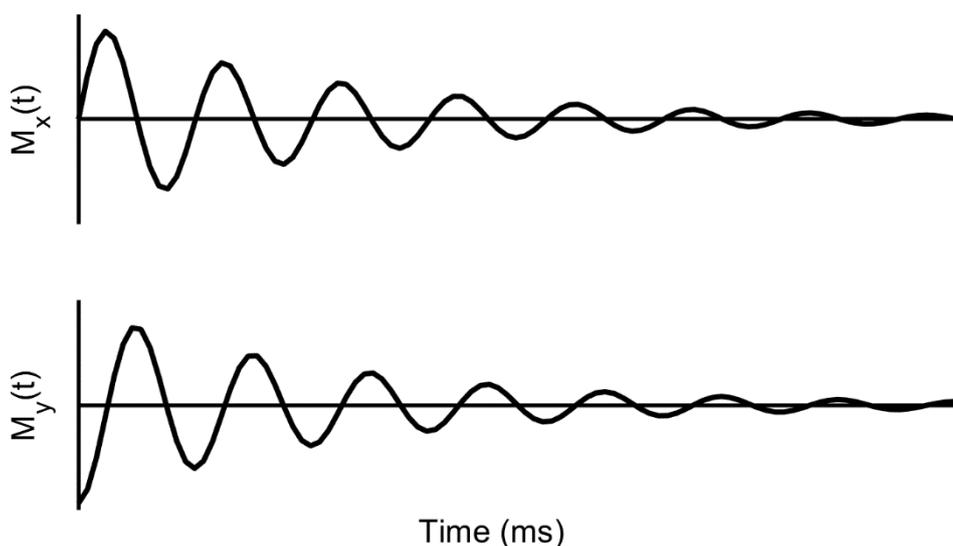
Figure 2.2 Transverse magnetisation. The magnetisation on the xy-plane oscillates, as the total magnetic moment rotates at ω_{obs}, and decays with a time constant T_2, due to the loss of phase coherence of the spins with time.

In consequence, the total transverse magnetisation becomes zero. In proteins, proton T_2 values typically range between 20 and 200 milliseconds (Kleckner and Foster, 2011), and also depend on the size of the protein, the magnitude of $\mathbf{B}_0$ and the sample

conditions. Small T_2 values, indicating the rapid loss of coherent magnetisation, is the primary obstacle of performing NMR spectroscopy in large proteins and biological assemblies (>20 kDa) (Tugarinov, Hwang and Kay, 2004; Foster, McElroy and Amero, 2007; Ruschak and Kay, 2010).

2.5.5 The free-induction decay and the chemical shift

NMR spectrometers are designed to detect the oscillating transverse magnetisation. The probe, located within the core of the spectrometer, contains a series of wire coils arranged perpendicularly around the sample. Due to Faraday's law, an oscillating magnetic field induces an electric current in the wire, which is recorded by a receiver. This oscillating electric current is called the free-induction decay (FID); this is the NMR signal recorded from the sample. The amplitude of an individual FID is small, sometimes in the same order of magnitude of the experimental noise. Hence, to enhance the signal-to-noise ratio of NMR experiments, multiple identical FIDs are acquired and added together. This process is called signal averaging. Given that the noise behaves randomly, the signal-to-noise ratio is proportional to the square root of the number of FIDs acquired (Levitt, 2008).

The FID has the same function and behaviour as the transverse magnetisation (Equation 2.5; Figure 2.2). The FID contains the ω_{obs} and the T_2 of all the spins from the sample. These values are extracted by applying a Fourier transform (FT), which translates the FID function from the time domain to the frequency domain. In other words, a Fourier transform converts a set of NMR signals (FIDs) into an NMR spectrum (Figure 2.3).
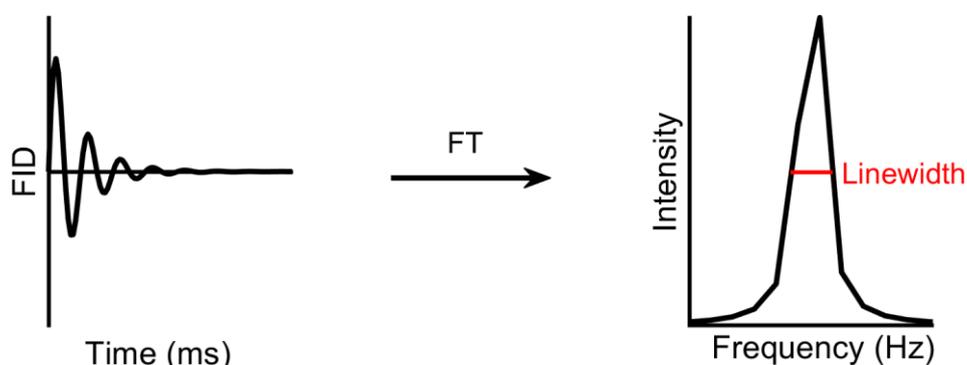
Figure 2.3 Fourier transform NMR spectroscopy. The application of a Fourier transform on the FID signal allows the analysis of the three main NMR observables: intensity, frequency and linewidth.

A given peak in an NMR spectrum has the following mathematical form:

$$I(\omega) = \frac{\frac{\lambda}{2}}{\left(\frac{\lambda}{2}\right)^2 + (\omega - \omega_{\text{obs}})^2} \qquad 2.6$$

where λ is the width of the peak, also known as the linewidth, and the peak maximum is centred at ω_{obs}. The linewidth is inversely proportional to the transverse relaxation:

$$\lambda = \frac{2}{T_2} \qquad 2.7$$

Commonly, the intensity of the peak (I) is also used to interpret NMR data. The intensity is defined as the peak height at its maximum.

The ω_{obs} and T_2 of an individual spin contains information on its local electronic environment and its interaction with neighbouring magnetic nuclei. Most of the structural and dynamical properties of molecules, including proteins, can be inferred from the ω_{obs} and T_2 of its spins (Mittermaier and Kay, 2009; Kleckner and Foster, 2011). However, ω_{obs} is also scales proportionally to $\mathbf{B}_0$. Therefore, the field-independent chemical shift (δ) is used instead to facilitate the comparison between experiments acquired using different spectrometers. The chemical shift is the scaled and normalised ω_{obs} with respect to a reference compound (ω_{ref}):

$$\delta_{obs} = \frac{\omega_{\text{obs}} - \omega_{\text{ref}}}{\omega_{\text{ref}}} \qquad 2.8$$

Chemical shifts are small numbers, usually reported in terms of parts per million (ppm). Trimethylsilylpropanoic acid (TSP) was used as the reference compound in this thesis.

2.5.6 Multidimensional NMR spectra

Due to technical limitations, it is not possible to measure simultaneously the chemical shift of all magnetically active nuclei within the sample, in the same experiment. Instead, biomolecular NMR experiments leverages the magnetic coupling between chemically bonded nuclei, named *J*-coupling, to obtain the heteronuclear chemical shift of adjacent atoms (Figure 2.4a). These experiments are referred to as multidimensional, as each dimension corresponds to a different nucleus. The intensity component of each *J*-coupled correlation is visualised by plotting multidimensional spectra as contour plots. For example, the spectrum of a standard two-dimensional $^1H^{15}N$ heteronuclear single quantum coherence (HSQC) experiment (Figure 2.4b). contains a peak for each backbone amide group (i.e., one for each non-proline peptide bond), a peak for each tryptophan sidechain indole group, pairs of peaks from each asparagine or glutamine sidechain carboxamide group and occasionally sidechain H-N groups from arginine, histidine and lysine residues that are protected from proton exchange.
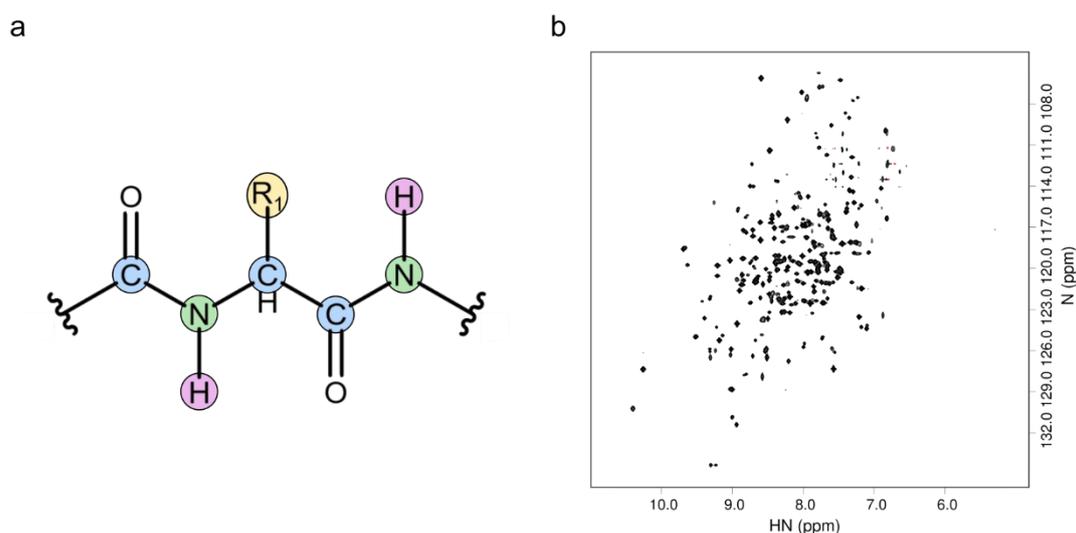
Figure 2.4 Multidimensional protein NMR spectroscopy. (a) Frequently observed nuclei in a multidimensional NMR experiment. ^1H–^{15}N correlations are preferred to track backbone conformational changes. ^1H–^{13}C are best suited for sidechain (yellow) inspection. (b) Example of a $^1H^{15}N$ HSQC spectrum of a 24 kDa (221 residue) protein.

2.5.7 Chemical exchange in NMR

Consider an individual protein molecule diffusing freely in solution. Even at equilibrium, the protein molecule experiences structural fluctuations and conformational changes at different timescales (Wei et al., 2016). Some, if not all, of these structural motions are relevant to the protein biological functions of proteins and enzymes (Henzler-Wildman and Kern, 2007). During these structural motions, the nuclei are interchanging between at least two distinct chemical environments, hereafter named states. This process is referred to as chemical exchange and it strongly influences their chemical shifts and their observed T_2. In this thesis, only two-state exchange processes were considered:

$$A \overset{k_B}{\underset{k_A}{\leftrightarrow}} B \qquad\qquad 2.9$$

where the equilibrium constant is $K_{eq} = k_B/k_A$. The final appearance of the NMR spectrum depends on the populations of each state, p_A and p_B, and the relative values of the absolute difference of chemical shift between states $\Delta\delta = |\delta_A - \delta_B|$ (in Hz), and the exchange rate $k_{ex} = k_A + k_B$. In NMR spectroscopy, three distinct exchange regimes (Figure 2.6) can be distinguished:

2.5.7.1 Slow exchange ($k_{ex} \ll \Delta\delta$)

The slow exchange regime is observed when k_{ex} is significantly smaller than $\Delta\delta$. For a particular nucleus, two distinct peaks appear in the NMR spectrum, each one reflecting the chemical shifts and linewidths of their respective states (Figure 2.6a). The intensity of each peak is proportional to the population of that state (although in practice the intensity is also affected by other experimental parameters). Slow exchange processes happen between the sub- and multi-second timescales. Common examples in proteins are protein folding, ligand tight binding and *cis-trans* proline isomerisation (Williamson, 2013; Lu et al., 2007).

2.5.7.2 Intermediate exchange ($k_{ex} \approx \Delta\delta$)

The intermediate exchange regime , or coalescence, occurs when k_{ex} and $\Delta\delta$ have similar values. Only one peak can be observed at a chemical shift in between δ_A and δ_B. However, the hallmark of intermediate exchange is the significant increase in peak linewidth, also known as exchange broadening (Figure 2.6b). The observed linewidth is proportional to k_{ex}. In some cases, the peak is broadened beyond detection, i.e., the

peak intensity is below the noise level. Processes that exhibit intermediate exchange are in the millisecond timescale, and include enzyme catalysis, domain movement, partial folding/unfolding and ligand binding (Boehr et al., 2006).

2.5.7.3 Fast exchange ($k_{ex} \gg \Delta\delta$)

A nucleus is in the fast exchange regime when k_{ex} is significantly larger than $\Delta\delta$. The observed peak has a population-weighted chemical shift ($\delta = p_A\delta_A + p_B\delta_B$) and linewidth (Figure 2.6c). Fast exchange processes occur on the picosecond to microsecond timescale. Some examples are ligand weak binding, loop motion and deprotonation.
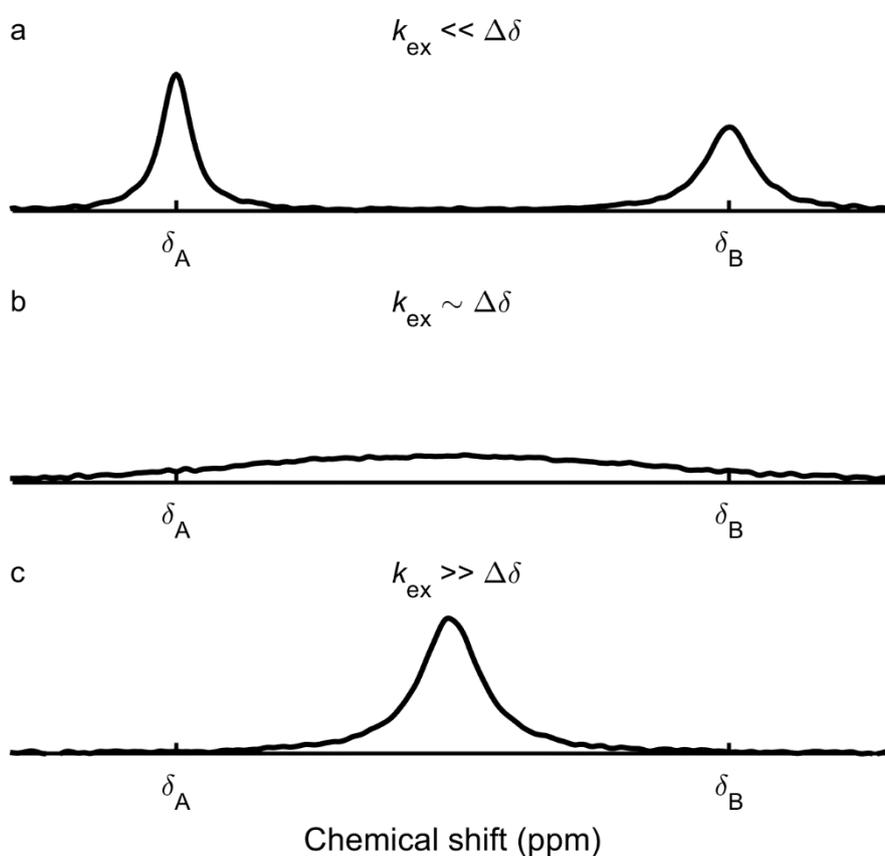
Figure 2.5 Exchange regimes in NMR spectroscopy. All the three NMR observables are affected by chemical exchange. (a) In slow exchange ($k_{ex} \ll \Delta\delta$), peaks for both states are observed showing their respective chemical shifts (δ_A and δ_B), intensities and linewidths. (b) At intermediate exchange ($k_{ex} \sim \Delta\delta$), only one peak with an intermediate chemical shift and broad linewidth is observed. Is common for the intensity of this broad peak to be below noise level. (c) In fast exchange ($k_{ex} \gg \Delta\delta$), an individual peak with a population-weighted chemical shift and linewidth is observed.

It is possible to determine the exchange regime of a certain process by shifting the equilibrium between the A and B states and following the change in the intensity, chemical shift and linewidth of the protein peaks (Williamson, 2013). For example, when following the titration of a protein with its ligand by ^{1}H^{15}N HSQC experiments (Figure 2.7), three different behaviours can be observed:

i. If the process is in slow exchange, the intensity of the unbound protein peak decreases, while a new protein-ligand complex peak appears (Figure 2.7a). The change in intensity of either state as a function of ligand concentration can be used to calculate the affinity between the protein and the ligand.

ii. In the intermediate exchange regime, the observed peak starts to change its chemical shift, i.e., moves, from the unbound to the bound state (Figure 2.7b). However, the peak linewidth also starts to increase due to the exchange between states, and the peak broadens beyond detection before reaching the chemical shift of the bound state. It is technically challenging to determine the protein-ligand affinity, as the signal disappears prior to reaching saturation.

iii. At the limit of the fast exchange, the peak moves from the unbound state to the bound state without a significant change in its linewidth (Figure 2.7c). The protein-ligand affinity can be estimated from the change in the chemical shift of the peak with respect to the ligand concentration.
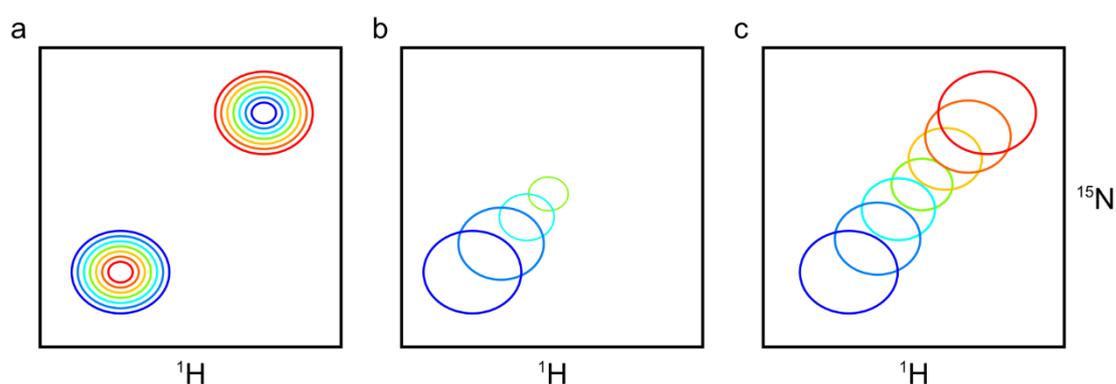
Figure 2.6 Chemical exchange in a ^{1}H^{15}N HSQC experiment. (a) In slow exchange, the intensity of the free protein (lower left) peak decreases while the (upper right) peak intensity increases with increasing concentration of ligand (blue to red rainbow). (b) In intermediate exchange, the free peak moves towards the peak, but it "broadens beyond detection" before reaching ligand saturation. (c) In fast exchange, the peak smoothly moves from the free protein to the protein-ligand complex chemical shift with increasing concentration of ligand (blue to red rainbow).

2.5.8 NMR methods and experiments

NMR spectroscopy has an array of approaches that provides insights into the structure and the conformational landscape of proteins at different. The next section gives a brief explanation of the main methods performed in this thesis.

2.5.8.1 Transverse relaxation-optimised spectroscopy.

Increased transverse relaxation is the main obstacle in the study of large proteins (>20 kDa) by NMR methods. It broadens the peaks, impairing both the sensitivity and the resolution of the NMR spectrum. At the high magnetic fields (600–1200 MHz) used in biomolecular NMR spectroscopy, a significant source of relaxation comes from dipole-dipole (DD) coupling, due to the magnetic interaction between neighbouring nuclei, and chemical shift anisotropy (CSA), which arises from the asymmetry of the electronic environment around nuclei. Both DD coupling and CSA have constructive and destructive interactions that differentially affect the final peak linewidth.

Transverse relaxation-optimised spectroscopy (TROSY) leverages both the DD coupling and CSA to increase the resolution of NMR spectra by reducing peak linewidths. The peaks of a $^1H^{15}N$ HSQC experiment (Figure 2.8a) can be decomposed in the constructive and destructive DD coupling and CSA interactions (Figure 2.8b). A TROSY-based $^1H^{15}N$ HSQC experiment (hereafter named $^1H^{15}N$ TROSY) selects for the narrowest component, where only the destructive CSA and DD coupling interactions are selected (Figure 2.8c). TROSY-based versions of both 2D and 3D heteronuclear experiments were used in this thesis.
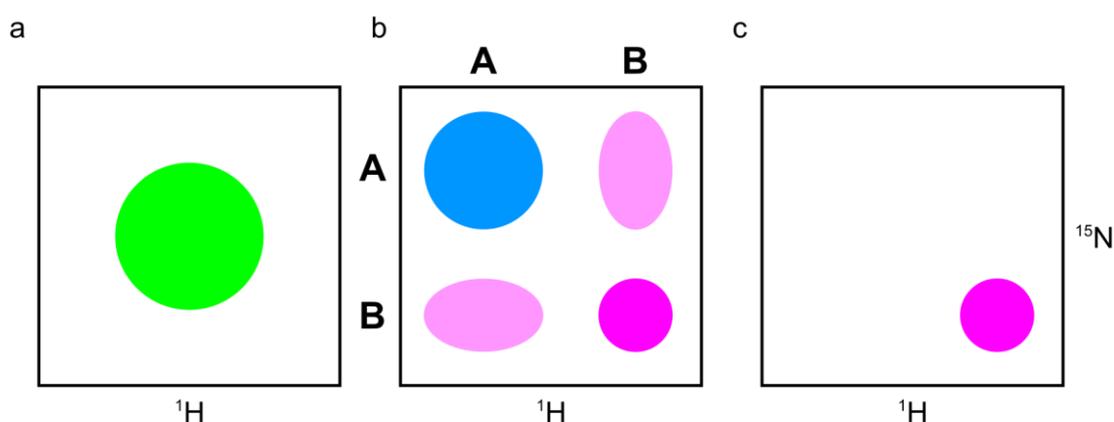
Figure 2.7   Transverse relaxation-optimised spectroscopy. (a) A peak in a conventional $^1H^{15}N$ HSQC spectrum. (b) A peak in a coupled $^1H^{15}N$ HSQC spectrum, where the constructive (**A**) and destructive (**B**) interactions of the DD coupling and CSA can be differentiated. (c) The TROSY-based $^1H^{15}N$ HSQC selects for the narrowest (**B**) component of the interaction between the DD coupling and CSA.

2.5.8.2 Protein backbone assignment

The chemical shifts of the protein backbone atoms, ^1H_N, ^{15}N, $^{13}C_\alpha$, $^{13}C_\beta$ and $^{13}C'$, are excellent reporters of secondary structure environments and exchange dynamics (Figure 2.4a). These structural features and motions can be investigated by NMR at atomic resolution. However, the peaks of an NMR spectrum need to be assigned to their respective atoms in the protein molecule before any major inferences can be made. Backbone assignment is possible by harnessing the *J*-coupling between adjacent backbone atoms to correlate their chemical shifts. Additionally, the assignment of these correlations is facilitated with the use of multidimensional NMR experiments (Tugarinov, Hwang and Kay, 2004).

One obstacle of multidimensional NMR spectroscopy of high molecular weight proteins and protein complexes (>20 kDa) is their small T_2 values. During a multidimensional NMR experiment, the magnetisation is transferred between adjacent nuclei (Kwan et al., 2011). At each step, the signal gets attenuated due to the transverse relaxation of each of the atoms involved. This is particularly true for ^{13}C nuclei. A common approach to increase T_2 values is to substitute carbon-bearing protons (^1H) with deuterons (^2H). The significant lower gyromagnetic ratio of ^2H compared to ^1H diminishes the strength of the nuclei interactions that cause transverse relaxation (Gardner and Kay, 1998).

In this thesis, 3D experiments that correlate the ^1H_N and ^{15}N atoms of residue *i* with the $^{13}C_\alpha$, $^{13}C_\beta$ or $^{13}C'$ atoms of both residues *i* and *i*-1, or just residue *i*-1, were used to unambiguously assign peaks in the $^1H^{15}N$ TROSY spectrum (Figure 2.9). For example, the HNCA experiment correlates the ^1H_N and ^{15}N peaks of residue *i* with the $^{13}C_\alpha$ peaks of both *i* and *i*-1 residues. The discrimination of the $^{13}C_\alpha$ chemical shifts is achieved by the HNCOCA experiment, which only records the $^{13}C_\alpha$ peak of residue *i*-1. The same procedure was used to identify the $^{13}C_\beta$ and $^{13}C'$ peaks of residues *i* and *i*-1, using the HNCACB/HNCOCACB and HNCACO/HNCO pairs of experiments. The peak picking process was performed using FELIX (Felix NMR, Inc.).

The next step is to frequency match the chemical shifts of the $^{13}C_\alpha$, $^{13}C_\beta$ and $^{13}C'$ backbone assignments. This matching process was achieved using the program *asstools* (Reed et al., 2003), which relies on a simulated-annealing algorithm to assign the chemical shifts to their respective residues in the protein molecule by frequency matching between sequential residues. The final lists of backbone chemical shifts ordered by residue number were deposited in the public repository BioMagResBank (http://www.bmrb.wisc.edu/).
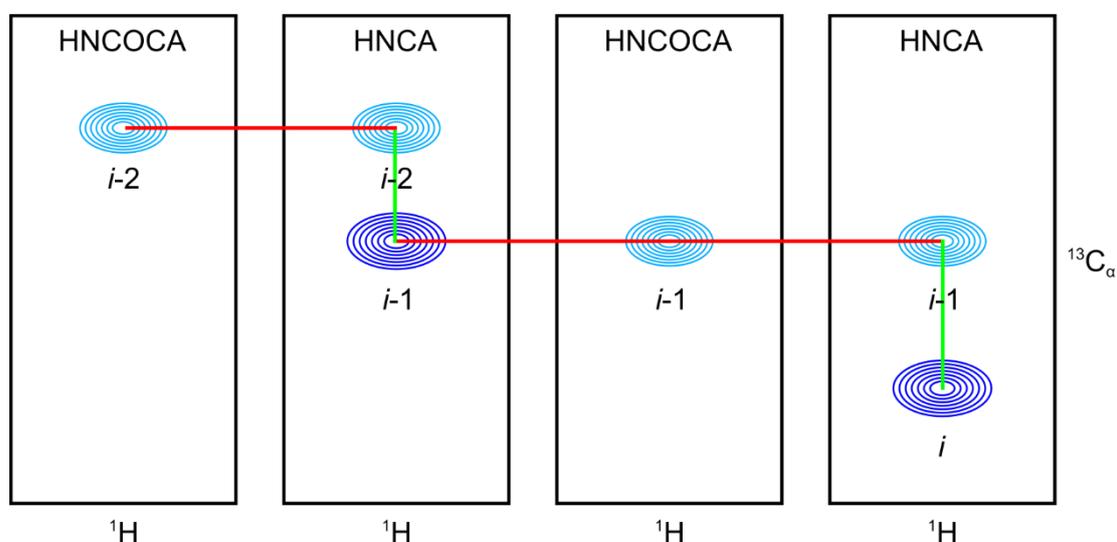
Figure 2.8 Protein backbone NMR assignment strategy. Schematic representation of the $^{1}H^{13}C$ slices of 3D HNCA and HNCOCA spectra. The simultaneous analysis of both experiments allows for the discrimination between $^{13}C_\alpha$ peaks of residues i and i-1 (green line). After all peaks have been picked, they are sequentially connected by frequency matching (red line).

2.5.8.3 Non-uniform sampling of NMR data

Protein NMR spectroscopy is known as an information rich but insensitive technique (Kwan et al., 2011). Despite the recent hardware advances, particularly the improvement of superconducting magnets and the incorporation of supercooled NMR probes (cryoprobes) (Flynn et al., 2000), the still relatively high sample concentration requirement creates a trade-off between spectral resolution, sensitivity and acquisition time, especially in time-sensitive samples and experiments. Non-uniform sampling (NUS) is a method to enhance resolution while maintaining accessible experimental times (Hyberts, Robson and Wagner, 2013). Alternatively, NUS was used in this thesis to maintain optimal signal sensitivity while reducing experimental acquisition time to

record transient, thermodynamically unstable species. Briefly, NUS is based on the assumption that in a conventional, uniformly sampled NMR spectrum, useful data is sparse compared to the actual amount of data typically recorded. Hence, in an NUS spectrum, only a proportion of the data is recorded, and the missing data is back-calculated using published reconstruction algorithms (Hyberts et al., 2012). The sampling schedules used in this thesis were generated using a Poisson Gap Sampling method with a sinusoidal weight value of 2 (Hyberts, Takeuchi and Wagner, 2010). The NUS spectra were reconstructed using the Compressed Sensing algorithm (Holland et al., 2011) implemented in TopSpin 4.0.6.

2.5.8.4 TALOS-N

Protein backbone chemical shifts contain information from the local peptide bond geometry, particularly the backbone torsion angles φ and ψ (Wishart, Sykes and Richards, 1991). Two different approaches can be used to establish quantitative relationships between heteronuclear chemical shifts and protein structure: quantum chemistry methods (De Dios, Pearson and Oldfield, 1993) and database-derived empirical methods. The program TALOS-N belongs to the latter category (Shen and Bax, 2013). TALOS-N relies on a curated database of assigned proteins with known structure. It predicts the torsion angles, and therefore the secondary structure, that corresponds to a given backbone chemical shift assignment by using a previously trained artificial neural network. Then it filters out those predictions by matching them to heptapeptides with the same sequence from the database. This two-step process improves coverage and reliability. Moreover, TALOS-N also outputs the random coil index order parameter (RCI-S^2) of the backbone residues (Berjanskii and Wishart, 2005), which allows for an independent validation the secondary structure prediction.

2.5.8.5 Real-time NMR with rapid mixing

Dynamic processes on the multi-second timescale can be monitored by quantifying the time-dependence of the peak intensities upon perturbation of the equilibrium state. Real-time NMR has been used to monitor protein folding (Zeeb and Balbach, 2004) and slow conformational changes (Corazza et al., 2010). Such conformational changes are triggered by the use of a special injection device (Franco et al., 2017) or a laser light (Harper et al., 2004) to perturb the sample inside the spectrometer and data acquisition is initiated rapidly.

To observe the species present immediately following the addition of phosphorylating agent to βPGM$_{WT}$, NMR experiments were recorded with the use of a homemade rapid mixing device. The equipment comprised a 2 m length of 0.8 mm internal diameter EFTE tubing (GE Healthcare), connected at one end to a 1 mL syringe and inserted at the other end through the lid of an NMR tube. The tubing was loaded with phosphorylating agent and a small, additional volume of air was drawn in to prevent premature mixing of the phosphorylating agent with the protein sample (Figure 2.10a). The rapid mixing device was loaded into a Bruker 600 MHz Neo spectrometer and allowed to equilibrate thermally at 298 K (Figure 2.10b). Following mixing by syringe action of the phosphorylating agent with the protein sample, the spectrometer was locked, tuned and shimmed, and the ^{1}H 90° pulse length was calibrated. A series of ^{1}H^{15}N BEST-TROSY spectra were recorded as 142 s experiments. The experimental dead-time was approximately 156 s.
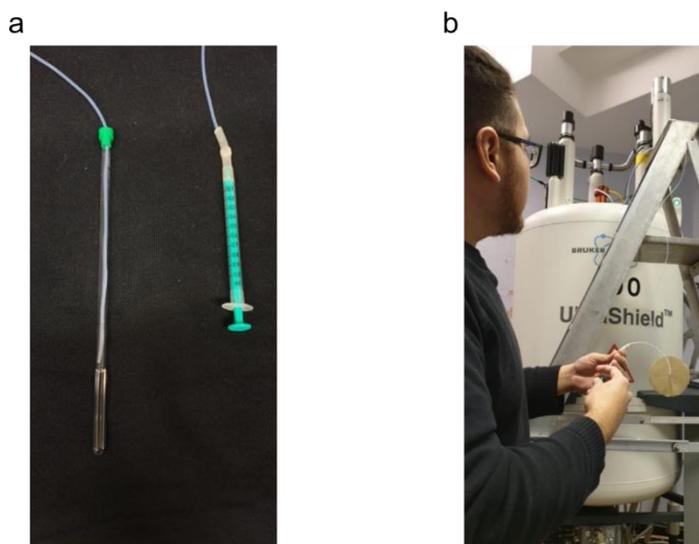
Figure 2.9 Injection device constructed for this thesis. (a) The EFTE tubing was connected at one end to a 1 mL syringe and inserted at the other end through the lid of an NMR tube. An approximately 5 mm space was left between the end of the tube and the sample, so the tube did not interfere with the probe (b) The rapid mixing device was manually loaded into the NMR spectrometer and allowed to equilibrate thermally for 5 min. The syringe and sample solutions were mixed by pump action for 1 min before initiating data acquisition.

2.5.8.6 Ligand binding

NMR spectroscopy allows for the simultaneous measurement of the ligand binding affinity and the structural changes prompted by the formation of the protein-ligand complex, through the analysis of the ligand concentration dependency of the protein chemical shifts or the peak intensities (Williamson, 2013). Due to the relative high protein concentration required, NMR spectroscopy is better suited for the determination of dissociation constants in the µM–mM range. Ligand binding was analysed assuming a one-site binding model:

$$P + L \overset{K_d}{\leftrightarrow} PL \qquad\qquad 2.10$$

Dilution effects were corrected by normalising the protein peak intensities at each ligand concentration using the TSP reference peak intensity:

$$I_{\text{corr}}([L]_T) = \frac{I([L]_T) - I(0)}{I_{\text{TSP}}([L]_T)} \qquad\qquad 2.11$$

The corrected intensities, $I_{\text{corr}}([L]_T)$, at each ligand concentration were then fitted using a non-linear least-squares fitting algorithm to:

$$I_{\text{corr}}([L]_T) = \frac{I_{\max}}{2[P]_T}\left(([P]_T + [L]_T + K_d) - \sqrt{([P]_T + [L]_T + K_d)^2 - 4[P]_T[L]_T}\right) \quad 2.12$$

where $I_{\max}$ is the maximum intensity at ligand saturation, $[P]_T$ and $[L]_T$ are the total protein and ligand concentrations, and K_d is the dissociation constant.

3 ^{1}H, ^{15}N and ^{13}C backbone resonance assignments of the P146A variant of β-phosphoglucomutase from *Lactococcus lactis* in its substrate-free form

[1]H, ^{15}N and ^{13}C backbone resonance assignments of the P146A variant of β-phosphoglucomutase from *Lactococcus lactis* in its substrate-free form

F. Aaron Cruz-Navarrete[1] · Nicola J. Baxter[1,2] · Henry P. Wood[1] · Andrea M. Hounslow[1] · Jonathan P. Waltho[1,2]

[1]   Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, The University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK

[2]   Manchester Institute of Biotechnology and School of Chemistry, The University of Manchester, 131 Princess Street, Manchester M1 7DN, UK

Corresponding Author

Jonathan P. Waltho

j.waltho@sheffield.ac.uk

ORCID 0000-0002-7402-5492

3.1 Abstract

β-Phosphoglucomutase (βPGM) is a magnesium-dependent phosphoryl transfer enzyme that catalyses the reversible isomerisation of β-glucose 1-phosphate and glucose 6-phosphate, via two phosphoryl transfer steps and a β-glucose 1,6-bisphosphate intermediate. Substrate-free βPGM is an essential component of the catalytic cycle and an understanding of its dynamics would present significant insights into βPGM functionality, and enzyme catalysed phosphoryl transfer in general. Previously, 30 residues around the active site of substrate-free βPGM$_{WT}$ were identified as undergoing extensive millisecond dynamics and were unassignable. Here we report ^{1}H, ^{15}N and ^{13}C backbone resonance assignments of the P146A variant (βPGM$_{P146A}$) in its substrate-free form, where the K145–A146 peptide bond adopts a *trans* conformation in contrast to all crystal structures of βPGM$_{WT}$, where the K145–P146 peptide bond is *cis.* In βPGM$_{P146A}$ millisecond dynamics are suppressed for all but 17 residues, allowing 92% of backbone resonances to be assigned. Secondary structure predictions using TALOS-N reflect βPGM crystal structures, and a chemical shift comparison between substrate-free βPGM$_{P146A}$ and βPGM$_{WT}$ confirms that the solution conformations are very similar, except for the D137–A147 loop. Hence, the isomerisation state of the 145–146 peptide bond has little effect on structure but the *cis* conformation triggers millisecond dynamics in the hinge (V12–T16), the nucleophile (D8) and residues that coordinate the transferring phosphate group (D8 and S114–S116), and the D137–A147 loop (V141–A142 and K145). These millisecond dynamics occur in addition to those for residues involved in coordinating the catalytic Mg^{II} ion and the L44–L53 loop responsible for substrate discrimination.

3.2 Biological context

β-Phosphoglucomutase (βPGM, EC 5.4.2.6) from *Lactococcus lactis* is a magnesium-dependent phosphoryl transfer enzyme of the haloacid dehalogenase superfamily (Lahiri et al., 2002a; Allen and Dunaway-Mariano, 2004; Dai et al., 2009). In the catabolism of maltose and trehalose, βPGM catalyses the reversible isomerisation of β-glucose 1-phosphate (βG1P) and glucose 6-phosphate (G6P). During catalysis, βG1P binds to phosphorylated βPGM (βPGM^P, phosphorylated on residue D8) forming β-glucose 1,6-bisphosphate (βG16BP), which is released to solution. Subsequent rebinding of βG16BP in the alternate orientation to non-phosphorylated, substrate-free βPGM results in dephosphorylation of βG16BP, with the formation of G6P (which enters glycolysis) and the regeneration of βPGM^P (Zhang et al., 2005; Dai et al., 2006). The βPGM gene (*pgmB*) is located on the trehalose operon and is induced by maltose or trehalose in the growth medium but is repressed by the presence of glucose or lactose (Qian et al., 1994; Qian et al., 1997). A βPGM knockout mutant strain of *L. lactis* shows impaired growth when maltose is used as the only carbon source, coupled with an intracellular accumulation of trehalose 6-phosphate and polysaccharide molecules composed of α-1,4-linked glucose units (Levander, Andersson and Radstrom, 2001). Such a perturbation of the metabolic flux highlights the crucial role that βPGM plays in mediating the efficient utilisation of carbohydrate species in *L. lactis* metabolism.

Wild-type βPGM (βPGM$_{WT}$) together with a series of variants have been studied extensively using kinetic experiments (Zhang et al., 2005; Dai et al., 2006; Dai et al., 2009; Golicnik et al., 2009), X-ray crystallography (Lahiri et al., 2002a; Lahiri et al., 2002b; Lahiri et al., 2003; Tremblay et al., 2005; Baxter et al., 2010; Griffin et al., 2012; Jin et al., 2014; Johnson et al., 2018), NMR spectroscopy (Baxter et al., 2006; Baxter et al., 2008; Baxter et al., 2009; Baxter et al., 2010; Griffin et al., 2012; Jin et al., 2014; Johnson et al., 2018) and density functional theory approaches (Webster, 2004; Marcos, Field and Crehuet, 2010; Elsasser, Dohmeier-Fischer and Fels, 2012; Barrozo et al., 2018) and it is considered as an archetypal system for enzyme catalysed phosphoryl transfer reactions. Structural analysis coupled with metal-fluoride ground state and transition state analogue (TSA) complexes have allowed the atomic resolution description of several discrete species found in the catalytic cycle i.e. substrate-free βPGM$_{WT}$ (PDB: 2WHE; Baxter et al., 2010), a ground state

βPGM$_{WT}^P$ analogue (βPGM$_{WT}$:BeF$_3$ complex; PDB: 2WFA; Griffin et al., 2012), two ground state βPGM$_{WT}^P$:G6P complexes (βPGM$_{WT}$:BeF$_3$:G6P complexes; PDB: 2WF8; PDB: 2WF9; Griffin et al., 2012), two βPGM$_{D10N}$:βG16BP complexes (PDB: 5OK1; PDB: 5OK0; Johnson et al., 2018), a βPGM$_{WT}^P$:G6P TSA complex (βPGM$_{WT}$:MgF$_3$:G6P TSA complex; PDB: 2WF5; Baxter et al., 2010) and a βPGM$_{WT}^P$:βG1P TSA complex (βPGM$_{WT}$:MgF$_3$:βG1CP TSA complex; PDB: 4C4R; Jin et al., 2014). The enzyme active site is located in the cleft formed between the α/β core domain (M1–D15 and S88–K216) and the α-helical cap domain (T16–V87). During catalysis, domain reorientation through hinge residue (D15–T16 and V87–S88) rearrangement results in closure and opening of the active site cleft facilitating substrate binding and product release. Two phosphate group binding sites are present, one in a *proximal* site adjacent to the carboxylate nucleophile (residue D8) (Lahiri et al., 2002a) and the catalytic Mg^{II} ion (Lahiri et al., 2002a), and the other in a *distal* site located ~8 Å away in the closed enzyme (Lahiri et al., 2003). The carboxylate group of the assigned general acid-base (residue D10) (Dai et al., 2009) populates two orientations depending on the degree of active site closure. In the structures of substrate-free βPGM$_{WT}$ and the βPGM$_{WT}^P$ analogue, the active site cleft is open and the D10 carboxylate group is not engaged in the active site, whereas in the closed βPGM$_{WT}^P$:G6P, βPGM$_{D10N}$:βG16BP, βPGM$_{WT}^P$:G6P TSA and βPGM$_{WT}^P$:βG1P TSA complexes, the carboxylate group is positioned to facilitate general acid-base catalysis promoting phosphoryl transfer (Johnson et al., 2018). Key roles for several residue segments in the active site have been identified including, coordination of the transferring phosphate group in the *proximal* site (V9, D10, S114, A115 and K145) (Lahiri et al., 2003), coordination of the phosphate group of the substrate in the *distal* site (R49, S116, K117 and N118) (Lahiri et al., 2003), substrate discrimination and binding (L44–L53) (Lahiri et al., 2004) and coordination of the catalytic Mg^{II} ion (D10, E169 and D170) (Lahiri et al., 2002a).

Previously, the solution behaviour of substrate-free βPGM$_{WT}$ was investigated by NMR spectroscopy and a backbone resonance assignment was determined (BMRB: 7235; Baxter et al., 2006). However, 30 residues (D8–T16, R38, L44–L53, S114–N118, V141–A142, K145 and S171–Q172) located primarily in the active site loops remained unassigned in the ^{1}H-^{15}N TROSY spectrum, most likely due to extensive conformational intermediate exchange dynamics occurring on the millisecond

timescale, which results in broadening of the correlations beyond the limits of detection. Substrate-free βPGM is an essential component of the catalytic cycle and an understanding of the dynamics of key residue segments would present significant insights into βPGM functionality. Accordingly, a series of single site variants of βPGM was screened to establish whether any improvement in spectral quality could be obtained. Of the variants tested, the P146A variant of βPGM (βPGM$_{P146A}$) reduced the intermediate exchange dynamics by the strongest extent and so was investigated further. On the basis of the conformational properties of alanine, βPGM$_{P146A}$ is expected to adopt a *trans* K145–A146 peptide bond as the dominant population. In contrast, all of the reported crystal structures described for βPGM$_{WT}$, indicate a *cis* K145–P146 peptide bond within the D137–A147 loop. Consequently, the isomerisation state of the 145–146 peptide bond presents a trigger for some of the intermediate exchange dynamics observed. Preliminary kinetics experiments using methods described previously (Johnson et al., 2018) indicate that βPGM$_{P146A}$ is active. Complete equilibration of 10 mM βG1P with G6P by 3 μM βPGM$_{P146A}$ was achieved in 1.5 h. Here, we report the ^{1}H$_N$, ^{15}N, ^{13}C$_\alpha$, ^{13}C$_\beta$ and ^{13}C' backbone resonance assignments of substrate-free βPGM$_{P146A}$, including the resonances of many residues that were previously unassigned in substrate-free βPGM$_{WT}$.

3.3 Methods and experiments

3.3.1 Protein expression and purification

Site-directed mutagenesis (QuikChange II Site-Directed Mutagenesis Kit, Agilent Technologies) of the *pgmB* from *Lactococcus lactis* cloned in the pET-22b(+) expression plasmid was employed to generate βPGM$_{P146A}$ using primers with single-site base changes. Successful mutagenesis was confirmed by DNA sequencing. The plasmid was transformed into *Escherichia coli* strain BL21(DE3) cells (Stratagene) and ^{2}H,^{15}N,^{13}C-labelled βPGM$_{P146A}$ (25 kDa) was expressed in defined isotopically labelled minimal media (Reed et al., 2003). The cells were grown at 37 °C with shaking until OD$_{600nm}$ = 0.6, at which point they were cooled to 25 °C and induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. Cells were incubated for a further 18 h and were harvested by centrifugation at 10,000 rpm for 10 min. The cell pellet was resuspended in ice-cold standard working buffer (50 mM K$^+$ HEPES pH 7.2, 5 mM MgCl$_2$, 2 mM NaN$_3$, 1 mM EDTA) supplemented with cOmplete™ protease inhibitor cocktail (Roche) (one tablet per 50 mL cell suspension).

The cell suspension was lysed on ice using 6 cycles of sonication with pulsation for 20 s followed by 60 s cooling intervals. The cell lysate was then separated by ultracentrifugation at 20,000 rpm (Beckman Coulter Avanti centrifuge using rotor JA-20) for 35 min at 4 °C. The cleared cell lysate was filtered using a 0.22 μm syringe filter (Merck Millipore) and loaded onto a DEAE-Sepharose fast flow anion-exchange column connected to an ÄKTA purification system (GE Healthcare) that had been washed previously with 1 column volume of 6 M guanidine chloride, 1 column volume of 1 M NaOH and equilibrated with 5 column volumes of standard working buffer. Proteins bound to the DEAE-Sepharose column were eluted with a gradient of 0 to 50% standard working buffer containing 1 M NaCl. Fractions were checked for the presence of βPGM_{P146A} by SDS-PAGE, pooled together and concentrated by Vivaspin (10 kDa MWCO, Sartorius). The protein sample was loaded onto a prepacked Hiload 26/60 Superdex 75 size-exclusion column connected to an ÄKTA purification system previously washed with 1 column volume of 1 M NaOH and equilibrated with 1.5 column volumes of standard working buffer containing 1 M NaCl. Fractions containing βPGM_{P146A} were checked for purity by SDS-PAGE, pooled together and buffer exchanged into standard working buffer and concentrated to ~1.6 mM by Vivaspin (10 kDa MWCO) for storage as 1 mL aliquots at -20 °C. No procedure was necessary to promote back exchange to amide protium atoms in perdeuterated βPGM_{P146A}. Protein concentrations were estimated by absorbance at 280 nm (ε_{280} = 19940 M^{-1} cm^{-1}). All reagents were of analytical grade and were purchased from Sigma-Aldrich (UK), except for the stable isotopically-labelled compounds $^{15}NH_4Cl$ (99%), ^{13}C, $^2H_7$-D-Glucose (U-$^{13}C_6$, 99%; 1,2,3,4,5,6,6-d_7 97–98%) and ^2H_2O (99.8%), which were purchased from CortecNet (France) and used as received.

3.3.2 NMR experiments

The NMR experiments were acquired using samples loaded into 5-mm NMR tubes, which contained 1.2 mM ^2H,^{15}N,^{13}C-labelled βPGM_{P146A} in standard working buffer supplemented with ^2H_2O (10% v/v) for the deuterium lock and 1 mM trimethylsilyl propanoic acid (TSP) as a chemical shift reference. All experiments were recorded at 298 K using an 800 MHz Bruker Avance I spectrometer fitted with a 5-mm TXI probe equipped with z-axis gradients and running TopSpin software version 2.1. For the backbone ^1H, ^{15}N, ^{13}C resonance assignment of substrate-free βPGM_{P146A}, 2D ^1H–^{15}N TROSY and TROSY-based 3D HNCA, HN(CO)CA, HNCACB, HN(CO)CACB,

HN(CA)CO and HNCO spectra were acquired using standard Bruker pulse sequences. ^{1}H chemical shifts were referenced relative to the internal TSP signal resonating at 0.0 ppm, whereas ^{15}N and ^{13}C chemical shifts were referenced indirectly using nuclei-specific gyromagnetic ratios.

3.3.3 Resonance assignments and data deposition

Backbone ^{1}H$_N$, ^{15}N, ^{13}C$_\alpha$, ^{13}C$_\beta$ and ^{13}C' chemical shifts were assigned for substrate-free βPGM$_{P146A}$ using standard triple resonance methodology (Gardner and Kay, 1998). The processing of spectra and peak picking were performed using FELIX (Felix NMR, Inc.). Frequency matching of the backbone assignments was achieved using a simulated annealing algorithm employed by the "asstools" assignment program (Reed et al., 2003). The backbone ^{1}H$_N$, ^{15}N, ^{13}C$_\alpha$, ^{13}C$_\beta$ and ^{13}C' chemical shifts have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) under the BMRB accession code 27920. Excluding the nine proline residues and the N-terminal methionine residue, 194 out of a possible 211 residues were assigned in the ^{1}H–^{15}N TROSY spectrum (Figure 3.1). In total, 93.4% of all backbone resonances were assigned (91.9% of ^{1}H$_N$, 91.9% of ^{15}N, 94.6% of ^{13}C$_\alpha$, 93.7% of ^{13}C$_\beta$ and 94.6% of ^{13}C' nuclei).
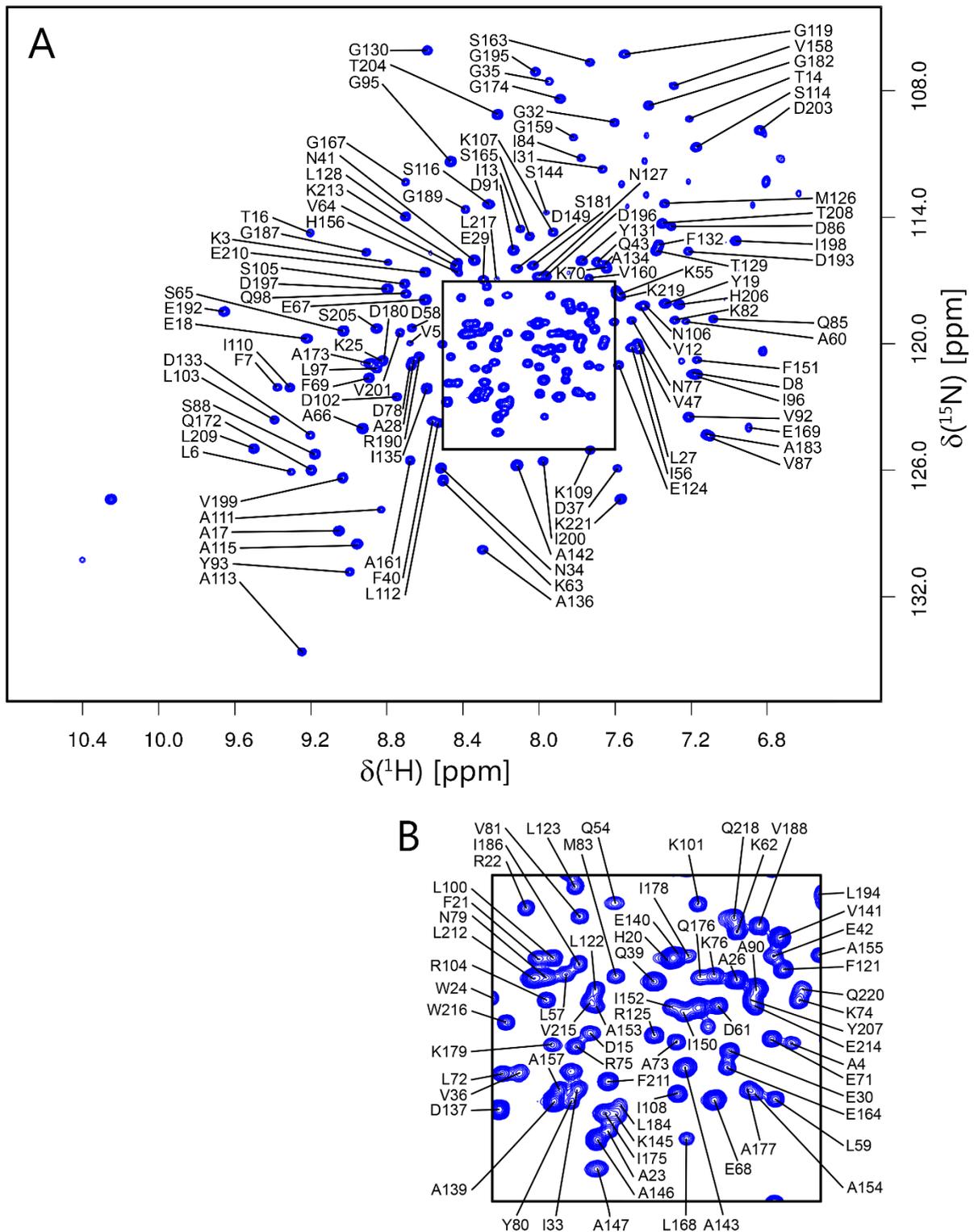
Figure 3.1 ^{1}H–^{15}N TROSY spectrum of ^{2}H,^{15}N,^{13}C-labelled substrate-free βPGM$_{P146A}$ in 50 mM K$^+$ HEPES pH 7.2, 5 mM MgCl$_2$, 2 mM NaN$_3$, 1 mM EDTA, 10% v/v ^{2}H$_2$O and 1 mM TSP recorded on an 800 MHz spectrometer at 298 K. (a) The full spectrum is shown together with (b) an expansion of the crowded region. The assignments of the backbone amide resonances are indicated by residue type and sequence number.

There are 17 residues that remain unassigned in the ^{1}H–^{15}N TROSY spectrum of substrate-free βPGM$_{P146A}$ (L9, D10, G11, R38, L44, K45, G46, S48, R49, E50, D51, S52, L53, K117, N118, D170 and S171) compared with 30 residues in substrate-free βPGM$_{WT}$ (Figure 3.2). From the crystal structures of substrate-free βPGM$_{WT}$ (PDB: 2WHE) and the βPGM$_{WT}$:MgF$_3$:G6P TSA complex (PDB: 2WF5), all of the unassigned residues in substrate-free βPGM$_{P146A}$ (except for R38) are situated within the active site and have significant roles in the catalytic cycle of the enzyme. Residues L9, D10 and G11 are key components of the *proximal* site, with D8 identified as the nucleophile, D10 assigned as the general acid-base and residues D10, D170 and S171 comprising the catalytic Mg^{II} ion binding site. Residues R49, K117 and N118 coordinate the phosphate group of the substrate in the *distal* binding site and the active site loop L44–L53 in the cap domain comprises part of a helix-loop-helix motif and is involved in substrate discrimination and binding. Due to the involvement of these residue segments in the catalytic cycle, it is likely that conformational exchange dynamics between two (or more) similarly populated forms is still occurring on the millisecond timescale in substrate-free βPGM$_{P146A}$, resulting in the attenuation of ^{1}H–^{15}N TROSY correlations beyond the limits of detection. Increased solvent exposure of the amide group of R38 through perturbation of hydrogen bonding with neighbouring sidechain groups may be coupled with helix-fraying exchange behaviour within the second α-helix (D37–E42) and the first turn of the third α-helix (S48–D58) of the cap domain, resulting in a loss of the ^{1}H–^{15}N TROSY correlation.

The secondary structure content and residue-specific random coil index order parameters (RCI-S^2) of substrate-free βPGM$_{P146A}$ were predicted by uploading the backbone ^{1}H$_N$, ^{15}N, ^{13}C$_\alpha$, ^{13}C$_\beta$ and ^{13}C' chemical shifts to the TALOS-N webserver (Shen and Bax, 2013). The predicted secondary structure for the solution conformation of substrate-free βPGM$_{P146A}$ compares well with the secondary structure present in the substrate-free βPGM$_{WT}$ crystal (PDB: 2WHE) (Figure 3.3). In addition, residues with the highest values of RCI-S^2 are located in well-defined secondary structure elements, whereas residues with the lowest values (having more random coil-like chemical shifts) correspond to loop regions in the crystal structure. Together, these data are in very good agreement, which indicates that the solution conformation is similar to the protein structure observed in the substrate-free βPGM$_{WT}$ crystal and provides confidence in the assignments of substrate-free βPGM$_{P146A}$.
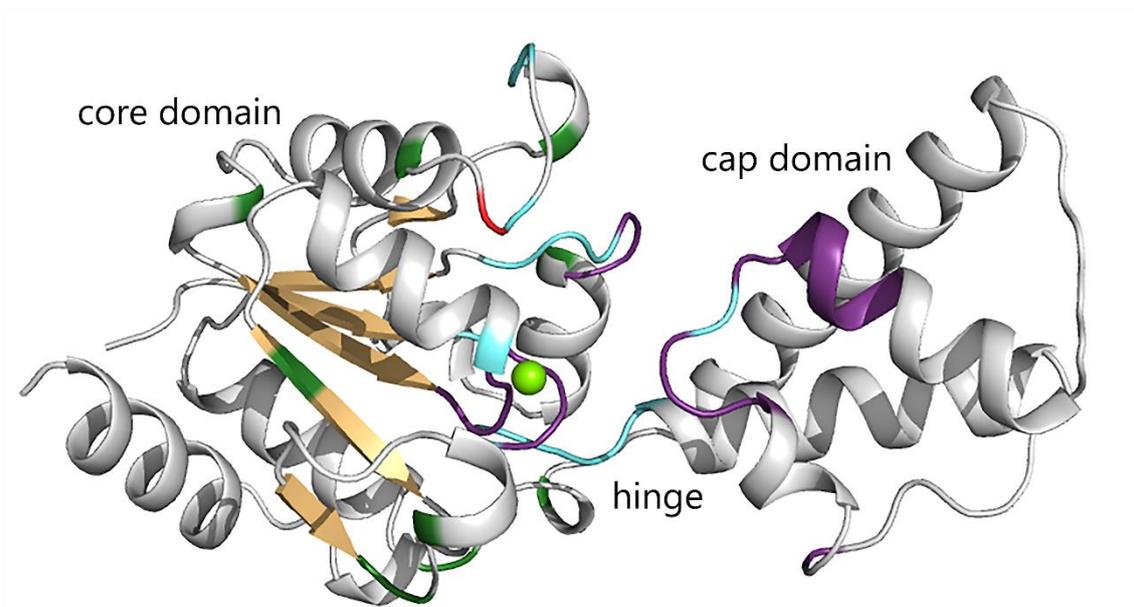
Figure 3.2 Cartoon representation of the substrate-free βPGM$_{WT}$ crystal structure (PDB: 2WHE) highlighting the extent of backbone amide resonance assignments for substrate-free βPGM$_{P146A}$ and substrate-free βPGM$_{WT}$. Assigned residues for βPGM$_{P146A}$ are coloured white (for loops and α-helices) and tan (for β-strands), with proline residues coloured green and residues that were unassigned in βPGM$_{WT}$ coloured cyan. Unassigned residues for βPGM$_{P146A}$ (L9, D10, G11, R38, L44, K45, G46, S48, R49, E50, D51, S52, L53, K117, N118, D170 and S171) are coloured purple and the 30 unassigned residues for βPGM$_{WT}$ are listed here for comparison (D8, L9, D10, G11, V12, I13, T14, D15, T16, R38, L44, K45, G46, V47, S48, R49, E50, D51, S52, L53, S114, A115, S116, K117, N118, V141, A142, K145, S171 and Q172). The location of the P146A mutation site is highlighted with a red backbone and the catalytic magnesium ion is shown as a green sphere.

A chemical shift comparison between substrate-free βPGM$_{P146A}$ and substrate-free βPGM$_{WT}$ (BMRB: 7235; Baxter et al., 2006) also supports the conclusion that the solution conformations of the two proteins are very similar. Negligible Δδ values are observed for all residues of the cap domain and small Δδ values are noted for the majority of residues present in the core domain (Figure 3.4). However, some larger Δδ values (0.08 < Δδ < 1.65 ppm) are observed primarily for two contiguous residue segments (D133–A153 and G174–S181), which are located within ~8 Å of the P146A mutation site. For a conservative single site amino acid substitution in a protein, chemical shift perturbations caused by changes in the local chemical environment are usually restricted to the immediate vicinity (within ~4 Å) of the mutation site (Baxter et al., 2017). Here, the size and the more widespread distribution of the chemical shift changes, together with a propagation of effects through several secondary structure elements, strongly suggest that the conformation of the D137–A147 loop is different

in the two proteins. Conformational heterogeneity in this loop is observed when comparing crystal structures of substrate-free βPGM$_{WT}$ (e.g. PDB: 1ZOL versus 2WHE; Zhang et al., 2005; Baxter et al., 2010). The source of this difference in structure is most probably associated with the isomerisation state of the K145–X146 peptide bond. For βPGM$_{WT}$, all the reported crystal structures have a *cis* K145–P146 peptide bond and this isomer is also populated in solution as P146 $\delta^{13}C_\beta$ = 34.8 ppm (Shen and Bax, 2010). For βPGM$_{P146A}$, a regular *trans* K145–A146 peptide bond is adopted according to TALOS-N, as expected on the basis of the conformational properties of alanine.
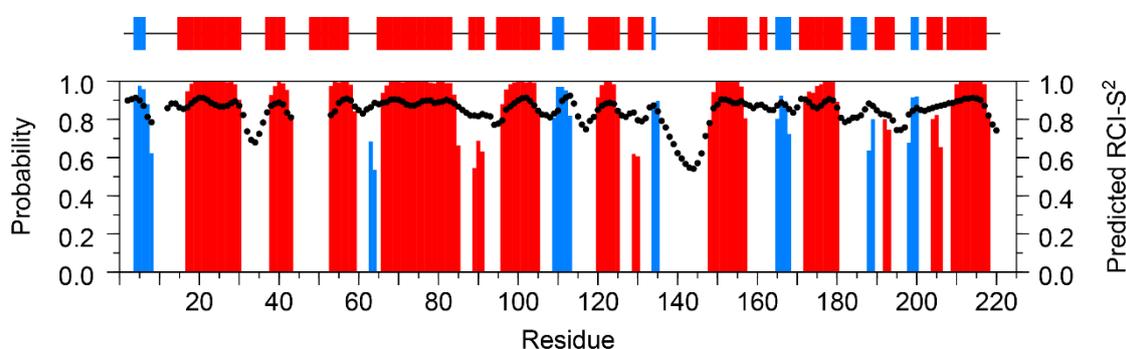
Figure 3.3 Predicted secondary structure content and residue-specific random coil index order parameters (RCI-S^2) of substrate-free βPGM$_{P146A}$ obtained with the TALOS-N webserver using the backbone ^1H_N, ^{15}N, $^{13}C_\alpha$, $^{13}C_\beta$ and $^{13}C'$ chemical shifts. The secondary structure prediction is shown as red bars for α-helices and blue bars for β-strands, with the height of the bars representing the probability assigned by the software. As a comparison, the secondary structure observed in the substrate-free βPGM$_{WT}$ crystal (PDB: 2WHE) is shown at the top of the figure in the same colour representation. The predicted RCI-S^2 values are shown as black circles.

In conclusion, the isomerisation state of the K145–X146 peptide bond appears to lead to a difference in the active site dynamics, where the *cis* conformation triggers millisecond exchange for residues of the hinge (V12–T16), the nucleophile (D8) and residues responsible for coordinating the transferring phosphate group in the *proximal* site (D8 and S114–S116), and residues of the D137–A147 loop (V141–A142 and K145). These occur in addition to millisecond dynamics for residues involved in the coordination of the catalytic Mg^{II} ion and for the L44–L53 loop responsible for substrate discrimination and binding.
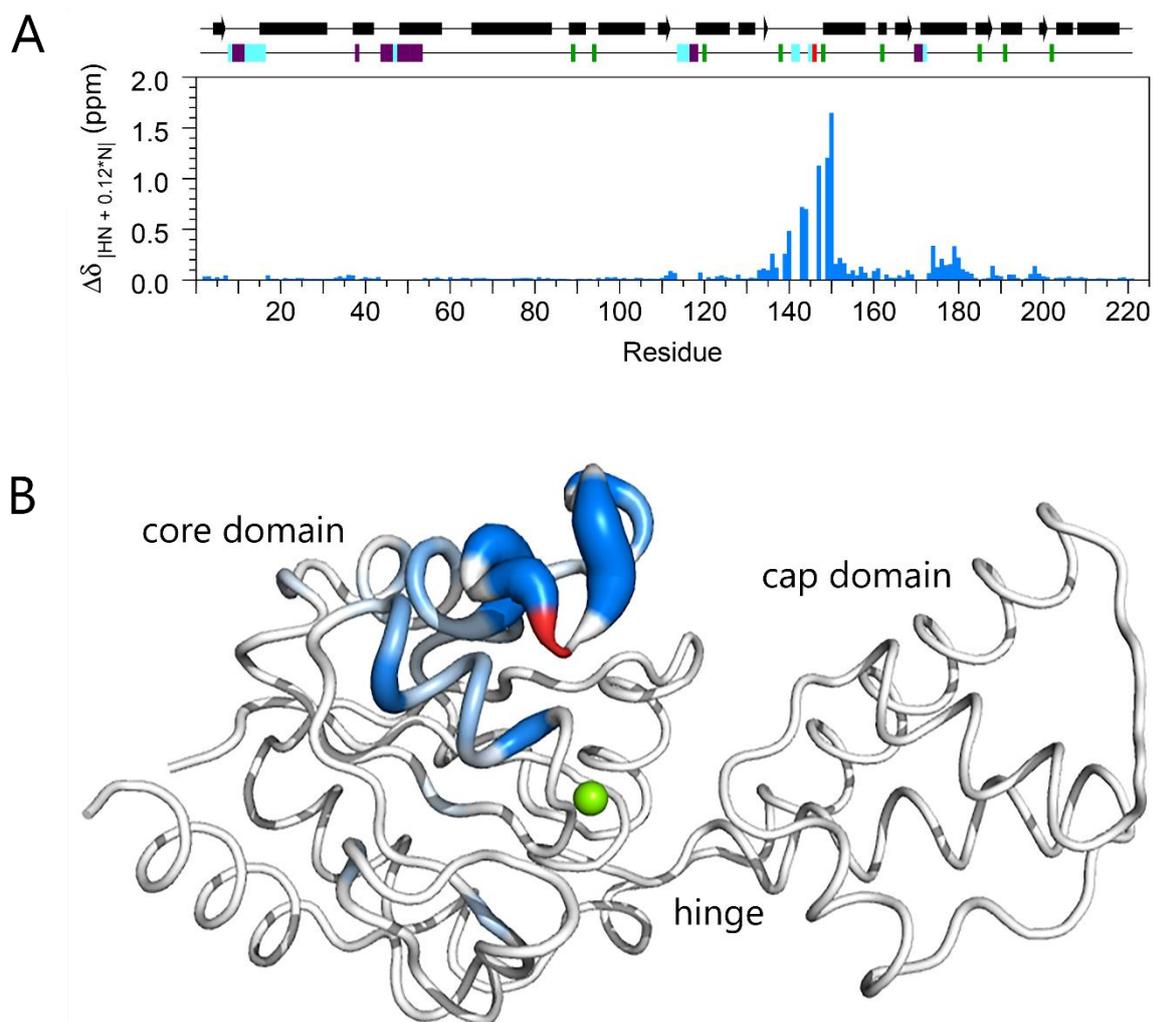
Figure 3.4   A chemical shift comparison between substrate-free βPGM$_{P146A}$ and substrate-free βPGM$_{WT}$. (a) Histogram of residue-specific chemical shift changes calculated between βPGM$_{P146A}$ and βPGM$_{WT}$ (BMRB: 7235; Baxter et al., 2006) as $\Delta\delta_{|HN + 0.12*N|} = [\Delta\delta_{HN}^2 + (0.12 \times \Delta\delta_N)^2]^{1/2}$, where $\Delta\delta_X = \delta_{X:\beta PGM:P146A} - \delta_{X:\beta PGM:WT}$ and $X = {}^1H_N$ or ${}^{15}N$ nuclei of the backbone amide group. Secondary structure elements from substrate-free βPGM$_{WT}$ (PDB: 2WHE) are indicated as bars for α-helices and arrows for β-strands at the top of the panel. Proline residues, the location of the P146A mutation site and unassigned residues in the ${}^1H$–${}^{15}N$ TROSY spectra of βPGM$_{P146A}$ and βPGM$_{WT}$ are shown as green, red, purple and cyan rectangles, respectively at the top of the panel. (b) Structure of substrate-free βPGM$_{WT}$ (PDB: 2WHE) with residues coloured according to $\Delta\delta_{|HN + 0.12*N|}$ ($\Delta\delta_{|HN + 0.12*N|} > 0.08$ ppm) with the intensity of colour and thickness of the backbone corresponding to larger values. The location of the P146A mutation site is highlighted with a red backbone and the catalytic magnesium ion is shown as a green sphere

3.4 Acknowledgements

3.5 Conflict of interest

The authors declare that they have no conflict of interest.

4 Allomorphy as a mechanism of post-translational control of enzyme activity

Allomorphy as a mechanism of post-translational control of enzyme activity

Henry P. Wood[a,1], F. Aaron Cruz-Navarrete[a,1,*], Nicola J. Baxter[a,b,1], Clare R. Trevitt[a], Angus J. Robertson[a,c], Samuel R. Dix[a], Andrea M. Hounslow[a], Matthew J. Cliff[b], and Jonathan P. Waltho[a,b,*]

[a]Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield, S10 2TN, United Kingdom; [b]Manchester Institute of Biotechnology and School of Chemistry, The University of Manchester, Manchester, M1 7DN, United Kingdom; [c]Present address: Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

[1]H.P.W., F.A.C.N. and N.J.B contributed equally to this work

[*]To whom correspondence may be addressed: Prof. Jonathan Waltho and F. Aaron Cruz-Navarrete, Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, S10 2TN, United Kingdom, +44 114 22717, j.waltho@sheffield.ac.uk, ORCID 0000-0002-7402-5492, facruznavarrete1@sheffield.ac.uk, ORCID 0000-0002-5233-581X

4.1 Abstract

Enzyme regulation is vital for metabolic adaptability in living systems. Fine control of enzyme activity is often delivered through post-translational mechanisms, such as allostery or allokairy. β-phosphoglucomutase (βPGM) from *Lactococcus lactis* is a phosphoryl transfer enzyme required for complete catabolism of trehalose and maltose, through the isomerisation of β-glucose 1-phosphate to glucose 6-phosphate via β-glucose 1,6-bisphosphate. Surprisingly for a gatekeeper of glycolysis, no fine control mechanism of βPGM has yet been reported. Herein, we describe allomorphy, a post-translational control mechanism of enzyme activity. In βPGM, isomerisation of the K145–P146 peptide bond results in the population of two conformers that have different activities owing to repositioning of the K145 sidechain. *In vivo* phosphorylating agents, such as fructose 1,6-bisphosphate, generate phosphorylated forms of both conformers, leading to a lag phase in activity until the more active phosphorylated conformer dominates. In contrast, the reaction intermediate β-glucose 1,6-bisphosphate, whose concentration depends on the β-glucose 1-phosphate concentration, couples the conformational switch and the phosphorylation step, resulting in the rapid generation of the more active phosphorylated conformer. In enabling different behaviours for different allomorphic activators, allomorphy allows an organism to maximise its responsiveness to environmental changes while minimising the diversion of valuable metabolites.

4.2 Introduction

Enzyme regulation is vital in maintaining the balance of catabolism and anabolism in living systems (Jacob and Monod, 1961; Monod, Changeux and Jacob, 1963; Wegner et al., 2015). Enzyme activity is subject to precise control, sometimes involving manifold layers of regulation, and failure often results in metabolic disorders and disease (Pardee, 2006; Iommarini et al., 2017). Regulatory mechanisms are divided into two broad categories: those relating to the control of enzyme concentration (coarse control) and those that modulate enzyme activity (fine control). In coarse control, concentration is determined by transcriptional modulation of gene expression and the balance between the rates of translation and degradation, with additional contributions from maturation, cellular compartmentalisation and local co-clustering (Schimke and Doyle, 1970; Welch, 1977; de Graffenried and Bertozzi, 2004; Castellana et al., 2014). Coarse control occurs on relatively long timescales (hours to days). In fine control, a diverse group of regulatory mechanisms act to modulate enzyme activity over much shorter timescales (< second to minutes). This group includes the binding of regulatory molecules and reversible covalent modification (Krebs and Beavo, 1979; Zhao et al., 2010), and often involves allosteric modulation, where an effector, acting somewhere other than the active site, stabilises forms of the enzyme with a reduced or enhanced activity (Monod, Changeux and Jacob, 1963; Monod, Wyman and Changeux, 1965; Koshland, Nemethy and Filmer, 1966; Changeux, 2013). Alternatively, allokairy is a fine control mechanism, where the activity of a monomeric enzyme is modulated by the near-equivalence of the conformational exchange rate and the catalytic rate in a substrate concentration-dependent manner (Whittington et al., 2015; Hilser, Anderson and Motlagh, 2015).

Precise enzyme regulation allows organisms to be responsive to environmental changes and to exploit multiple energy sources. *Lactococcus lactis* (*L. lactis*) is a Gram-positive bacterium that has worldwide usage in the manufacture of fermented dairy products and in the commercial production of lactic acid (Neves et al., 2005). It can grow on a variety of carbohydrate media including trehalose and maltose (Qian et al., 1994; Qian et al., 1997; Andersson et al., 2005). Trehalose is transported into *L. lactis* by the phosphoenolpyruvate-dependent phosphotransferase system, yielding trehalose 6-phosphate (T6P), which is phosphorolysed by P_i-dependent trehalose 6-phosphate phosphorylase to β-glucose 1-phosphate (βG1P) and glucose 6-phosphate

(G6P) (Andersson, Levander and Radstrom, 2001) (Figure 5.1). In contrast, maltose enters cells by the ATP-dependent permease system and is phosphorolysed by the action of P_i-dependent maltose phosphorylase to βG1P and glucose (Sjoberg and Hahnhagerdal, 1989). Glucose is subsequently phosphorylated to G6P by glucokinase and enters glycolysis via fructose 1,6-bisphosphate (F16BP). For complete catabolism of both trehalose and maltose, the isomerisation of βG1P to G6P is catalysed by β-phosphoglucomutase (βPGM, EC 5.4.2.6, 25 kDa). βPGM deficient *L. lactis* is unable to grow or has impaired growth, when the sole carbon source is trehalose or maltose, respectively (Levander, Andersson and Radstrom, 2001). With maltose, βG1P accumulates intracellularly and is excreted into the growth medium. Correspondingly, both P_i-dependent trehalose 6-phosphate phosphorylase and P_i-dependent maltose phosphorylase (Figure 5.1) operate in the reverse sense to their physiological roles in wild-type *L. lactis*, resulting in βG1P being combined with G6P to form T6P or polymerised to form amylose (α(1→4)-linked glucose units). In trehalose and maltose metabolism, therefore, βPGM acts as the gatekeeper to and from glycolysis, and is expected to be subject to tight regulation. In terms of coarse control, transcription of the βPGM gene (*pgmB*), which is located in the *tre* operon, is subject to negative transcriptional control by glucose and lactose (Qian et al., 1997). When *L. lactis* switches from metabolising glucose to metabolising maltose (or by implication, trehalose), there is a significant rise in the specific activity of βPGM over a period of several hours. However, no fine control mechanism has yet been identified at basal levels of βPGM, which would allow the cell to compete more successfully during a transition between carbohydrate sources.

βPGM is a monomeric magnesium-dependent phosphoryl transfer enzyme of the haloacid dehalogenase (HAD) superfamily (Lahiri et al., 2002b; Zhang et al., 2005; Baxter et al., 2006; Dai et al., 2009; Golicnik et al., 2009; Baxter et al., 2010; Jin et al., 2014; Johnson et al., 2018). The active site is located in the cleft between the α/β core domain (M1–D15, S88–K216) and the α-helical cap domain (T16–V87), with closure of the cleft through domain reorientation occurring during catalysis. Two phosphate group binding sites are present, a proximal site adjacent to the carboxylate nucleophile and the catalytic Mg^{2+} ion, and a distal site located ~8 Å away in the closed enzyme (Baxter et al., 2010). During steady-state catalysis, βG1P binds to phosphorylated βPGM (βPGM^P, phosphorylated on D8) and forms β-glucose 1,6-bisphosphate

(βG16BP). Release to solution and subsequent rebinding of βG16BP in the alternate orientation (Dai et al., 2006) leads to the formation of G6P and the regeneration of βPGM^P (Figure 4.1). In vitro, a phosphorylating (priming) agent is required to initiate the catalytic cycle since the half-life of βPGM^P is ~30 s (Golicnik et al., 2009). *In vivo*, potential candidates for this agent include F16BP, βG1P, G6P, α-glucose 1,6-bisphosphate (αG16BP), and acetyl phosphate (AcP), as well as the reaction intermediate, βG16BP. However, only βG16BP allows βPGM to reach its maximum catalytic rate, and a significant lag phase is observed in the reaction with αG16BP as the phosphorylating agent (Golicnik et al., 2009), until the βG16BP concentration greatly exceeds its resting concentration in the cell. In the current kinetic model for βPGM catalysis, αG16BP is also required to act as a very strong inhibitor of βPGM. αG16BP is a close structural analogue of βG16BP (Zhang et al., 2005), but very similar kinetic behaviour is observed when AcP is used as the phosphorylating agent (Johnson et al., 2018), suggesting that other factors are contributing to post-translational control of βPGM. Here we show, through combined use of NMR spectroscopy, X-ray crystallography, site-directed mutagenesis and kinetic assays, that a regulatory mechanism is operating in βPGM, which we have termed allomorphy to illustrate its relationship to and distinction from allostery and allokairy. In the substrate-free enzyme, the isomerisation of proline 146 results in the near-equal population of two conformers that have different activities. Alternative phosphorylating agents such as F16BP and AcP generate phosphorylated forms of both conformers, resulting in a lag phase in βPGM activity until the more active phosphorylated conformer dominates. In contrast, the βG16BP reaction intermediate is able to couple the conformational switch and the phosphorylation step, resulting in the rapid generation of the more active phosphorylated species. This allows the βG16BP concentration to effectively act as a surrogate of the βG1P concentration and modulate the activity of βPGM according to the carbohydrate source available to *L. lactis*.
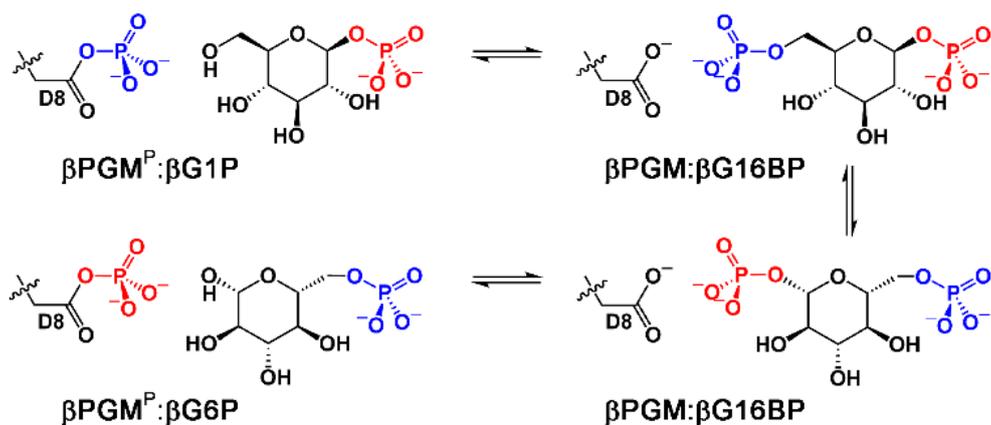
Figure 4.1 βPGM catalytic cycle. βPGM reaction scheme for the enzymatic conversion of βG1P to G6P via a βG16BP intermediate. The phosphoryl transfer reaction between phospho-enzyme (βPGM^P, phosphorylated at residue D8) and βG1P is illustrated with the transferring phosphate (blue) in the proximal site and the 1-phosphate (red) of βG1P in the distal site. The equivalent reaction between βPGM^P and G6P is shown with the transferring phosphate (red) in the proximal site and the 6-phosphate (blue) of G6P in the distal site.

73

4.3 Methods

4.3.1 β-Phosphoglucomutase (βPGM) expression and purification

Wild type βPGM (βPGM$_{WT}$) and the P146A variant (βPGM$_{P146A}$) proteins were expressed using either ^{15}N or ^{2}H^{15}N^{13}C isotopic enrichment (Reed et al., 2003) and purified using methodology that minimised the presence of contaminating phosphoryl transfer enzymes (e.g. phosphoglucose isomerase and βPGM from *E. coli*) (Johnson et al., 2018; Cruz-Navarrete et al., 2019). The βPGM$_{WT}$ and βPGM$_{P146A}$ genes were cloned into the pET-22b(+) plasmid, transformed into *E. coli* strain BL21(DE3) cells and expressed in defined isotopically labelled minimal media. Cells were grown at 37 °C with shaking until OD$_{600nm}$ = 0.6, then cooled at 25 °C and induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for a further 18 h. Cells were harvested by centrifugation at 10,000 rpm for 10 min. The cell pellet was resuspended in ice-cold lysis buffer (50 mM K$^+$ HEPES (pH 7.2), 2 mM NaN$_3$, 1 mM EDTA) supplemented with cOmplete™ protease inhibitor cocktail and lysed by 6 cycles of sonication. The cell lysate was cleared by centrifugation at 20,000 rpm for 35 min at 4 °C. The supernatant was filtered using a 0.22 µm syringe filter and loaded onto a DEAE-Sepharose fast flow anion-exchange column. Proteins bound to the column were eluted with a gradient of 0 to 50% lysis buffer containing 1 M NaCl. Fractions containing βPGM were purified further using a Hiload 26/60 Superdex 75 size-exclusion column previously equilibrated with lysis buffer containing 1 M NaCl. Final fractions were pooled, buffer exchanged into 50 mM K$^+$ HEPES buffer (pH 7.2) containing 2 mM NaN$_3$ and concentrated to ~1.6 mM for storage at -20 °C.

4.3.2 Reagents

Unless otherwise stated, reagents were purchased from Sigma-Aldrich, GE Healthcare, Melford Laboratories or CortecNet. βG1P was synthesised enzymatically from maltose using maltose phosphorylase (EC 2.4.1.8) (Johnson et al., 2018). A solution of maltose (600 mM) was incubated overnight with 1.2 units mL^{-1} maltose phosphorylase in 0.5 M sodium phosphate buffer (pH 7.0) at 30 °C and βG1P production was confirmed using ^{31}P NMR spectroscopy. Maltose phosphorylase was removed using a Vivaspin (5 kDa MWCO) and the resulting flow-through was used without further purification. βG16BP was produced enzymatically from βG1P and AcP using the D170N variant of βPGM (βPGM$_{D170N}$; expressed and purified as detailed

above) (Wood et al., 2021). βG1P and AcP were incubated with βPGM$_{D170N}$ for 4 h at 25 °C and the reaction was quenched by heating at 90 °C for 10 min. βG16BP was purified using barium salt precipitation.

4.3.3 NMR spectroscopy

^{1}H^{15}N-TROSY NMR spectra of βPGM$_{WT}$ and βPGM$_{P146A}$ were acquired at 298 K using 0.5–1 mM ^{15}N-βPGM in standard NMR buffer (50 mM K$^+$ HEPES (pH 7.2), 5 mM MgCl$_2$, 2 mM NaN$_3$ with 10% (v/v) ^{2}H$_2$O and 1 mM TSP). Typically, ^{1}H^{15}N-TROSY spectra were accumulations of 16 transients, with 256 increments and spectral widths of 32–36 ppm centred at 120 ppm in the indirect ^{15}N dimension. ^{1}H^{15}N-TROSY-based ZZ-exchange experiments were performed at mixing times of 100, 300, 500 and 900 ms. Rapid acquisition ^{1}H^{15}N BEST-TROSY experiments to monitor the steady-state behaviour of ^{15}N-βPGM$_{WT}$ (0.2 mM) and ^{15}N-βPGM$_{P146A}$ (0.2 mM) were acquired in standard kinetic buffer (200 mM K$^+$ HEPES (pH 7.2), 5 mM MgCl$_2$, 2 mM NaN$_3$ with 10% (v/v) ^{2}H$_2$O and 1 mM TSP) containing either 50–100 mM F16BP, 60–100 mM AcP or 35 mM βG16BP. The ^{1}H^{15}N BEST-TROSY spectra were recorded at 298 K using a Bruker 600 MHz Neo spectrometer equipped with a 5-mm TCI cryoprobe and z-axis gradients as 11 min experiments (16 transients, 128 increments and a recycle delay of 0.2 s) with selective ^{1}H pulses centred on the amide region (8.3 ppm). Excitation pulses (90°) were 1.7 ms (pulse shape Eburp2), whereas refocusing pulses (180°) were 1.4 ms (pulse shape Reburp). The experimental dead-time was ~5 min.

For βPGM$_{P146A}$ prepared in standard kinetic buffer containing 50 mM F16BP, B^P dephosphorylation was monitored at 298 K by consecutive one dimensional ^{1}H NMR experiments recorded with 16 transients, a 1 s recycle delay and a spectral width of 32 ppm centred on the water signal. Following 0.3 Hz Lorentzian apodisation and baseline correction, normalised integral values of the F16BP peak (4.22-4.15 ppm) were plotted against time to give a kinetic profile. The initial linear steady-state portion of the kinetic profile was fitted using a linear least-squares fitting algorithm included in MATLAB 2018a to derive an apparent dephosphorylation rate constant.

To observe the species present immediately following the addition of phosphorylating agent to βPGM$_{WT}$, NMR experiments were recorded with the use of a homemade rapid mixing device. The equipment comprised a 2 m length of 0.8 mm internal diameter EFTE tubing (GE Healthcare), connected at one end to a 1 mL syringe and inserted

at the other end through the lid of an NMR tube. The tubing was loaded with phosphorylating agent (550 mL 100 mM F16BP or 250 mL 320 mM AcP, prepared in standard kinetic buffer) and a small, additional volume of air was drawn in to prevent premature mixing of the phosphorylating agent with the 550 mL 1.2 mM ^{15}N-βPGM$_{WT}$ sample prepared in standard kinetic buffer. The rapid mixing device was loaded into a Bruker 600 MHz Neo spectrometer and allowed to equilibrate thermally at 298 K. Following mixing by syringe action of the phosphorylating agent (final concentration: 50 mM F16BP or 100 mM AcP) with the βPGM$_{WT}$ sample, the spectrometer was locked (with ~7% (v/v) ^2H_2O), tuned and shimmed, and the ^1H 90° pulse length was calibrated. A series of $^1H^{15}N$ BEST-TROSY spectra were recorded as 142 s experiments (4 transients, 128 increments and a recycle delay of 0.15 s). The experimental dead-time was ~156 s.

Multi-dimensional heteronuclear NMR spectra for ^1H, ^{15}N and ^{13}C backbone resonance assignment of $^2H^{15}N^{13}C$-βPGM$_{WT}$ in standard NMR buffer containing 10 mM tris were acquired at 298 K on a Bruker 800 MHz Avance III spectrometer equipped with a 5-mm TCI cryoprobe and z-axis gradients. The standard Bruker suite of $^1H^{15}N$-TROSY and 3D TROSY-based constant time experiments were acquired (HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, HN(CO)CACB) using non-uniform sampling (NUS) with a multi-dimensional Poisson Gap scheduling strategy with exponential weighting (Hyberts, Robson and Wagner, 2013). NUS data were reconstructed using multi-dimensional decomposition in TopSpin3 (Hyberts et al., 2012). Backbone resonance assignments for conformer A and conformer B present simultaneously in the spectra were obtained using a simulated annealing algorithm employed by the asstools assignment program (Reed et al., 2003) and assignments were confirmed using sequential backbone amide to amide correlations obtained from TROSY-based (H)N(COCA)NNH and H(NCOCA)NNH experiments (Sun et al., 2005). Multi-dimensional heteronuclear NMR spectra for ^1H, ^{15}N and ^{13}C backbone resonance assignment of the $^2H^{15}N^{13}C$-βPGM$_{P146A}$:MgF$_3$:G6P TSA complex in standard NMR buffer containing 15 mM NaF and 10 mM G6P were acquired at 298 K on a Bruker 800 MHz Avance I spectrometer equipped with a 5-mm TXI probe and z-axis gradients. $^1H^{15}N$-TROSY and 3D TROSY-based constant time experiments were acquired (HN(CA)CO and HNCACB) and backbone resonance assignments were obtained using asstools (Reed et al., 2003). Multi-dimensional heteronuclear NMR

spectra for ^{1}H, ^{15}N and ^{13}C backbone resonance assignment of phosphorylated ^{2}H^{15}N^{13}C-βPGM$_{P146A}$ in standard kinetic buffer containing 75-100 mM F16BP were acquired at 298 K on a Bruker 800 MHz Neo spectrometer equipped with a 5-mm TXI probe and z-axis gradients. ^{1}H^{15}N-TROSY and 3D TROSY-based constant time experiments were acquired (HNCO and HNCA) using non-uniform sampling (NUS) with a multi-dimensional Poisson Gap scheduling strategy with exponential weighting (Hyberts, Robson and Wagner, 2013). NUS data were reconstructed using multi-dimensional decomposition in TopSpin4 (Hyberts et al., 2012). TROSY resonances were assigned by comparing the correlated ^{13}C chemical shifts with those of βPGM$_{P146A}$ (BMRB 27920; Cruz-Navarrete et al., 2019) and the βPGM$_{WT}$:BeF$_3$ complex (BMRB 17851; Griffin et al., 2012). Experiments were processed using TopSpin (Bruker) or FELIX (Felix NMR, Inc.) and NMR figures were prepared using FELIX. ^{1}H chemical shifts were referenced relative to the internal TSP signal resonating at 0.0 ppm, and ^{13}C and ^{15}N and chemical shifts were referenced indirectly using nuclei-specific gyromagnetic ratios. Differences in chemical shift were calculated as: $\Delta\delta = [(\delta_{HN-X} - \delta_{HN-Y})^2 + (0.12 \times (\delta_{N-X} - \delta_{N-Y}))^2]^{1/2}$, where X and Y are the two species being compared.

Reaction kinetics for βPGM$_{P146A}$-catalysed reactions were followed using a Bruker 500 MHz Avance DRX spectrometer (operating at 202.456 MHz for ^{31}P) equipped with a room-temperature broadband probe. The equilibration of 10 mM βG1P with G6P by 1–3 µM βPGM$_{P146A}$ was measured in standard kinetic buffer at 298 K. The reaction was initiated by and timed from the addition of 20 mM AcP and monitored by the acquisition of consecutive ^{31}P spectra without proton-phosphorus decoupling with 256 transients and a 1 s recycle delay. A spectral width of 50 ppm centred at -10 ppm enabled the observation of the relevant phosphorus signals. Normalised integral values of both the βG1P and G6P peaks following baseline correction and 5 Hz Lorentzian apodisation were plotted against time to give kinetic profiles. The linear steady-state portion of the G6P integral data was fitted using a linear least-squares fitting algorithm to derive the observed rate constant, k_{obs}.

One dimensional ^{19}F NMR experiments were acquired at 298 K on a Bruker 500 MHz Avance III spectrometer (operating at 470.536 MHz for ^{19}F) equipped with a 5-mm QCI-F cryoprobe and z-axis gradients. Samples were prepared using 0.5 mM ^{15}N-βPGM$_{WT}$ or ^{15}N-βPGM$_{P146A}$ in standard NMR buffer (without 10% (v/v) ^{2}H$_2$O)

containing 15 mM NaF and 10 mM G6P. The spectrometer lock was provided by ^2H_2O sealed inside a glass capillary tube inserted into the sample tube. Typically, ~6000 transients were acquired without ^1H decoupling over a spectral width of 120 ppm and were processed with sinebell functions shifted by 60°.

For NMR experiments conducted in filtered milk, fresh skimmed bovine milk purchased from a local supermarket was filtered using a 10 kDa MWCO Vivaspin (Sartorius) to remove fat micelles and milk proteins. A 1 mM ^{15}N-βPGM$_{WT}$ sample prepared in standard NMR buffer was diluted 5-fold with the resulting milk flow-through and 10% (v/v) ^2H_2O and 1 mM TSP were added to the sample. Experiments were acquired at 298 K on a Bruker 800 MHz Avance I spectrometer equipped with a 5-mm TXI probe and z-axis gradients. The pH of the sample was estimated as pH 6.8 using the residual ^1H resonances of HEPES buffer originating from the standard NMR buffer.

For NMR experiments conducted in the absence of salt, a 1.4 mM ^{15}N-βPGM$_{WT}$ sample in standard NMR buffer containing an additional 200 mM NaCl was buffer-exchanged into freshly produced deionised water (18.2 MΩ.cm, Purelab Classic, Elga-Veolia), with an equivalent dilution of the previous buffer by a factor of 18.7×10^6. The resulting sample contained 0.9 mM βPGM$_{WT}$ at pH 6.3 and was supplemented with 10% (v/v) ^2H_2O for the deuterium lock. Experiments were acquired at 298 K on a Bruker 800 MHz Avance I spectrometer equipped with a 5-mm TXI probe and z-axis gradients.

4.3.4 Kinetic experiments using coupled assays

All kinetic assays for βPGM$_{WT}$ and βPGM$_{P146A}$ were conducted at 298 K using a FLUOstar OMEGA microplate reader and the BMG LABTECH Reader Control Software (version 5.11) (BMG Labtech) in standard kinetic buffer (200 mM K$^+$ HEPES (pH 7.2), 5 mM MgCl$_2$, 1 mM NaN$_3$) in a 160 μL reaction volume. The rate of G6P production was measured indirectly using a glucose 6-phosphate dehydrogenase (G6PDH) coupled assay, in which G6P is oxidised and concomitant NAD$^+$ reduction is monitored by the increase in absorbance at 340 nm (NADH extinction coefficient = 6220 M^{-1} cm^{-1}). βPGM$_{WT}$ and βPGM$_{P146A}$ concentrations were determined using a NanoDrop OneC spectrophotometer (Thermo Scientific) and diluted accordingly (βPGM extinction coefficient = 19 940 M^{-1} cm^{-1}). For the determination of k_{cat} and K_m values for βPGM$_{WT}$, the reaction was initiated by dilution of the enzyme prepared in

standard kinetic buffer to a final concentration of 1 nM βPGM$_{WT}$ in solutions of 1 mM NAD$^+$ and 5 units mL^{-1} G6PDH and variable concentrations of βG1P (10, 20, 30, 50, 70, 100, 150, 200, 300, 500, 700 μM) and βG16BP (0.4, 1, 2, 5, 10 μM). For the determination of k_{cat} and K_m values for βPGM$_{P146A}$, the reaction was initiated by dilution of the enzyme prepared in standard kinetic buffer to a final concentration of 100 nM βPGM$_{P146A}$ in solutions of 1 mM NAD$^+$ and 5 units mL^{-1} G6PDH and variable concentrations of βG1P (5, 10, 15, 20, 30, 50, 70, 100, 200, 300, 500 μM) and βG16BP (2, 5, 10, 35, 50, 100 μM). The initial rate of G6P production was fitted using a linear least-squares fitting algorithm to determine the reaction velocity (v_0) at each βG1P and βG16BP concentration at a total enzyme concentration (E_T). Mean data from triplicate measurements were subsequently globally fitted to Equation 4.1 (Golicnik et al., 2009), which is derived for a ping–pong mechanism and adapted to account for βG1P inhibition (K_i) to calculate k_{cat} and individual K_m values ($K_{βG1P}$ and $K_{βG16BP}$), with their corresponding standard deviations, using an in-house python non-linear least-squares fitting program.

$$v_0 = \frac{k_{cat}[E_T][\beta G1P][\beta G16BP]}{[\beta G1P][\beta G16BP] + K_{\beta G1P}[\beta G16BP] + K_{\beta G16BP}[\beta G1P]\left(\frac{K_i + [\beta G1P]}{K_i}\right)} \quad 4.1$$

Kinetic experiments demonstrating the effect of different phosphorylating agents were conducted by the addition of either 5 nM βPGM$_{WT}$ or 200 nM βPGM$_{P146A}$ to solutions containing either 1 mM F16BP, 8 mM AcP or 10 μM βG16BP, together with 1 mM NAD$^+$, 5 units mL^{-1} G6PDH and 50 μM βG1P. F16BP represents an equilibrium mixture of an α-anomer (15%), a β-anomer (81%) and two open chain forms with an interconversion rate of 8 s^{-1} (MidelfortGupta and Rose, 1976).

4.3.5 X-ray crystallography

For the crystallisation experiments of βPGM$_{WT}$, βPGM$_{P146A}$, the βPGM$_{P146A}$:MgF$_3$:G6P TSA complex and the βPGM$_{WT}$:citrate complex, frozen aliquots of βPGM$_{WT}$ or βPGM$_{P146A}$ in standard native buffer (50 mM K$^+$ HEPES (pH 7.2), 5 mM MgCl$_2$, 2 mM NaN$_3$) were thawed on ice and centrifuged briefly to pellet insoluble material. For the βPGM$_{P146A}$:MgF$_3$:G6P TSA complex, 15 mM NaF and 10 mM G6P were added to the βPGM$_{P146A}$ sample, whereas for the βPGM$_{WT}$:citrate complex, 50 mM citrate was added to the βPGM$_{WT}$ sample. Solutions were adjusted to a final protein concentration

of 0.4-0.6 mM, incubated for 1 h and mixed 1:1 with precipitant (24–34% (w/v) PEG 4000, 200 mM sodium acetate and 100 mM tris-HCl (pH 7.5)). Crystals were grown at 290 K by hanging-drop vapour diffusion using a 2 ml drop suspended on a siliconised glass cover slip above a 700 mL well. Rod-shaped or large plate crystals grew typically over several days. Crystals were harvested using a mounted LithoLoop (Molecular Dimensions Ltd) and were cryo-protected in their mother liquor containing an additional 25% (v/v) ethylene glycol (and 50 mM citrate for the βPGM_{WT}:citrate crystals) prior to plunging into liquid nitrogen. Diffraction data were collected at 100 K on the MX beamlines at the Diamond Light Source (DLS), Oxfordshire, United Kingdom.

Data were processed using the xia2 pipeline (Winter, 2010) and resolution cut-offs were applied using CC-half values and Aimless (Evans and Murshudov, 2013). The crystals diffracted in the $P2_12_12_1$ spacegroup, with cell dimensions that varied depending on the degree of enzyme closure. Structures were determined by molecular replacement with MolRep (version 11) (Vagin and Teplyakov, 1997) using the highest resolution model with the most appropriate cap and core domain relationship as a search model. Model building was carried out in COOT (version 0.8.8) (Emsley et al., 2010) with ligands not included until the final rounds of refinement with REFMAC5 (Murshudov, Vagin and Dodson, 1997) so that they could be built into unbiased difference Fourier maps. The βPGM_{P146A}:MgF_3:G6P TSA complex structure was refined with anisotropic B-factors, whereas both the βPGM_{WT} structures and the substrate-free βPGM_{P146A} structure were refined isotropically. Structure validation was carried out in COOT and MolProbity (version 4.4) (Chen et al., 2010). Superpositions and crystallographic figures were prepared using PyMOL (The PyMOL Molecular Graphics System, version 1.8, Schrödinger, LLC). To confirm the isomerisation state of the K145–A146 peptide bond in the structures of substrate-free βPGM_{P146A} and the βPGM_{P146A}:MgF_3:G6P TSA complex, difference density maps (Fo – Fc) were generated using REFMAC5 with the S144–P148 segment omitted from the final structures. Omit map figures were prepared using CCP4mg (version 2.10.9) (McNicholas et al., 2011). Additional details for X-ray crystallography data collection, data processing and refinement are provided in Table 5.1.

4.4 Results

4.4.1 βPGM$_{WT}$ exchanges slowly between two stable conformations

The observation of a lag phase when using either αG16BP or AcP (Golicnik et al., 2009; Johnson et al., 2018) as the phosphorylating agent implies that the target of phosphorylation, the substrate-free enzyme, has a role in post-translational control. Hence, the solution properties of substrate-free wild-type βPGM (βPGM$_{WT}$) were investigated using NMR spectroscopy. In the previous backbone resonance assignment of βPGM$_{WT}$ (BMRB 7235; Baxter et al., 2006) performed in standard NMR buffer (50 mM K$^+$ HEPES (pH 7.2), 5 mM MgCl$_2$, 2 mM NaN$_3$, 10% (v/v) ^{2}H$_2$O and 1 mM trimethylsilyl propanoic acid (TSP)) containing 10 mM NH$_4$F, two features were apparent during the analysis: (1) peaks of 30 active site residues were missing from the spectra owing to line-broadening resulting from conformational exchange on the millisecond timescale and (2) a large number of unassigned ^{1}H$_N$, ^{15}N, ^{13}C$_α$, ^{13}C$_β$ and ^{13}C' resonances were present with a low intensity. To test whether HEPES or NH$_4$F were contributing to the millisecond conformational exchange, spectra were recorded in tris buffer (50 mM tris (pH 7.2), 5 mM MgCl$_2$, 2 mM NaN$_3$, 10% (v/v) ^{2}H$_2$O and 1 mM TSP), and mixtures of HEPES and tris buffers in order to transfer the assignment between conditions (Wong, Khirich and Loria, 2013). It was noticed that the inclusion of 5 mM tris in the standard NMR buffer increased the intensity of the unassigned resonances significantly and therefore all observable resonances in the spectra were re-assigned using standard triple resonance TROSY-based methodology (Gardner and Kay, 1998). Excluding the ten proline residues and the N-terminal methionine, the backbone resonances of 193 out of a total of 210 residues (92%) were assigned. Seventeen residues located in the vicinity of the active site remained unassigned (L9, D10, G11, R38, L44, K45, G46, S48, R49, E50, D51, S52, L53, K117, N118, D170 and S171). Notably, 102 of the assigned residues displayed pairs of resonances in the ^{1}H^{15}N-TROSY spectrum (Figure 4.2a, Figure 5.2a), consistent with the population of two βPGM$_{WT}$ conformers (70% conformer A, BMRB 28095 and 30% conformer B, BMRB 28096). A further five residues (K145, A147, D149, I150 and Q176) have assignments in conformer A, but are missing assignments in conformer B, owing to some differential millisecond conformational exchange occurring in the two species. The βPGM$_{WT}$ conformers are present in the spectra as a result of slow conformational exchange rather than as chemically distinct species, as the addition of 3 mM BeCl$_2$

and 10 mM NH$_4$F to the βPGM$_{WT}$ sample induced the population of a single βPGM$_{WT}$:BeF$_3$ complex (an analogue of phosphorylated conformer A; BMRB 17851; Griffin et al., 2012) (Figure 5.3a). The exchange between conformer A and conformer B is on the multi-second timescale, with $k_{ex} \leq 1.0$ s^{-1} from ZZ-exchange measurements. Differences in chemical shift between the two conformers (Figure 5.4a) indicate that the regions of βPGM$_{WT}$ involved in the multi-second conformational exchange process are located primarily in the core domain and comprise the D137–A147 loop, the β-strands (K109–A113 and D133–A136) at the outer edge of the β-sheet and the I152–S163 and Q172–A183 α- and 3$_{10}$-helical regions (Figure 4.3). Predicted random coil index order parameters (RCI-S^2) (Shen and Bax, 2013) show a decrease in value for conformer B in two regions (G32–R38 in the cap domain and D133–K145 in the core domain) (Figure 5.5a), which indicates increased conformational flexibility compared with conformer A.
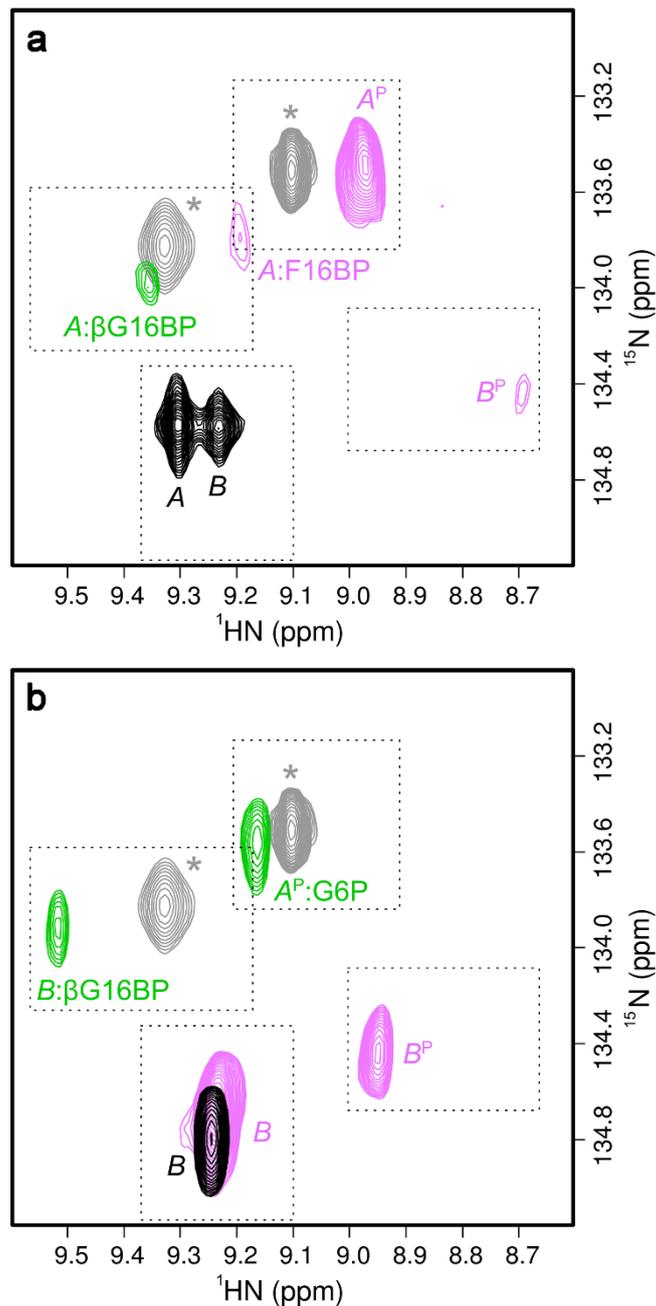
Figure 4.2 Effect of different phosphorylating agents on βPGM. a, b, Overlays of a section of $^1H^{15}N$-TROSY spectra highlighting the behaviour of residue A113. (a) $βPGM_{WT}$ (black) populates conformer A and conformer B in slow exchange. $βPGM_{WT}$ supplemented with F16BP (pink) populates phosphorylated conformer A (A^P) as the dominant species, phosphorylated conformer B (B^P) and a $βPGM_{WT}$:F16BP species (A:F16BP). $βPGM_{WT}$ supplemented with βG16BP (green) populates an A:βG16BP complex. (b) $βPGM_{P146A}$ (black) populates one conformer (conformer B). $βPGM_{P146A}$ supplemented with F16BP (pink) populates conformer B and B^P. $βPGM_{P146A}$ supplemented with βG16BP (green) populates an A^P:G6P complex and a B:βG16BP complex. Peaks indicated in grey asterisks correspond to the $βPGM_{WT}$:BeF_3 complex (grey; $δ_N$ = 133.5 ppm) (BMRB 17851), an analogue of A^P, and the Mg^{2+}-saturated $βPGM_{D10N}$:βG16BP complex (grey; $δ_N$ = 133.8 ppm) (BMRB 27174), a mimic of the A:βG16BP complex, and are shown for comparison.

4.4.2 Influence of physiological factors on the conformational exchange

An investigation of factors that could potentially affect the population distribution of conformer A and conformer B was performed using ^{1}H^{15}N-TROSY spectra of βPGM$_{WT}$ recorded under different conditions of temperature, pH, hydrostatic pressure, MgCl$_2$ (0–100 mM), NaCl (0–200 mM), K$^+$ HEPES buffer (0–200 mM) and βPGM$_{WT}$ concentration (0.1–1.2 mM). All of these perturbations had little or no effect, apart from the addition of either MgCl$_2$ (100 mM) or NaCl (200 mM) to standard NMR buffer, which shifted the population of βPGM$_{WT}$ primarily to conformer A (Figure 5.6a-d). Buffer exchange into deionised water resulted in conformer B being the dominant population. However, both conformer A and conformer B remained populated when Mg^{2+} was removed from the NMR buffer solution, showing that the multi-second conformational exchange is not simply a result of incomplete saturation of the catalytic Mg^{2+} binding site. These observations indicate that chloride anions perturb the population distribution.
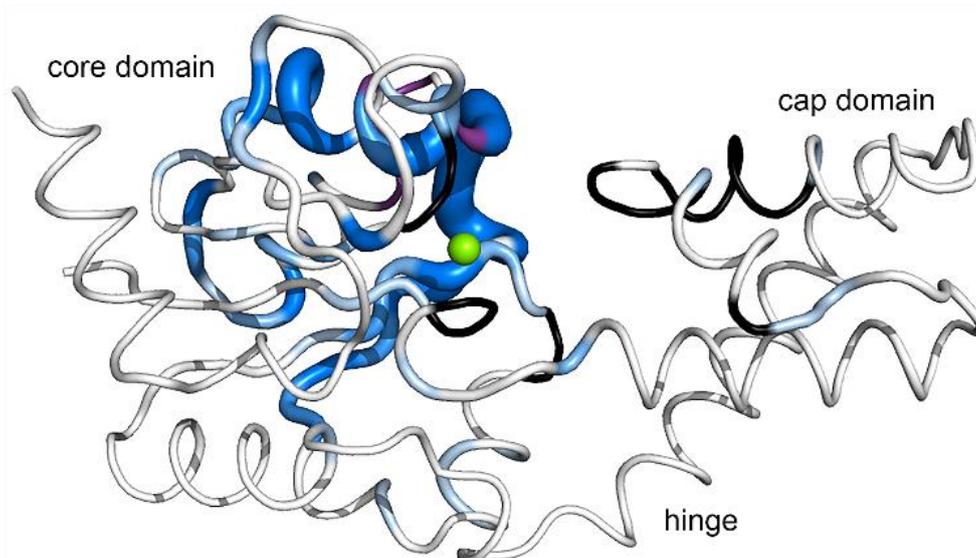
Figure 4.3 Exchange behaviour in βPGM$_{WT}$. Crystal structure of βPGM$_{WT}$ (PDB 2WHE) showing residues of βPGM$_{WT}$ undergoing conformational exchange on different timescales. Residues which populate two conformations in slow exchange are coloured in shades of blue according to chemical shift differences between conformer A and conformer B, with the intensity of colour and thickness of the backbone corresponding to larger values. Residues in conformer A and conformer B with missing backbone amide peaks in the ^{1}H^{15}N-TROSY spectrum of βPGM$_{WT}$ are coloured black, whereas missing backbone amide peaks in conformer B only are coloured purple. The amide ^{1}H^{15}N coherences are likely broadened beyond detection due to intermediate exchange on the millisecond timescale. The catalytic Mg^{2+} ion is indicated as a green sphere.

The inorganic ionic composition of *L. lactis* cytoplasm (~2 mM Mg^{2+}, ~50 mM Na^+, ~400 mM K^+, ~50 mM Cl^-) (Goel et al., 2012) overlaps with the concentration ranges tested, where the population distribution between conformer A and conformer B remained unaffected. Therefore, it is expected that both conformer A and conformer B are populated in cytoplasm. However, the intracellular milieu is a complex mix of metabolites that could influence this equilibrium. This environment was mimicked through the use of bovine skimmed milk, a medium in which *L. lactis* thrives within the dairy industry. It is anticipated that the organic components in milk will also be present in cytoplasm. Moreover, the inorganic ionic composition (~5 mM Mg^{2+}, ~24 mM Na^+, ~38 mM K^+, ~28 mM Cl^-) (Zamberlin et al., 2012) is similar to cytoplasm (except for K^+, which has no effect on the equilibrium between conformer A and conformer B), so any effects will be due to the influence of metabolites. βPGM$_{WT}$ was diluted 5-fold into fresh skimmed milk, which had been filtered to remove species with a molecular weight larger than 10 kDa. The $^1H^{15}N$-TROSY spectrum revealed that both conformer A and conformer B were populated with a similar ratio (60% conformer A and 40% conformer B) to βPGM$_{WT}$ recorded in standard NMR buffer (Figure 5.2c, Figure 5.7a). However, minor chemical shift changes in the active site loops and the sharpening of some peaks that were line-broadened under standard conditions indicated that one of the milk components was binding in the vicinity of the active site. The two dominant organic components of the filtered milk were lactose and citrate (Figure 5.2d). Titration of lactose into βPGM$_{WT}$ had no effect on the $^1H^{15}N$-TROSY spectrum, whereas titration of citrate led to equivalent chemical shift changes and sharpening of line-broadened peaks to those observed in milk. Similar effects were observed in both conformer A and conformer B. Hence, βPGM$_{WT}$ was crystallised in the presence of citrate and the structure was determined to 2.1 Å resolution (PDB 6YDM; Figure 5.8a-c, Table 5.1). Two chains are present in the crystallographic asymmetric unit, one of which has citrate and acetate bound, whilst the other has tris and acetate bound. Citrate is coordinated in the active site by residues T16, H20, V47–R49 and A115–K117 and mimics substrate binding to some extent. Both monomers share a similar fold and overlay closely with a previously reported substrate-free βPGM$_{WT}$ structure (PDB 2WHE; Baxter et al., 2010; non-H atom RMSDs of 0.56 Å and 0.95 Å). Although only one of the two conformers observed in solution is represented in the crystal, the NMR experiments show that both conformer A and conformer B remain well-populated under physiological conditions.

4.4.3 The conformational exchange involves *cis-trans* proline isomerisation

Exchange phenomena on multi-second timescales in proteins are often a consequence of *cis-trans* isomerisation of Xaa-Pro peptide bonds (Grathwohl and Wüthrich, 1981; Lu et al., 2007). The largest differences in chemical shift between conformer A and conformer B are observed for residues in a loop (D137–A147) containing two proline residues (P138 and P146) (Figure 5.4a). From the crystal structures of the substrate-free form of the enzyme (PDB 6YDL (Table 5.1) determined to 1.5 Å resolution, which compares closely with PDB 2WHE; Baxter et al., 2010 (non-H atom RMSD = 0.53 Å) and PDB 1ZOL; Zhang et al., 2005 (non-H atom RMSD = 0.65 Å)), nine *trans* Xaa-Pro peptide bonds are present in βPGM$_{WT}$, whereas the K145–P146 peptide bond adopts a *cis* conformation. Proline residues with *cis* peptide bonds have $^{13}C_\beta$ nuclei that resonate 2.0–2.5 ppm downfield from those with *trans* peptide bonds (Shen and Bax, 2010) and therefore the isomerisation state of the Xaa-Pro peptide bonds for βPGM$_{WT}$ in solution was investigated. All but one of the assigned proline residues in conformer A and conformer B possess $^{13}C_\beta$ chemical shifts in the range 30.4–31.9 ppm consistent with the population of *trans* Xaa-Pro peptide bonds (Figure 5.5b). In contrast, the $^{13}C_\beta$ chemical shift for P146 (34.7 ppm) corroborates the presence of a *cis* K145–P146 peptide bond in solution for conformer A. However for conformer B, the absence of proline $^{13}C_\beta$ resonances for P146 and P148, owing to millisecond conformational exchange in the K145–I150 region, precluded an identification of the isomerisation state for these proline residues using NMR methods.

To explore whether proline isomerisation at the K145–P146 peptide bond is the source of the multi-second conformational exchange in βPGM$_{WT}$, the βPGM variant P146A (βPGM$_{P146A}$) was prepared, and the solution properties of the substrate-free form were investigated. A ^{1}H^{15}N-TROSY spectrum shows that only a single species is present (Figure 4.2b, Figure 5.2b) and 194 out of a total of 211 residues (92%) were assigned using standard TROSY-based methodology (BMRB 27920; Cruz-Navarrete et al., 2019). The same seventeen residues as βPGM$_{WT}$ remain unassigned owing to millisecond conformational exchange. The chemical shifts of βPGM$_{P146A}$ were compared with those of conformer A and conformer B of βPGM$_{WT}$ (Figure 5.4b, c). While the largest perturbations relate to the mutation site (together with an associated propagation of effects through the P148–V158 and S171–G182 α-helices), additional significant and widespread chemical shift differences are present between conformer

A and βPGM$_{P146A}$, especially in the D137–A147 loop. In contrast, much smaller chemical shift changes are observed between conformer B and βPGM$_{P146A}$, indicating that the solution conformations for these species are closely similar. However, although the K145–A146 peptide bond in βPGM$_{P146A}$ is likely to adopt a *trans* conformation as the dominant population, the isomerisation state remains ambiguous using NMR methods. Therefore, βPGM$_{P146A}$ was crystallised and the structure was determined to 2.0 Å resolution (PDB 6YDK; Table 5.1). The cap and the core domains of the crystal structure superimpose closely with those of βPGM$_{WT}$ (non-H atom RMSD = 0.33 Å, PDB 2WHE; Baxter et al., 2010; non-H atom RMSD = 0.48 Å, PDB 6YDL; non-H atom RMSD = 0.51 Å, PDB 1ZOL; Zhang et al., 2005). The D137–A147 loop exhibits elevated temperature factors, consistent with the lower predicted RCI-S^2 values derived from NMR chemical shifts (Figure 5.5c). The electron density is best fit by the *trans* conformation of the K145–A146 peptide bond (ω dihedral angle = -177°) (Figure 4.4a, c). In comparison to βPGM$_{WT}$, the D137–A147 loop adopts a different conformation, although both a 3$_{10}$-helix (D137–V141) and a β-turn hydrogen bond (A147$_{HN}$–S144$_{CO}$) are retained. These perturbations in structure are consistent with the chemical shift changes observed between conformer A and βPGM$_{P146A}$ and support the βPGM$_{P146A}$ structure being a close model of conformer B.
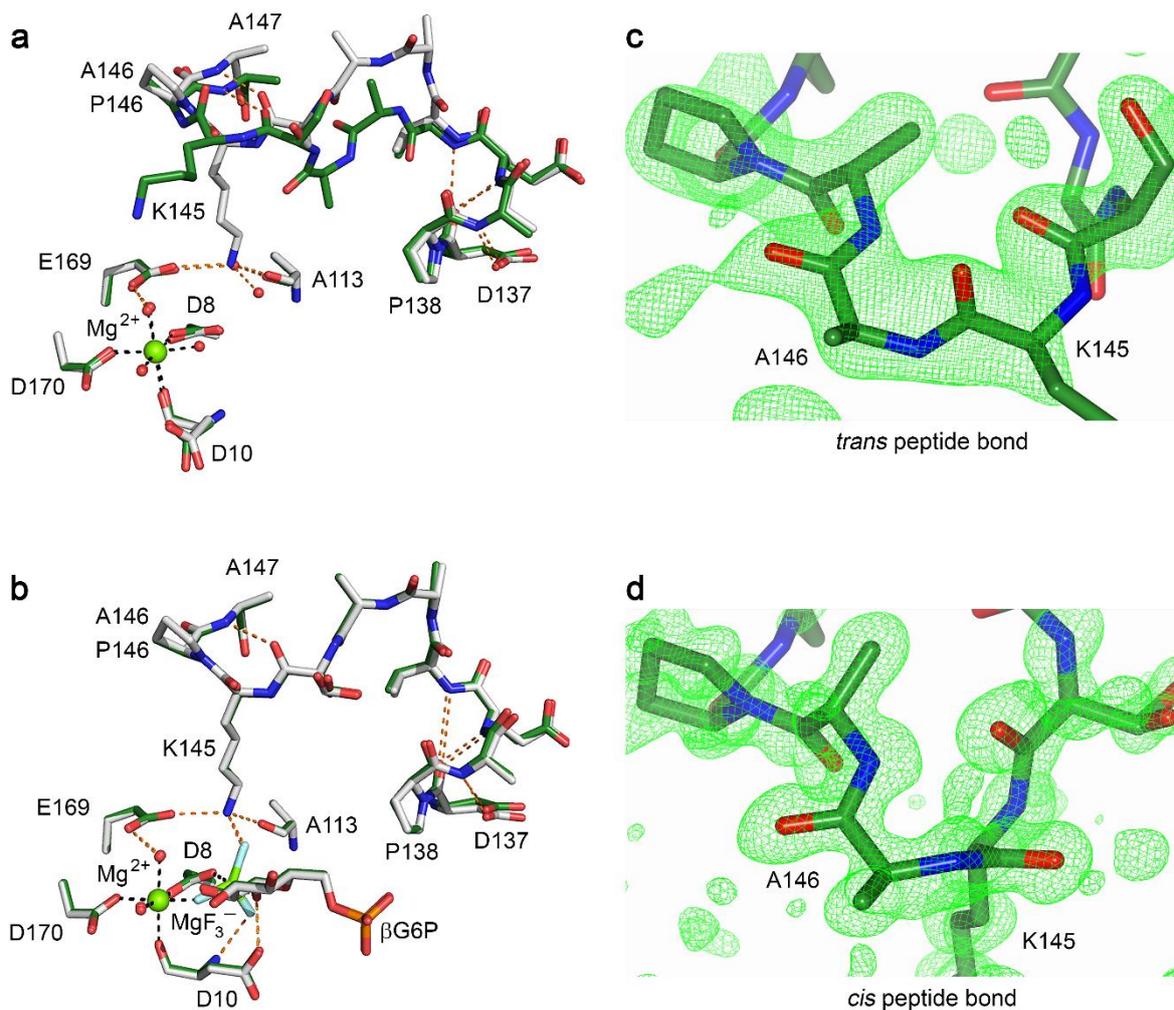
Figure 4.4 Conformational plasticity of the active site of βPGM. a, b, Active sites of βPGM$_{WT}$ (as conformer A) and βPGM$_{P146A}$ superposed on the core domain. (a) Selected residues are shown as sticks for the crystal structures of βPGM$_{WT}$ (grey carbon atoms; PDB 6YDL) and βPGM$_{P146A}$ (dark green carbon atoms; PDB 6YDK). In βPGM$_{WT}$, a *cis* K145–P146 peptide bond allows coordination of the K145 sidechain by E169 and A113, whereas in βPGM$_{P146A}$ a *trans* K146–A146 peptide bond changes significantly the backbone conformation of the D137–A147 loop, which precludes active site engagement of the K145 sidechain. The catalytic Mg^{2+} ion is drawn as a green sphere; black dashes indicate metal ion coordination and orange dashes show probable hydrogen bonds. (b) Selected residues, the MgF$_3^-$ moiety and G6P are shown as sticks for the crystal structures of the βPGM$_{WT}$:MgF$_3$:G6P TSA complex (grey carbon atoms; PDB 2WF5) and the βPGM$_{P146A}$:MgF$_3$:G6P TSA complex (dark green carbon atoms; PDB 6YDJ). βPGM$_{WT}$ maintains the *cis* K145 K146–A146 peptide bond, whereas βPGM$_{P146A}$ changes the isomerisation state of the K145–A146 peptide bond from a *trans* conformation in the substrate-free enzyme to a *cis* conformation in the transition state. c, d, omit map generated by refinement in the absence of residues S144–P148 in βPGM$_{P146A}$. (c) The S144–P148 segment, containing a *trans* K145–A146 peptide bond, with positive difference density (Fo – Fc; green mesh contoured at +2.5σ) in substrate-free βPGM$_{P146A}$. (d) The S144–P148 segment, containing a *cis* K145–A146 peptide bond, with positive difference density (Fo – Fc; green mesh contoured at +2.5σ) in the βPGM$_{P146A}$:MgF$_3$:G6P TSA complex.

The most pronounced consequence of the change in isomerisation state of the K145–A146 peptide bond is the failure of the K145 sidechain in βPGM_{P146A} to engage in the active site (Figure 4.4a). Instead, this sidechain is positioned in the open cleft between the cap and core domains and is exposed to solvent. In βPGM_{WT}, the ε-amino group of K145 is coordinated by the carbonyl oxygen atom of A113, the carboxylate sidechain of E169, and a water molecule that is replaced in the transition state analogue (TSA) complex (PDB 2WF5; Baxter et al., 2010) by a fluoride ion that mimics an oxygen atom of the transferring phosphoryl group. An electrostatic relationship also exists between the ε-amino group and the carboxylate group of D8. In βPGM_{P146A}, the position of the missing ε-amino group of K145 and its βPGM_{WT} hydrogen bonding are satisfied by a water molecule. The predicted RCI-S^2 order parameters for βPGM_{P146A} and βPGM_{WT} (Figure 5.5a, c) share similar profiles apart from the D137–A147 loop region, where the RCI-S^2 values for βPGM_{P146A} indicate increased dynamic properties that broadly mirror those of conformer B in βPGM_{WT}. Together, these data reveal that βPGM_{P146A} reflects the properties of conformer B and link the chemical shift and RCI-S^2 differences between conformers to the isomerisation state of the K145–X146 peptide bond. Thus, the multi-second exchange between conformer A and conformer B in solution involves *cis-trans* proline isomerisation of the K145–P146 peptide bond.

4.4.4 βPGM_{WT} lag phase depends on the phosphorylating agent

To ensure that the extent of the lag phase observed previously with AcP is not a method dependent observation (Johnson et al., 2018), the effect of different phosphorylating agents on the mutase activity of βPGM_{WT} was investigated by monitoring the conversion of 50 μM βG1P to G6P with either F16BP (1 mM), AcP (8 mM) or βG16BP (10 μM) present as phosphorylating agents, using the standard glucose 6-phosphate dehydrogenase coupled assay (Golicnik et al., 2009). Despite the crucial involvement of βG16BP as the reaction intermediate in the catalytic cycle, its concentration in the cell can vary markedly and is dependent upon the concentration of βG1P. Therefore, F16BP is the most likely phosphorylating agent of βPGM *in vivo* when *L. lactis* is growing on glucose-rich media (~50 mM F16BP (Neves et al., 2005) versus K_m ~100 μM (Zhang et al., 2005)). AcP is also a potential activator *in vivo*, as although it is present at lower concentrations (1–3 mM in *Escherichia coli* (*E. coli*) (Bennett et al., 2009; Klein et al., 2007) versus K_m ~800 μM (Zhang et al., 2005)), it is inherently a much faster phosphorylating agent. In the coupled assay

experiments with βPGM$_{WT}$ (Figure 4.5a), when either F16BP or AcP was used as the phosphorylating agent, their progression curves display significant lag phases. The lag is considerably more pronounced in the F16BP experiment, and consequently the maximum rate of βG1P to G6P conversion is not achieved before the substrate is exhausted. When βG16BP was used as the phosphorylating agent the kinetic profile shows a linear, fast initial rate. Consequently, initial rate measurements were made at several βG1P and βG16BP concentrations (10–700 μM and 0.4–100 μM, respectively) and were globally fitted to an equation derived for a ping–pong mechanism with βG1P inhibition (Golicnik et al., 2009). Accurate fits were obtained at βG16BP concentrations up to 10 μM, as above this concentration the model no longer describes the data. At elevated βG16BP concentrations, the back reaction from βG16BP to βG1P becomes significant, and the free βG16BP concentration is attenuated owing to a multimeric interaction between βG16BP and Mg^{2+} ions (Johnson et al., 2018). Accordingly, the data above 10 μM βG16BP were omitted from the fitting. This analysis yielded values for k_{cat} of 382 ± 12 s^{-1}, K_m (βG1P) of 91 ± 4 μM, K_m (βG16BP) of 8.5 ± 0.3 μM and K_i (βG1P) of 1510 ± 100 μM (Figure 5.9a). These values are all higher than those previously reported (Zhang et al., 2005; Golicnik et al., 2009) for βPGM owing to the extension of the analysis to higher βG1P and Mg^{2+} concentrations.

4.4.5   βPGM catalysis utilises a *cis* K145–X146 peptide bond

To assess whether βPGM is active as conformer B, the effect of different phosphorylating agents on the mutase activity of βPGM$_{P146A}$ was investigated by monitoring the conversion of 50 μM βG1P to G6P with either F16BP (1 mM), AcP (8 mM) or βG16BP (10 μM) present as phosphorylating agents using the standard glucose 6-phosphate dehydrogenase coupled assay. As for βPGM$_{WT}$, the kinetic profiles for βPGM$_{P146A}$ display significant lag phases with F16BP and AcP, while the progression curve with βG16BP shows a linear, fast initial rate (Figure 4.5b). Consequently, initial rate measurements were made at several βG1P and βG16BP concentrations (5–500 μM and 2–100 μM, respectively) and were globally fit to the equation used for βPGM$_{WT}$ above. For βPGM$_{P146A}$, the fitting yielded values for k_{cat} of 19.2 ± 0.2 s^{-1}, K_m (βG1P) of 157 ± 3 μM, K_m (βG16BP) of 175 ± 3 μM and K_i (βG1P) of 933 ± 32 μM (Figure 5.9b). In addition, the equilibration of βG1P and G6P for both βPGM$_{WT}$ and βPGM$_{P146A}$ was monitored by ^{31}P NMR spectroscopy using AcP as a phosphorylating agent (Figure 5.9c, d) (Johnson et al., 2018). The time courses show

a similar overall profile together with the presence of the lag phase and subsequent fitting of the linear segments yielded a k_{obs} of 70 ± 30 s^{-1} for βPGM$_{WT}$ and a k_{obs} of 1.1 ± 0.2 s^{-1} for βPGM$_{P146A}$. The variation between the kinetic parameters derived using the two methods is caused by inhibition resulting from different levels of phosphate-containing species present in the assays. However, the data clearly demonstrate that βPGM$_{P146A}$ is active, with a ~20-fold reduction in k_{cat}, a ~21-fold increase in K_m for βG16BP and a similar K_m and K_i for βG1P, when compared with βPGM$_{WT}$.
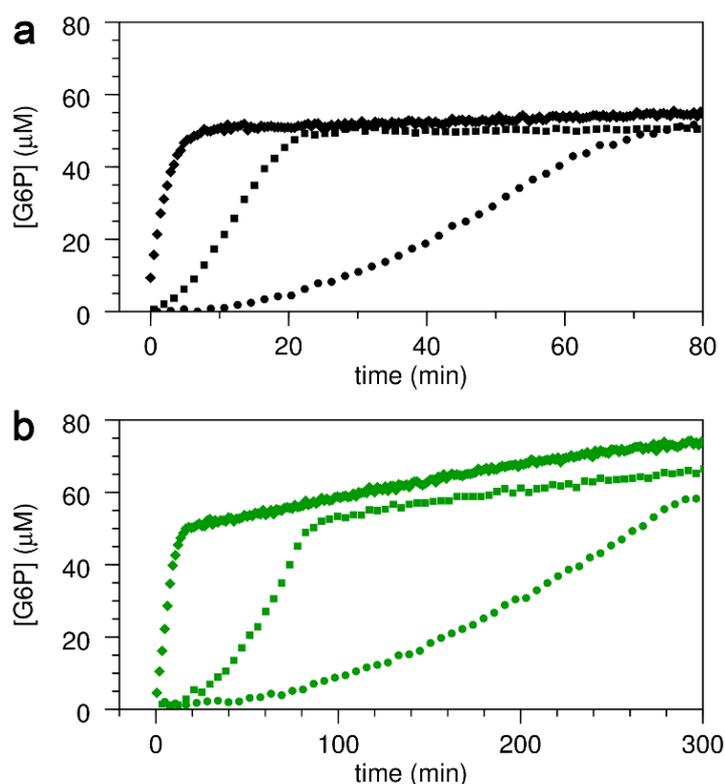
Figure 4.5 Kinetic profiles of βPGM activity. a, b, Reaction kinetics for the conversion of βG1P to G6P catalysed by βPGM$_{WT}$ and βPGM$_{P146A}$. The rate of G6P production was measured indirectly using a glucose 6-phosphate dehydrogenase coupled assay, in which G6P is oxidised and concomitant NAD$^+$ reduction is monitored by the increase in absorbance at 340 nm. Reaction catalysed by either (a) βPGM$_{WT}$ or (b) βPGM$_{P146A}$ in standard kinetic buffer using either F16BP (circles), AcP (squares) or βG16BP (diamonds) as a phosphorylating agent. For clarity, between 100% and 8% of the data points are included in the kinetic profiles. Following βG1P substrate depletion, the kinetic profiles show an apparent increase in G6P concentration, which results from: (1) the concentration of the reaction ingredients through evaporation from the assay plate wells and (2) for the reactions recorded using βG16BP, the enzyme-dependent conversion of remaining βG16BP to G6P via βPGM^P, occurring at a rate proportional to the amount of enzyme.

The mechanism of βPGM$_{P146A}$ catalytic activity was explored by preparing a TSA complex containing MgF$_3^-$ and G6P (Baxter et al., 2008; Baxter et al., 2010) and the resulting βPGM$_{P146A}$:MgF$_3$:G6P TSA complex was investigated using NMR spectroscopy. The observed ^{19}F chemical shifts are indistinguishable from those of the βPGM$_{WT}$:MgF$_3$:G6P TSA complex (Figure 5.9e, f) and a ^{1}H^{15}N-TROSY spectrum peak comparison (BMRB 7234; Baxter et al., 2006) indicates an almost identical correspondence between frequencies. Such close agreement allowed a backbone resonance assignment (211 residues – 100%) using 3D HNCACB and 3D HN(CA)CO spectra (BMRB 28097). Residues with the largest chemical shift differences between the βPGM$_{P146A}$:MgF$_3$:G6P and βPGM$_{WT}$:MgF$_3$:G6P TSA complexes are located within 4 Å of residue 146 (K145, A147 and A177) and within 5 Å of residue K145 (S48, V141 and A142) (Figure 5.4d). Taken together, these results confirm that βPGM$_{P146A}$ can assemble a stable and wild-type like βPGM$_{P146A}$:MgF$_3$:G6P TSA complex in solution. The βPGM$_{P146A}$:MgF$_3$:G6P TSA complex was crystallised, and the structure was determined to 1.0 Å resolution (PDB 6YDJ; Table 5.1). This complex superimposes very closely with the βPGM$_{WT}$:MgF$_3$:G6P TSA complex (non-H atom RMSD = 0.18 Å, PDB 2WF5; Baxter et al., 2010) and identifies both the positioning of the K145 sidechain in the active site and the *cis* K145–A146 peptide bond (ω dihedral angle = 14°; compared with ω dihedral angle = 12° for the K145–P146 peptide bond in the βPGM$_{WT}$:MgF$_3$:G6P TSA complex) (Figure 4.4b, d). The 6-phosphate group of G6P is in the distal site and the trigonal MgF$_3^-$ moiety mimicking the transferring phosphoryl group is coordinated in the proximal site between D8 (atom Oδ1) and the 1-OH group of G6P. The donor–acceptor distance and the angle of alignment are 4.1 Å and 174°, respectively (compared with 4.3 Å and 176°, respectively for the βPGM$_{WT}$:MgF$_3$:G6P TSA complex). The catalytic Mg^{2+} ion coordination also has comparable octahedral geometry to the βPGM$_{WT}$:MgF$_3$:G6P TSA complex and to substrate-free βPGM$_{P146A}$. Together, these data demonstrate that βPGM$_{P146A}$ is able to populate a *cis* K145–A146 peptide bond and achieve full domain closure with concomitant formation of transition state geometry. Additionally, assuming that βG16BP binding is diffusion controlled, the increase in K_m for βG16BP in βPGM$_{P146A}$ reflects the energetic cost of the *trans* to *cis* isomerisation of the K145–A146 peptide bond (Odefey, Mayr and Schmid, 1995). As in βPGM$_{WT}$, these results imply that conformer A of βPGM$_{P146A}$ represents the more active form.

4.4.6 βPGM forms two different transient phospho-enzyme species

The possible involvement of conformer B in the modulation of enzyme activity was investigated using real-time NMR methods by comparing the phosphorylation of βPGM under saturating conditions of either F16BP (50–100 mM), AcP (60–100 mM) or βG16BP (35 mM). Residue A113 is a well-resolved reporter of the relevant species – conformer A and conformer B, and their phosphorylated counterparts, A^P and B^P. The carbonyl group of A113 is coordinated by the ε-amino group of K145 (in conformer A) or a water molecule (in conformer B) and its amide proton is hydrogen bonded to the carbonyl group of F7 (adjacent to the D8 nucleophile) (Figure 4.4a, b).

On addition of F16BP to $βPGM_{P146A}$, the two dominant species observed are conformer B and B^P (Figure 4.2b). The presence of conformer B shows that the phosphorylation rate of $βPGM_{P146A}$ is very similar to the dephosphorylation rate of B^P (through hydrolysis), and only an apparent rate constant can be measured. The apparent rate constant for dephosphorylation was determined to be $0.003 ± 0.00002$ s^{-1} from the rate of reduction of the free F16BP concentration in ^1H NMR experiments. The ^1H and ^{15}N chemical shifts of B^P, assigned using fast acquisition 3D HNCO and 3D HNCA NMR experiments, mirror those of conformer B, except for the active site residues F7–D8, A113–A115 and hinge residues T16–E18, owing to their proximity to phosphorylated D8 (Figure 5.4e). Resonances from the D137–A147 loop show no significant differences between both forms, indicating that the K145–A146 peptide bond is in a *trans* conformation in B^P (the conformer B to A^P transition results in large chemical shift changes for the D137–A147 loop; Figure 5.4f). Conformer B and B^P are also observed when AcP was used as the phosphorylating agent, and a minor population of A^P is present, correlating with a small increase in the population of B^P relative to conformer B (Figure 5.7b). Identification of A^P is based on the assignment and structure of the $βPGM_{WT}:BeF_3$ complex (BMRB 17851; PDB 2WFA; Griffin et al., 2012), where the K145–P146 peptide bond is in a *cis* conformation and the K145 sidechain is engaged in the active site. Notably, when βG16BP was used as the phosphorylating agent, B^P is not observed (Figure 4.2b, Figure 5.7b). Instead, the A^P:G6P and B:βG16BP complexes are the primarily populated species. The A^P:G6P complex has similar chemical shifts to the $βPGM_{WT}:BeF_3$ complex, and the slow exchange between the B:βG16BP and the A^P:G6P complexes correlates with the measured k_{cat} values for $βPGM_{P146A}$.

In βPGM$_{WT}$, A^P is the dominant species observed on addition of F16BP (Figure 4.2a, Figure 5.3c). Therefore, the phosphorylation rate of βPGM$_{WT}$ by F16BP under these conditions must be faster than the hydrolysis rate of A^P ($k_{hydrolysis}$ = 0.06 ± 0.006 s^{-1}) (Johnson et al., 2018). A minor population of the A:F16BP complex is also present, indicating that the phosphorylation rate is slower than the chemical shift difference between the A^P and A:F16BP peaks (140 Hz). Significantly, a minor population of B^P is also observed. This species is populated transiently (~5 min) and disappears at longer timeframes, whereas A^P and the A:F16BP complex populations remain dominant while the phosphorylating agent is at high concentration. Hence, the B^P population is converting to the more stable A^P species with a rate constant of ≥0.003 s^{-1}, which mirrors the *trans* to *cis* isomerisation rate constants of Xaa-Pro peptide bonds in model peptides (Grathwohl and Wüthrich, 1981). Equivalent behaviour is observed when AcP was used as the phosphorylating agent (Figure 5.3b, Figure 5.7a), except that an A:AcP complex is not detected. When βG16BP was used as a phosphorylating agent, B^P does not accumulate at any point in the 3 hour time course. The only detectable species is an A:βG16BP complex (Figure 4.2a, Figure 5.3b, Figure 5.7a), which is identified by the similarity of chemical shift distribution with the βPGM$_{D10N}$:βG16BP complex (BMRB 27174; PDB 5OK1; Johnson et al., 2018). The low intensity of the A113 peak, along with peaks of other active site residues (Figure 4.2a, Figure 5.3b, Figure 5.7a), results from millisecond conformational exchange between species within the catalytic cycle, which correlates with the measured k_{cat} values for βPGM$_{WT}$. Overall therefore, the consequence of phosphorylation by the reaction intermediate βG16BP is markedly different to that of other phosphorylating agents, in that it generates no detectable B^P or lag phase, even when the initial population of conformer B is high.

4.5 Discussion

Substrate-free βPGM_{WT} exists in solution as two distinct conformers with near-equal populations, which differ in the isomerisation state of the K145–P146 peptide bond and interconvert at a rate between 0.003 s^{-1} and 1.0 s^{-1}. Conformer A contains the *cis*-isomer of this peptide bond, as observed in the crystal structures of substrate-free βPGM_{WT}, whereas conformer B contains the *trans*-isomer, as mimicked by the βPGM_{P146A} variant. In the crystal structure of βPGM_{P146A}, the *trans* K145–A146 peptide bond positions the sidechain of K145 away from the site of phosphoryl transfer, which is significantly different to its location in all other substrate-free and TSA complex structures reported for βPGM_{WT}. The removal of this positively charged amine group from the active site of conformer B disrupts the charge balance in the vicinity of the D8 nucleophile and therefore transition state stability will be severely impaired (Cliff et al., 2010; Jin et al., 2016). However, kinetics data for the mutase reaction indicate that βPGM_{P146A} is only ~20-fold less active than βPGM_{WT} and has a similar affinity for $\beta G1P$. In the crystal structure of the $\beta PGM_{P146A}:MgF_3:G6P$ TSA complex, which mimics the transition state of the phosphoryl transfer step, the K145–A146 peptide bond adopts the *cis*-isomer and the sidechain of K145 is able to engage in the active site. Hence, conformer A remains the preferred route for phosphoryl transfer in βPGM_{P146A}, despite it being ~1000 times less stable than in βPGM_{WT} because of the *trans* to *cis* isomerisation of the K145–A146 peptide bond (Odefey, Mayr and Schmid, 1995).

This disruption of the active site has regulatory significance, as conformer B of βPGM_{WT} constitutes a substantial population of the resting substrate-free enzyme. In vitro, linear initial kinetics and full activation of the enzyme are achieved only when $\beta G16BP$ is used as the phosphorylating agent. Phosphorylation of both βPGM_{WT} and βPGM_{P146A} with either F16BP or AcP leads to a lag phase, with the lag caused by F16BP lasting ~3.5 times longer than that observed with AcP. Moreover, real-time NMR experiments establish that the phosphorylation of both βPGM_{WT} and βPGM_{P146A} with either F16BP or AcP result in the generation of B^P. In βPGM_{WT}, B^P isomerises completely into A^P in <5 min, whereas in βPGM_{P146A}, B^P is consistently more populated than A^P. In contrast, B^P is not observed for either βPGM_{WT} or βPGM_{P146A} when $\beta G16BP$ is used as the phosphorylating agent. This result demonstrates that phosphorylation with $\beta G16BP$ leads to the stabilisation of conformer A, resulting in

production of A^P, regardless of the initial βPGM conformation, even for the βPGM$_{P146A}$ variant despite the considerable energetic cost of the conformer B to conformer A transition. Thus, conformer B and B^P are significantly less-active forms of βPGM, and the slow transition from B^P to A^P is part of the characteristic lag phase observed in the coupled assay kinetics. The longer lag in the F16BP experiments and the observation of a βPGM$_{WT}$:F16BP complex are consistent with F16BP having a slower phosphorylation rate than AcP. Therefore, βPGM is able to follow alternative kinetic pathways depending on the phosphorylating agent present, with its overall catalytic rate determined by the rate-limiting step in each pathway (Figure 4.6a, b).
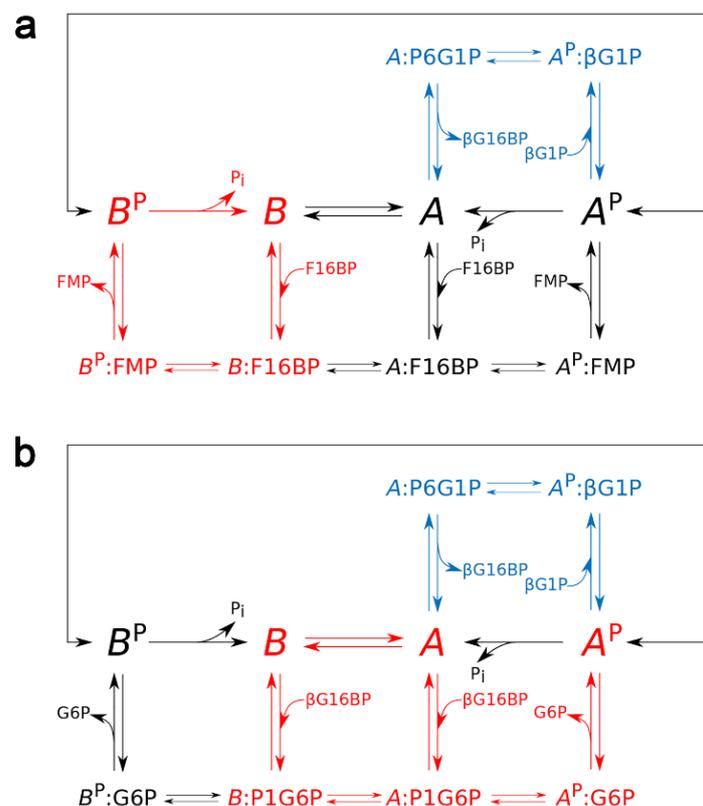
Figure 4.6 Kinetic model of βPGM activity. a, b, Reaction schemes for βPGM$_{WT}$ as conformer A or conformer B with different phosphorylating agents. The favoured pathways are shown (red text) for βPGM$_{WT}$ with (a) F16BP as a phosphorylating agent and (b) βG16BP as a phosphorylating agent. The βG16BP generating steps are highlighted in blue text. Fructose monophosphate (FMP) is either fructose 6-phopsphate or fructose 1-phosphate. The complexes X:P1G6P (X = A or B) and A:P6G1P denote explicitly the orientation of the βG16BP bound in the active site. The double-headed arrows connecting A^P and B^P indicate that these species interconvert with a multi-second exchange rate, similar to that described for the interconversion of conformer A and conformer B.

The response of βPGM to different phosphorylating agents also has functional significance. In *L. lactis*, the sole source of βG16BP is βPGM itself. In contrast, F16BP accumulates during glycolysis, reaching a concentration of up to ~50 mM (Neves et al., 2005). Hence, F16BP is likely to be the primary source of βPGM activation *in vivo*, with AcP contributing to a lesser extent (Bennett et al., 2009; Klein et al., 2007). Correspondingly, a proportion of βPGM will populate the less active B^P species. Following a switch from glucose to maltose or trehalose metabolism, which will greatly increase the βG1P concentration (Figure 5.1), βG16BP will begin to accumulate (Figure 4.1, Figure 4.6a, b). As a result, the conformer B and B^P populations will be recruited into the more active A^P species. This two-state control mechanism allows *L. lactis* to effectively catabolise maltose or trehalose, while the increase in transcription of *pgmB* is in progress (Andersson et al., 2005). Furthermore, when carbohydrate levels are low, a significant proportion of the basal βPGM population will be maintained as conformer B, which will consequently reduce the undesirable dephosphorylation of F16BP and also hinder the conversion of G6P to βG1P.

The multi-second substrate-dependent non-allosteric conformational exchange mediated through *cis-trans* proline isomerisation seen in βPGM represents a mechanism of post-translational enzyme regulation. This regulation mechanism relies on the existence of alternative pathways with different rate limiting steps, where the catalytic rate depends on the capacity of an activator, acting as a substrate, to bias the enzyme population towards the fastest pathway, by stabilising the most active conformation (Figure 4.7). Similar to allostery and allokairy, this mechanism depends on the ability of the enzyme to adopt at least two conformations with distinct activities but does not require the binding of an additional effector to the protein, nor an equivalence between the conformational exchange rate and the catalytic rate. We suggest the name allomorphy for this mechanism, from the greek *allos* meaning other and *morphe* meaning form, in keeping with the previously described allostery and allokairy mechanisms (Hilser, Anderson and Motlagh, 2015).
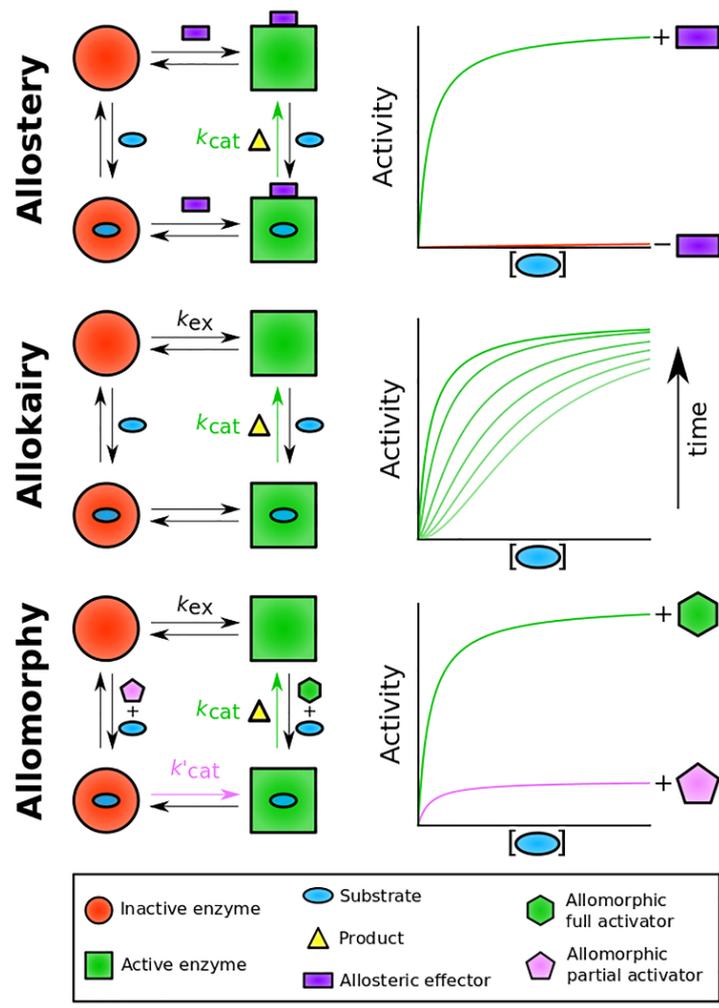
Figure 4.7 Mechanisms of regulation and activity profiles in monomeric enzymes.
In allostery, binding (or reaction) of an allosteric effector (purple rectangle) outside of the active site shifts the enzyme population from an inactive form (red circle and red profile) to an active form (green square and green profile), which stimulates the transformation of substrate (blue oval) to product (yellow triangle) at the catalytic rate (k_{cat}, green arrow). In allokairy, binding of substrate in the active site shifts the enzyme population from an inactive form to an active form, at an exchange rate (k_{ex}) that is similar to k_{cat}, resulting in time-dependent activity profiles (gradient of light green to dark green profiles). Following exhaustion of substrate, the enzyme population returns to the original equilibrium position. In allomorphy, reaction of the activating substrate, termed here allomorphic full activator (green hexagon), in the active site shifts the enzyme population from an inactive form to an active form, which stimulates the transformation of the native substrate (blue oval) to product (yellow triangle) at the maximal catalytic rate (k_{cat}, green arrow and green profile). However, reaction of alternatives substrates, termed here allomorphic partial activators (pink pentagon), in the active site are unable to shift the enzyme population from an inactive form to an active form, resulting in a slower overall catalytic rate (k'_{cat}, pink arrow and pink profile). The exchange rate (k_{ex}) between the two enzyme forms is much slower than k_{cat}. Following exhaustion of the allomorphic activator, the enzyme population returns to the original equilibrium position.

Allomorphy may modulate the activity of other monomeric enzymes with hysteretic behaviour, i.e., those that exhibit a burst or lag phase in their kinetic profile (Frieden, 1979). Several theoretical models have been put forward to rationalise hysteretic behaviour, such as the mnemonic (Ricard, Meunier and Buc, 1974) and the ligand-induced slow transition (Ainslie, Neet and Shill, 1972) models, but detailed structural-based molecular mechanisms have proved elusive. To our knowledge, only one such mechanism, allokairy in human glucokinase, has been described in detail (Whittington et al., 2015; Hilser, Anderson and Motlagh, 2015). Allomorphy is a different fine control regulatory mechanism and is potentially widespread, at least across phosphomutases; for example, both rabbit muscle and *L. lactis* α-phosphoglucomutases appear to be hysteretic enzymes (Chae and Markley, 2000; Zhang, Allen and Dunaway-Mariano, 2018), but belong to very different protein superfamilies. Like βPGM, these enzymes require a phosphorylating agent to initiate the catalytic cycle and, for the latter, the use of the reaction intermediate results in linear kinetics, whereas alternative phosphorylating agents produce a lag phase in their kinetic profiles. Similarly, α-phosphomannomutase from *Galdieria sulphuraria*, which also requires the addition of a phosphorylating agent to initiate the catalytic cycle, exhibits linear kinetics when α-mannose 1-phosphate and α-mannose 1,6-bisphosphate (or α-glucose 1-phosphate and αG16BP) are included in the reaction, but has a lag phase when there is a mismatch between substrate and phosphorylating agent, or when F16BP is used as the phosphorylating agent (Oesterhelt, Schnarrenberger and Gross, 1997). All of these observations are consistent with the presence of allomorphic control.

In summary, allomorphy is a fine control mechanism by which part of an enzyme population is maintained in a more latent state and is quickly switchable between high and low activity levels, without allosteric effectors. It delivers important control with which *L. lactis* is able to reconcile two seemingly contradictory requirements: the need to maximise its responsiveness to changes in carbohydrate source and the need to minimise unproductive diversion of valuable metabolites.

4.6 Data availability

Data supporting the findings of this manuscript are available from the corresponding author upon reasonable request. The atomic coordinates and structure factors have been deposited in the Protein Data Bank (www.rcsb.org) with the following codes: βPGM:citrate complex (PDB 6YDM), substrate-free βPGM (PDB 6YDL), substrate-free $βPGM_{P146A}$ (PDB 6YDK) and $βPGM_{P146A}$:MgF_3:G6P TSA complex (PDB 6YDJ). The NMR chemical shifts have been deposited in the BioMagResBank (www.bmrb.wisc.edu) with the following accession numbers: substrate-free βPGM conformer A (BMRB 28095), substrate-free βPGM conformer B (BMRB 28096) and $βPGM_{P146A}$:MgF_3:G6P TSA complex (BMRB 28097).

4.7 Code availability

Code developed in Python3 and bash for this study is publicly available under an MIT license and can be found on GitHub [http://doi.org/10.5281/zenodo.4022248].

4.8 Acknowledgements

4.9 Author contributions

H.P.W., F.A.C.N., N.J.B., C.R.T., A.J.R., J.P.W. designed research; H.P.W., F.A.C.N., C.R.T., A.J.R. produced isotopically enriched protein; H.P.W. performed coupled assay kinetic experiments; H.P.W., C.R.T. produced and purified βG1P and βG16BP; H.P.W., F.A.C.N., C.R.T., A.M.H., M.J.C. acquired NMR experiments; H.P.W., F.A.C.N., N.J.B., C.R.T. analysed NMR data; F.A.C.N., N.J.B., C.R.T. performed backbone resonance assignments; H.P.W., A.J.R., S.R.D. performed and analysed X-ray crystallography experiments; F.A.C.N. conceived and developed the allomorphy mechanism; H.P.W., F.A.C.N., N.J.B., C.R.T., J.P.W. wrote the paper with help from all authors.

4.10 Competing interests

The authors declare no competing interest.

5 Supplementary Information

SUPPLEMENTARY INFORMATION

Allomorphy as a mechanism of post-translational control of enzyme activity

Henry P. Wood[a,1], F. Aaron Cruz-Navarrete[a,1,*], Nicola J. Baxter[a,b,1], Clare R. Trevitt[a], Angus J. Robertson[a,c], Samuel R. Dix[a], Andrea M. Hounslow[a], Matthew J. Cliff[b], and Jonathan P. Waltho[a,b,*]

[a]Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield, S10 2TN, United Kingdom; [b]Manchester Institute of Biotechnology and School of Chemistry, The University of Manchester, Manchester, M1 7DN, United Kingdom; [c]Present address: Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

[1]H.P.W., F.A.C.N. and N.J.B contributed equally to this work

[*]To whom correspondence may be addressed: Prof. Jonathan Waltho and F. Aaron Cruz-Navarrete, Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, S10 2TN, United Kingdom, +44 114 22717, j.waltho@sheffield.ac.uk, ORCID 0000-0002-7402-5492, facruznavarrete1@sheffield.ac.uk, ORCID 0000-0002-5233-581X

Figure 5.1 Pathways for trehalose and maltose metabolism in *L. lactis*. Trehalose is transported into cells by the phosphoenolpyruvate-dependent phosphotransferase system yielding T6P, which is phosphorolysed by P_i-dependent trehalose 6-phosphate phosphorylase to βG1P and G6P. In contrast, maltose enters cells by the ATP-dependent permease system and is phosphorolysed by the action of P_i-dependent maltose phosphorylase to βG1P and glucose. Glucose is subsequently phosphorylated to G6P by glucokinase and enters glycolysis via fructose 6-phosphate (F6P) and F16BP. βPGM catalyses the isomerisation of βG1P to G6P, allowing complete catabolism of both trehalose and maltose. βPGM deficient *L. lactis* is unable to grow on trehalose and when cultured on maltose (disaccharide composed of α(1→4)-linked glucose units) the cells excrete βG1P into the growth medium and accumulate intracellularly βG1P (~0.7 M), T6P (~2.7 M) and amylose (polysaccharides composed of α(1→4)-linked glucose units). This observation is consistent with both P_i-dependent trehalose 6-phosphate phosphorylase and P_i-dependent maltose phosphorylase operating in the reverse sense (grey dotted arrows) to their physiological roles in wild-type *L. lactis*, resulting in excess βG1P being combined with G6P to form T6P or polymerised as α(1→4) glucose units to form amylose.
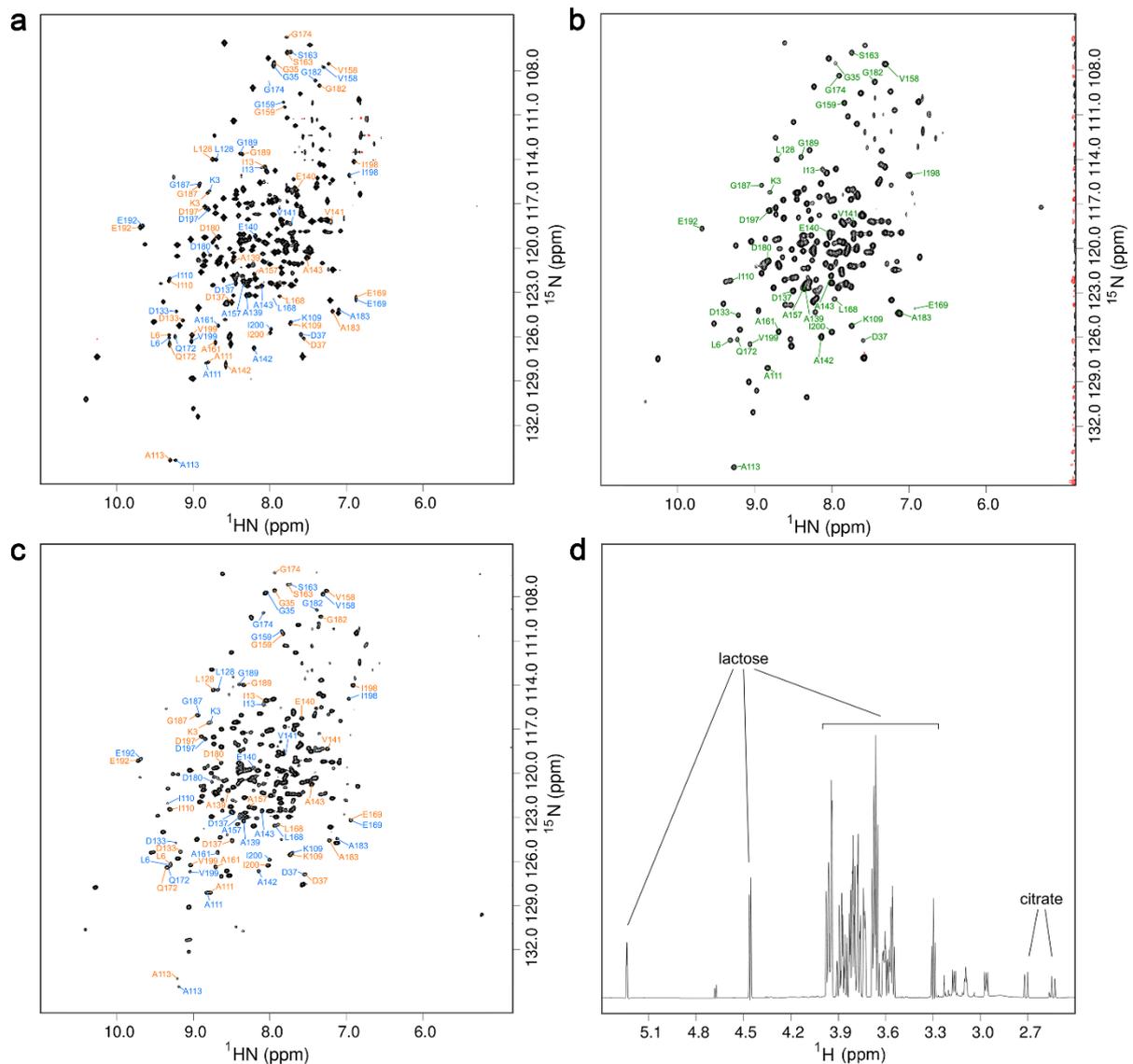
Figure 5.2 Solution behaviour of βPGM. (a) $^1H^{15}N$-TROSY spectrum of $^2H^{15}N^{13}C$-βPGM$_{WT}$ in standard NMR buffer containing 5 mM tris. A selection of well-resolved residues are labelled which populate conformer A (orange labels) and conformer B (blue labels) in slow exchange. (b) $^1H^{15}N$-TROSY spectrum of $^2H^{15}N^{13}C$-βPGM$_{P146A}$ in standard NMR buffer. βPGM$_{P146A}$ populates one conformer and the same selection of residues has been labelled in dark green for comparison. (c) $^1H^{15}N$-TROSY spectrum of ^{15}N-βPGM$_{WT}$ in filtered milk where both conformers are populated. (d) ^1H NMR spectrum of ^{15}N-βPGM$_{WT}$ in filtered milk showing the major milk components. The concentrations of lactose and citrate are estimated as 17 mM and 5 mM, respectively.
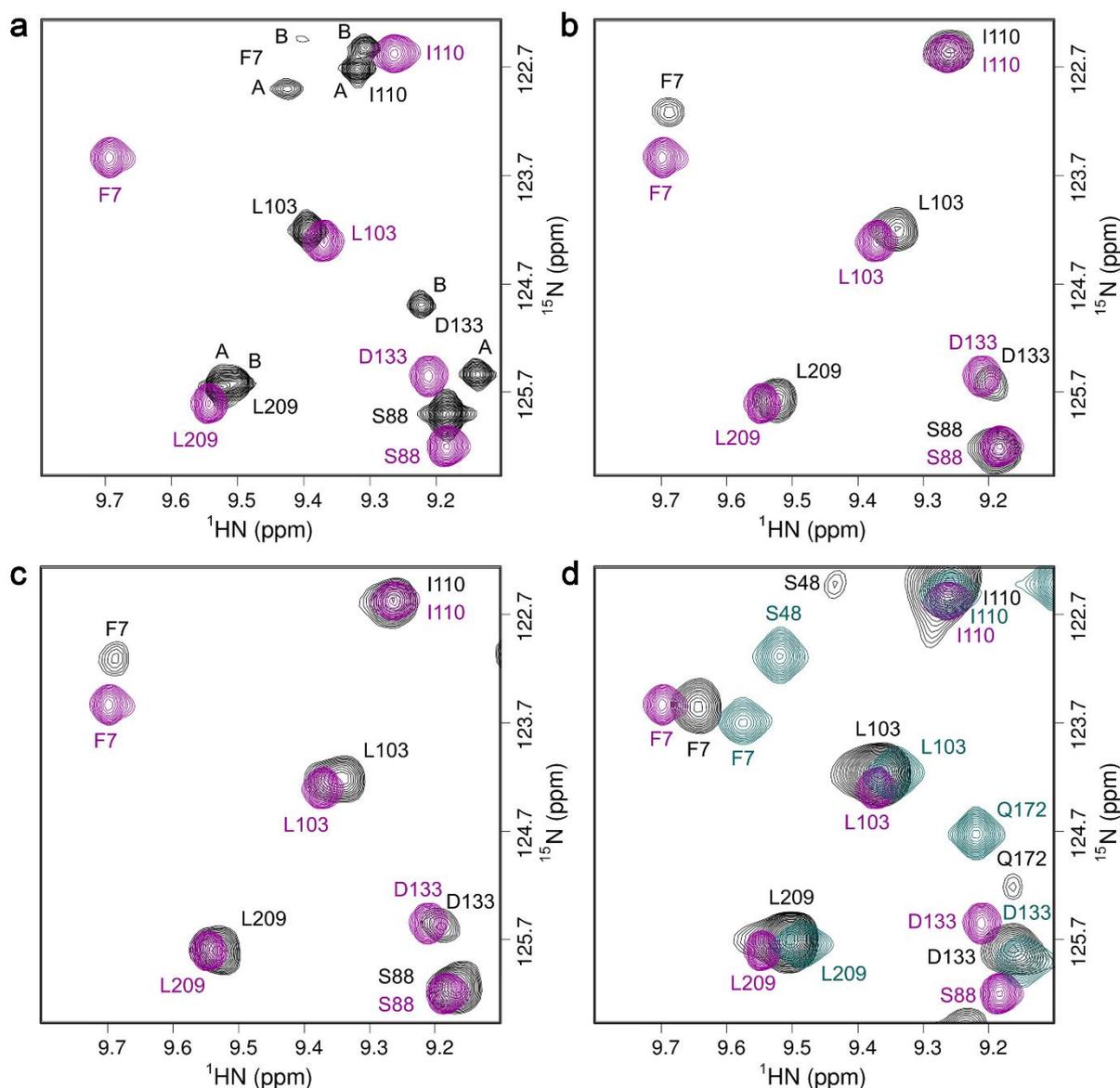
Figure 5.3 Effect of different phosphorylating agents on βPGM$_{WT}$. a–d, Comparative overlays of a section of ^{1}H^{15}N-TROSY spectra highlighting the behaviour of residues F7, S48, S88, L103, I110, D133, Q172 and L209. (a) βPGM$_{WT}$ (black) populates conformer A and conformer B in slow exchange for a subset of these residues. The addition of 3 mM BeCl$_2$ and 10 mM NH$_4$F to the βPGM$_{WT}$ sample induces the population of a single βPGM$_{WT}$:BeF$_3$ complex (magenta) (BMRB 17851), which is an analogue of phosphorylated conformer A (A^P). (b) βPGM$_{WT}$ supplemented with AcP (black) populates A^P as the dominant species, which shows a good degree of correspondence with the A^P analogue (magenta). (c) βPGM$_{WT}$ supplemented with F16BP (black) populates A^P as the dominant species, which again overlays well with the A^P analogue (magenta). (d) βPGM$_{WT}$ supplemented with βG16BP (black) populates a βPGM$_{WT}$:βG16BP complex (A:βG16BP), which shares a better correspondence with the Mg^{2+}-saturated βPGM$_{D10N}$:βG16BP complex (pale blue) (BMRB 27174) than with the A^P analogue (magenta).
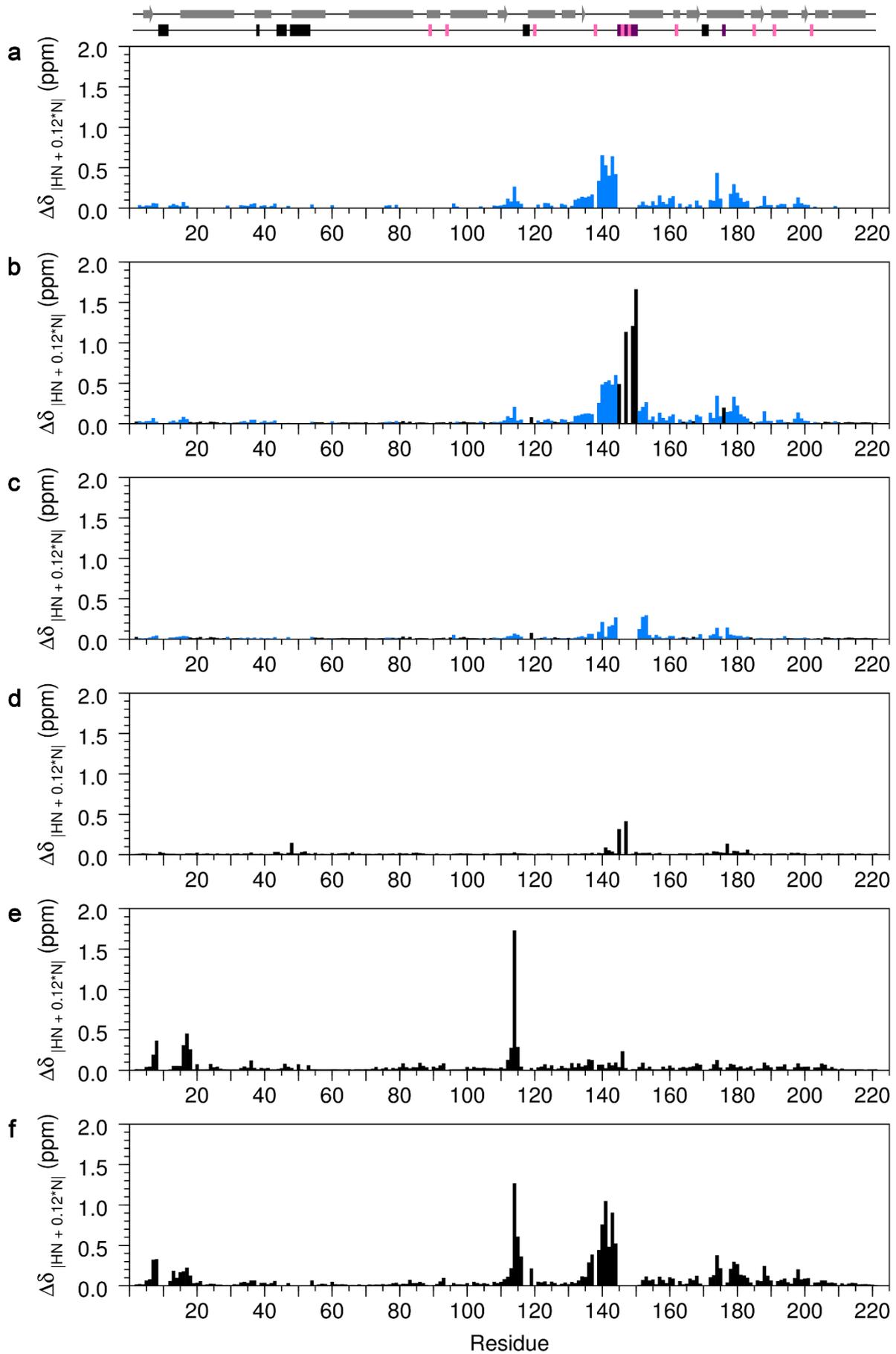
Figure 5.4 Residue specific backbone amide group chemical shift differences ($\Delta\delta$) for βPGM$_{WT}$ and βPGM$_{P146A}$. (a) Chemical shift differences between conformer A and conformer B. (b) Chemical shift differences between conformer A and βPGM$_{P146A}$. (c) Chemical shift differences between conformer B and βPGM$_{P146A}$. (d) Chemical shift differences between the βPGM$_{WT}$:MgF$_3$:G6P TSA and the βPGM$_{P146A}$:MgF$_3$:G6P TSA complexes. (e) Chemical shift differences between βPGM$_{P146A}$ and phosphorylated βPGM$_{P146A}$ (B^P). (f) Chemical shift differences between conformer B and the βPGM$_{WT}$:BeF$_3$ complex (A^P analogue). In panels (b) and (c), bars are coloured blue if residues in (a) showed a $\Delta\delta$ value greater than zero, otherwise bars are coloured black. Disregarding the dominant effect of the P146A substitution (large black bars), there are smaller and fewer chemical shift perturbations (blue bars) in (c) than in (b) suggesting that conformer B adopts a conformation closely similar to βPGM$_{P146A}$. At the top of the panel, secondary structure elements from βPGM$_{WT}$ (PDB 2WHE) are indicated by grey bars (α-helices) and arrows (β-strands). Residues in conformer A and conformer B with missing backbone amide peaks in the ^{1}H^{15}N-TROSY spectrum of βPGM$_{WT}$ are shown by black rectangles, whereas missing backbone amide peaks in conformer B only are shown by purple rectangles. Proline residues in βPGM$_{WT}$ are denoted by pink rectangles.
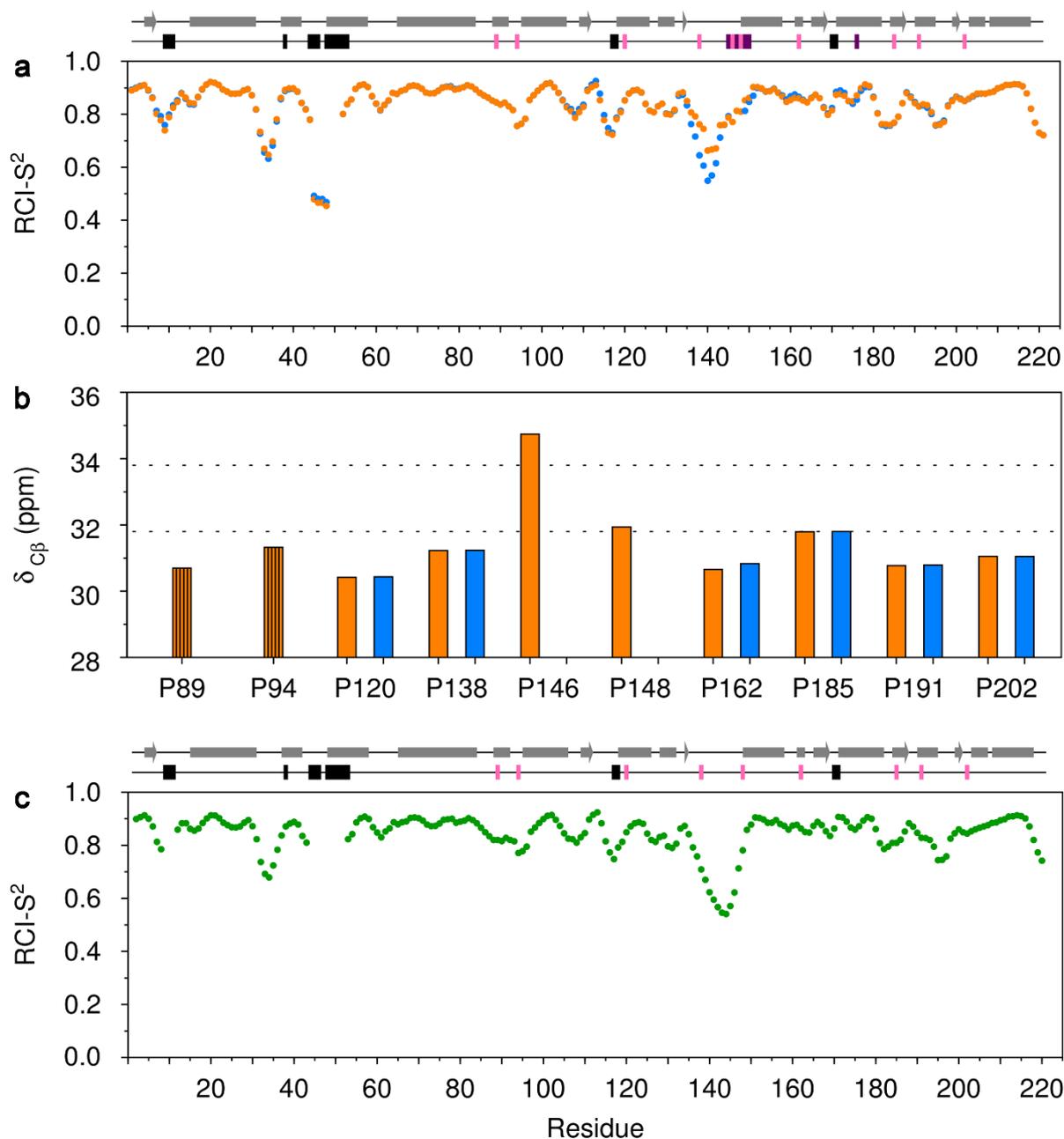
Figure 5.5 Chemical shift analysis of βPGM_WT and βPGM_P146A. (a) Random coil index order parameter (RCI-S^2) prediction of conformer A (orange circles) and conformer B (blue circles) obtained using TALOS-N. Secondary structure elements, the extent of assignment and proline residue locations are presented at the top of the panel for βPGM_WT as described previously. (b) Comparison of assignable proline $^{13}C_\beta$ chemical shifts in βPGM_WT for conformer A (orange bars) and conformer B (blue bars). Orange bars with vertical black shading indicate identical chemical shifts for conformer A and conformer B. The $^{13}C_\beta$ resonances of P146 and P148 in conformer B are likely to be broadened beyond detection due to conformation exchange in the K145–I150 region occurring on the millisecond timescale. Dotted horizontal lines represent average proline $^{13}C_\beta$ chemical shift values with *trans* (31.8 ± 1.0 ppm) and *cis* (33.8 ± 1.2 ppm) Xaa-Pro peptide bonds. (c) RCI-S^2 prediction of βPGM_P146A (dark green circles) obtained using TALOS-N. Secondary structure elements, the extent of assignment and proline residue locations are presented at the top of the panel for βPGM_P146A as described previously.

109
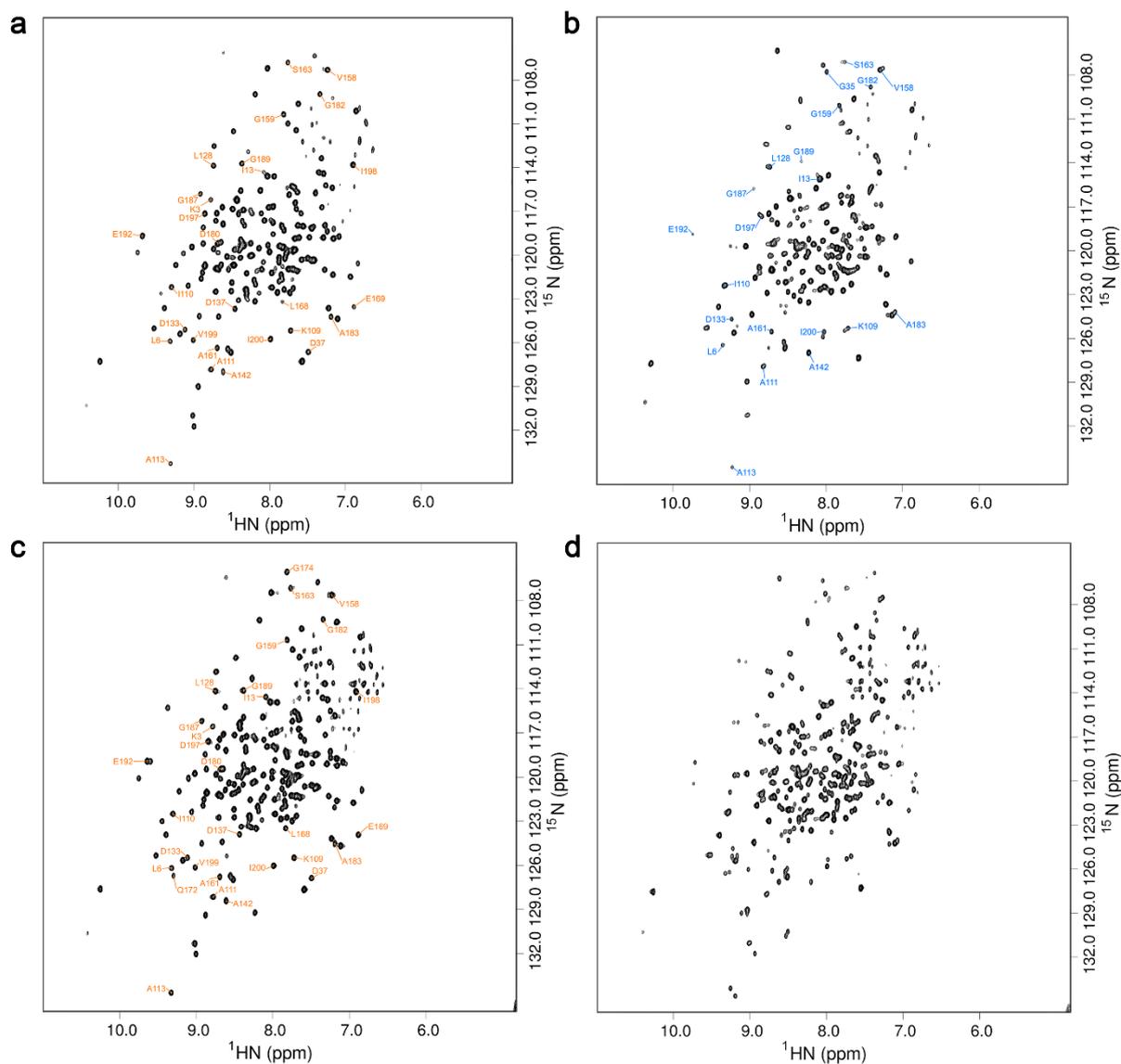
Figure 5.6 Solution behaviour of βPGM_WT under variable ion concentrations. (a) $^1H^{15}N$-TROSY spectrum of ^{15}N-βPGM_WT in standard NMR buffer containing 200 mM NaCl. The dominant population of βPGM_WT is conformer A. (b) $^1H^{15}N$-TROSY spectrum of ^{15}N-βPGM_WT in deionised water. The dominant population of βPGM_WT is conformer B. (c) $^1H^{15}N$-TROSY spectrum of ^{15}N-βPGM_WT in standard NMR buffer containing 100 mM MgCl₂. The dominant population of βPGM_WT is conformer A. (d) $^1H^{15}N$-TROSY spectrum of ^{15}N-βPGM_WT in Mg^{2+}-free standard NMR buffer. Both conformer A and conformer B are populated.
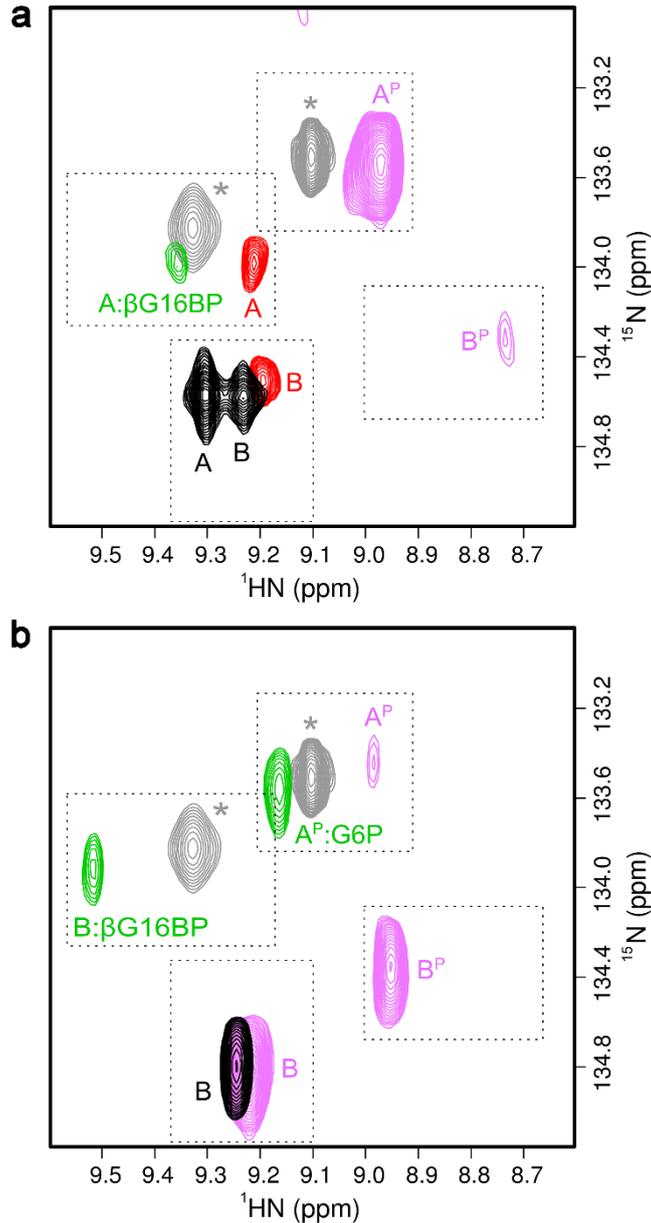
Figure 5.7 Effect of different phosphorylating agents on βPGM. a, b, Overlays of a section of $^1H^{15}N$-TROSY spectra for βPGM$_{WT}$ and βPGM$_{P146A}$ highlighting the behaviour of residue A113. (a) βPGM$_{WT}$ (black) populates conformer A and conformer B in slow exchange. βPGM$_{WT}$ recorded in filtered milk (red) populates conformer A and conformer B. βPGM$_{WT}$ supplemented with AcP (pink) populates A^P as the dominant species and B^P. βPGM$_{WT}$ supplemented with βG16BP (green) populates an A:βG16BP complex. (b) βPGM$_{P146A}$ (black) populates conformer B. βPGM$_{P146A}$ supplemented with AcP (pink) populates conformer B, A^P and B^P. βPGM$_{P146A}$ supplemented with βG16BP (green) populates an A^P:G6P complex and a B: βG16BP complex. Peaks indicated by grey asterisks correspond to the βPGM$_{WT}$:BeF$_3$ complex (grey; $δ_N$ = 133.5 ppm; BMRB 17851), which is an analogue of A^P, and the Mg^{2+}-saturated βPGM$_{D10N}$:βG16BP complex (grey; $δ_N$ = 133.8 ppm; BMRB 27174), which is a mimic of the A:βG16BP complex, and are shown for comparison.

111
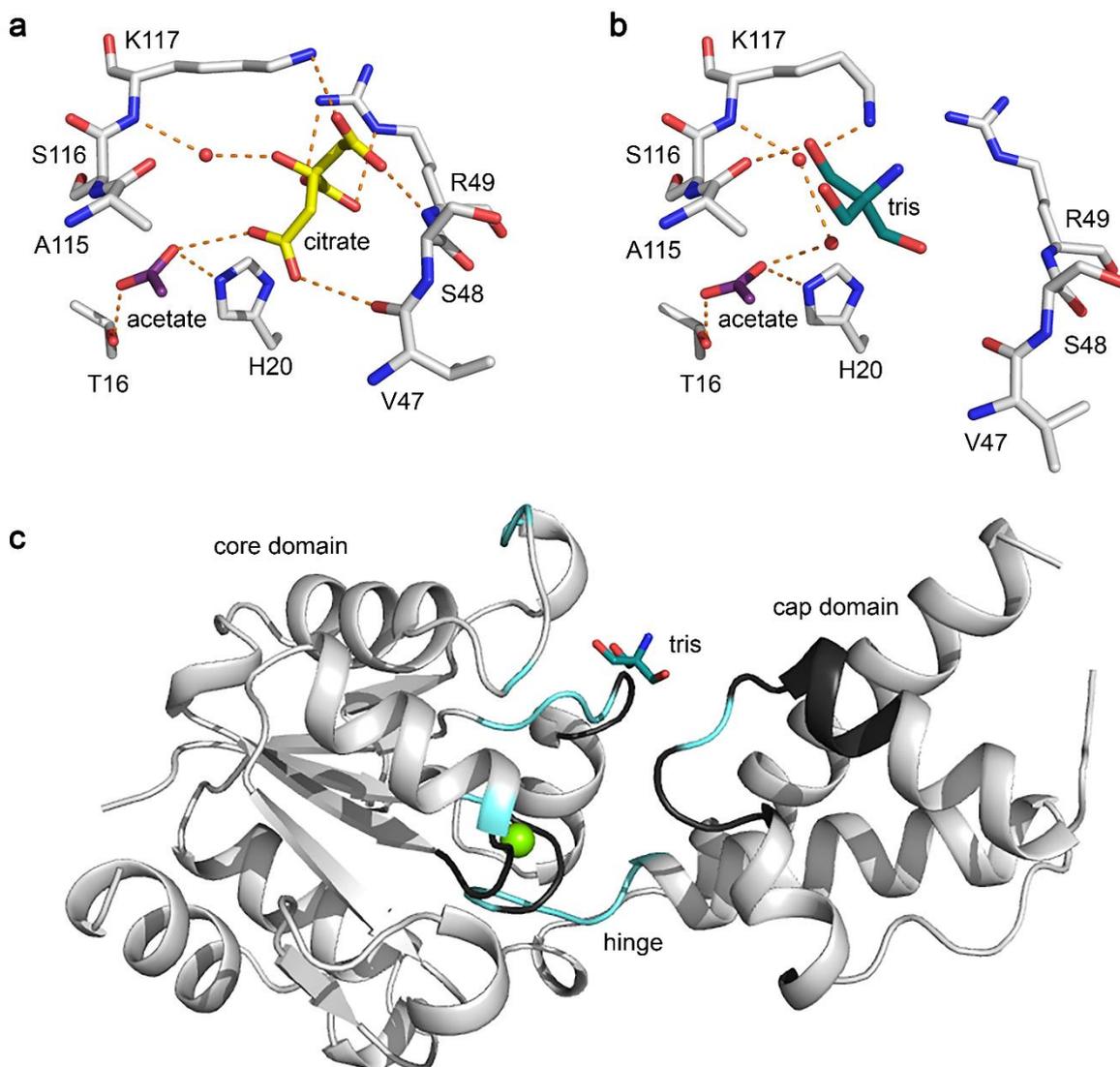
Figure 5.8 Binding of small molecules to βPGM_WT. a, b, Active site coordination for citrate, acetate and tris in the βPGM_WT:citrate complex (PDB 6YDM). βPGM_WT was crystallised in the presence of citrate in a crystallisation buffer containing acetate and tris. Selected active site residues and ligands are shown as sticks with βPGM_WT (grey carbon atoms), citrate (gold carbon atoms), acetate (purple carbon atoms) and tris (teal carbon atoms) for both chains of the crystallographic asymmetric unit. Red spheres indicate structural water molecules and orange dashes show probable hydrogen bonds. (c) Changes in intermediate exchange behaviour on tris binding. Cartoon representation of the βPGM_WT:citrate complex highlighting the extent of active site residues undergoing intermediate exchange behaviour in the original βPGM_WT assignment (without tris; cyan backbone, D8, L9, D10, G11, V12, I13, T14, D15, T16, R38, L44, K45, G46, V47, S48, R49, E50, D51, S52, L53, S114, A115, S116, K117, N118, V141, A142, K145, S171 and Q172; BMRB 7235) and in the assignment of conformer A (with tris; black backbone, L9, D10, G11, R38, L44, K45, G46, S48, R49, E50, D51, S52, L53, K117, N118, D170 and S171; BMRB 28095). The coordination of tris (teal carbon atoms) in the active site cleft perturbs the exchange behaviour for some loop residues, with the result that the corresponding resonances are no longer broadened beyond detection and can therefore be assigned. The catalytic Mg^{2+} ion is shown as a green sphere.
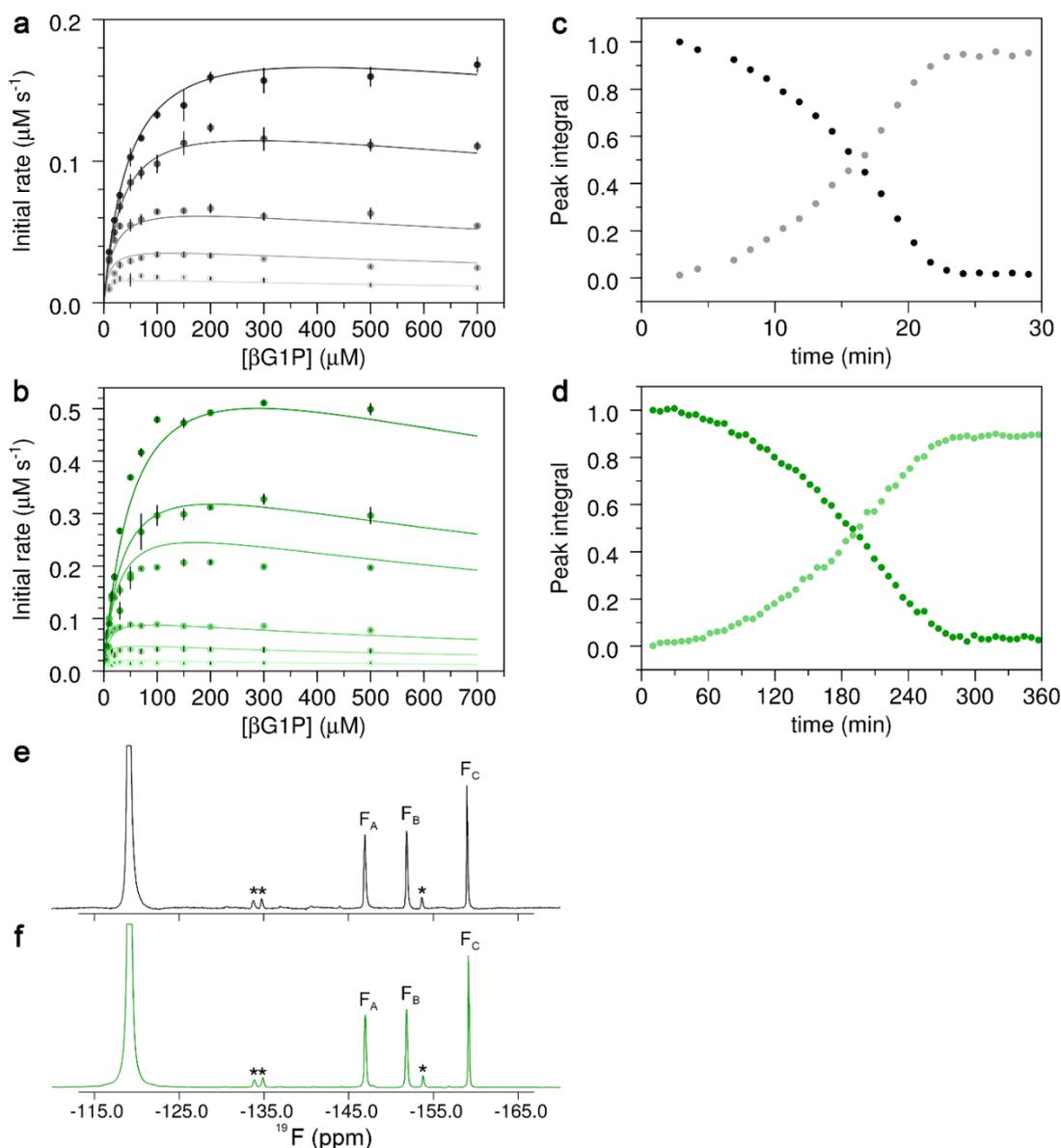
Figure 5.9 Activity of βPGM$_{WT}$ and βPGM$_{P146A}$. a, b, Initial rate measurements for βPGM$_{WT}$ and βPGM$_{P146A}$ using the coupled assay. The solid lines represent a global fit of the data to Equation 1 and vertical black lines indicate standard error of the mean of three replicate measurements. (a) βPGM$_{WT}$ initial rate measurements at a range of βG1P concentrations (10, 20, 30, 50, 70, 100, 150, 200, 300, 500, 700 μM) and βG16BP concentrations (0.4, 1, 2, 5, 10 μM, grey gradient increasing with concentration). (b) βPGM$_{P146A}$ initial rate measurements at a range of βG1P concentrations (5, 10, 15, 20, 30, 50, 70, 100, 200, 300, 500 μM) and βG16BP concentrations (2, 5, 10, 35, 50, 100 μM, green gradient increasing with concentration). c, d, Reaction kinetics monitored by ^{31}P NMR spectra for βPGM$_{WT}$ (grey tones) and βPGM$_{P146A}$ (green tones) for the equilibration of βG1P with G6P in standard kinetic buffer. The reactions were initiated by the addition of 20 mM AcP. Normalised integral values for the βG1P peak (black / dark green) and the G6P peak (grey / light green) are plotted as a function of time. e, f, ^{19}F NMR spectra of the βPGM:MgF$_3$:G6P TSA complexes formed using either βPGM$_{WT}$ (black) or βPGM$_{P146A}$ (green) in standard NMR buffer, supplemented with 15 mM NaF and 10 mM G6P. Chemical shifts are given in ppm for each ^{19}F resonance: βPGM$_{WT}$:MgF$_3$:G6P TSA complex (F$_A$ = −146.9, F$_B$ = −151.9 and F$_C$ = −159.0) and βPGM$_{P146A}$:MgF$_3$:G6P TSA complex (F$_A$ = −147.0, F$_B$ = −151.8 and F$_C$ = −159.2). Resonances indicated by black asterisks correspond to an alternative conformation of the βPGM:MgF$_3$:G6P TSA complexes. Free F$^−$ resonates at −119.1 ppm and the full peak intensity has been truncated for clarity.

Table 5.1 Data collection, data processing and refinement statistics for the βPGM complexes.

Complex	βPGM$_{WT}$:citrate	βPGM$_{WT}$	βPGM$_{P146A}$	βPGM$_{P146A}$:MgF$_3$:G6P
PDB Code	PDB 6YDM	PDB 6YDL	PDB 6YDK	PDB 6YDJ
Crystallisation conditions	0.6 mM βPGM$_{WT}$ 50 mM citrate	0.6 mM βPGM$_{WT}$	0.5 mM βPGM$_{P146A}$	0.4 mM βPGM$_{P146A}$ 10 mM G6P 15 mM NaF
Crystal morphology	Rod shaped crystals	Rod shaped crystals	Rod shaped crystals	Large plate crystals
Wavelength (Å) Beamline, Facility	0.97179 Beamline i03, DLS	0.92819 Beamline i04-1, DLS	0.97950 Beamline i04, DLS	0.91587 Beamline i04-1, DLS
Resolution (Å) [1]	46.57 – 2.10 (2.16 – 2.10)	44.65 – 1.52 (1.56 – 1.52)	43.95 – 2.02 (2.05 – 2.02)	54.25 – 1.04 (1.06 – 1.04)
Space group	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$
Cell dimensions: a, b, c (Å) α, β, γ (°)	53.1, 76.6,117.3 90.0, 90.0, 90.0	53.3, 54.1, 81.9 90.0, 90.0, 90.0	53.3, 56.2, 77.7 90.0, 90.0, 90.0	37.1, 54.3, 104.3 90.0, 90.0, 90.0
Total reflections [1]	239876 (15971)	264843 (16619)	103063 (2568)	680305 (21980)
Unique reflections [1]	27995 (2095)	36815 (2658)	15683 (686)	101730 (5021)
Multiplicity [1]	8.6 (7.6)	7.2 (6.3)	6.6 (3.7)	6.7 (4.4)
Completeness (%) [1]	97.5 (90.1)	99.3 (99.3)	98.8 (88.9)	99.9 (99.3)
$<I/\sigma I>$ [1]	13.7 (3.7)	18.8 (1.3)	14.4 (1.2)	12.4 (1.1)
Wilson B factor (Å^2)	24.2	21.2	32.2	7.4
R$_{merge}$ [1]	0.093 (0.536)	0.044 (1.190)	0.075 (0.953)	0.067 (1.138)
R$_{pim}$ [1]	0.033 (0.198)	0.019 (0.561)	0.031 (0.506)	0.028 (0.609)
CC-half [1]	0.999 (0.899)	1.000 (0.536)	0.999 (0.556)	0.999 (0.544)
Molecular replacement model	PDB 2WHE	PDB 2WHE	PDB 2WHE	PDB 2WF5
R$_{factor}$	0.230	0.183	0.214	0.149
R$_{free}$	0.290	0.216	0.253	0.169
Number of atoms: Protein Ligands Metal ions Water	3379 38 2 187	1733 0 1 148	1708 0 1 36	1704 29 2 209
Protein residues	438	219	221	218
RMS deviations: Bonds (Å) Angles (°)	0.0089 1.482	0.0145 1.497	0.0125 1.522	0.0093 1.473
Average B factors (Å^2) Main chain Side chains Ligands Metal ions Water	32 36 50 27 33	25 29 N/A 27 36	15 27 N/A 44 39	11 13 10 10 24
Ramachandran analysis Favoured/allowed (%) Disallowed (%) Favoured rotamers (%) Poor rotamers (%)	97.7 0.0 94.1 1.10	98.6 0.0 96.2 0.54	97.7 0.0 96.7 1.11	97.7 0.0 95.1 0.55
MolProbity score	1.29 (99th percentile, 2.10 ± 0.25 Å)	0.93 (100th percentile, 1.52 ± 0.25 Å)	0.98 (100th percentile, 2.02 ± 0.25 Å)	0.66 (99th percentile, 1.04 ± 0.25 Å)

[1] Values for the higher resolution shell are in parentheses

6 Enzymes that catalyse intrinsically difficult reactions harvest substrate binding energy for stabilisation of the closed state but at the risk of introducing substrate inhibition

Enzymes that catalyse intrinsically difficult reactions harvest substrate binding energy for stabilisation of the closed state but at the risk of introducing substrate inhibition

Angus J. Robertson,[1,#] F. Aaron Cruz-Navarrete,[1,#] Henry P. Wood,[1] Nikita Vekaria,[2] Andrea M. Hounslow,[1] Claudine Bisson,[1,4] Matthew J. Cliff,[2] Nicola J. Baxter,[1,2] and Jonathan P. Waltho[*,1,2]

[1] Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield, S10 2TN, United Kingdom

[2] Manchester Institute of Biotechnology and School of Chemistry, The University of Manchester, Manchester, M1 7DN, United Kingdom

[#]A.J.R. and F.A.C.N. contributed equally to this work

[*]To whom correspondence may be addressed: Prof. Jonathan P. Waltho, Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield, S10 2TN, +44 114 22717, j.waltho@sheffield.ac.uk, ORCID 0000-0002-7402-5492

6.1 Abstract

Understanding the factors that underpin the enormous catalytic proficiencies of enzymes is a primary aim of fundamental catalysis and enzyme design. Enzymes are able achieve high catalytic proficiencies through utilisation of the binding energy derived from non-reacting portions of the substrate. In particular, enzymes that have substrates with non-reacting phosphodianion groups use this binding energy to facilitate the conformational change from an open active from to a closed active form, rather than to either induce ground state destabilisation or stabilise the transition state. β-Phosphoglucomutase (βPGM) catalyses the isomerisation of β-glucose 1-phosphate to glucose 6-phosphate via phosphoryl transfer reactions in the proximal site, whilst coordinating the non-reacting phosphodianion group of the substrate(s) in the distal site. Additionally, βPGM has one of the largest catalytic proficiencies measured and undergoes significant domain closure during its catalytic cycle. Therefore, this enzyme is a suitable model to investigate the relationship between non-reacting phosphodianion group binding, enzyme closure and catalytic proficiency. Here we show that perturbation of the non-reacting phosphodianion group coordination results in decreased substrate binding that hinders a successful enzyme closure. These results, together with those derived from other non-reacting phosphodianion group binding enzymes, uncover a trend, whereby enzymes that catalyse intrinsically difficult chemistry exhibit a greater requirement to utilise the substrate binding energy to stabilise the closed active form. Additionally, no significant communication is present between the proximal and distal sites, allowing the structural characterisation of substrate inhibition, which reveals that distal site phosphodianion group binding facilitates enzyme closure even in a non-productive binding mode.
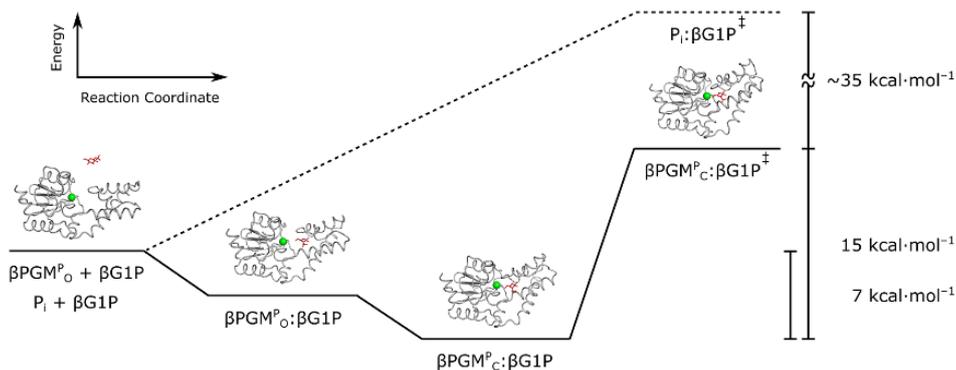
Figure 6.1 Free energy reaction profile for βPGM. Free energy levels of the first half of the catalytic cycle of βPGM. First, the open phosphoenzyme (βPGM^P_O) binds βG1P to form a transient open βPGM^P_O:βG1P complex. The substrate binding energy is then utilised to facilitate the adoption of an active closed βPGM^P_C:βG1P complex as the dominant Michaelis complex. The free energy of stabilisation of the Michaelis complex is 7 kcal mol^{-1} (calculated using the apparent K_d (G6P) of the βPGM_{WT}:AlF$_4$:G6P complex). The free energy difference between the βPGM^P_C:βG1P complex and the βPGM^P_C:βG1P‡ transition state is 15 kcal mol^{-1} (calculated using the observed catalytic rate constant of βPGM_{WT}). Through its catalytic mechanism, βPGM stabilises the transition state of phosphoryl transfer by ~35 kcal mol^{-1}, compared to the same reaction in water (dashed lines) (Lad, Williams and Wolfenden, 2003).

118

6.2 Introduction

The ability of enzymes to achieve enormous catalytic proficiencies remains the subject of intense investigation, leading to continual progress in the understanding of enzyme active site electronics, structure and dynamics. Electrostatic stabilisation of the chemical transition state (Warshel, Arieh et al., 2006; Garcia-Viloca et al., 2004), ground state destabilisation (Jencks, 1975; Andrews, Fenn and Herschlag, 2013), efficient formation of near attack conformers in the ground state (Hur and Bruice, 2003) and contributions from conformational motions (Benkovic and Hammes-Schiffer, 2003; Hammes-Schiffer and Benkovic, 2006; Schnell, Dyson and Wright, 2004) are all argued to contribute to catalytic proficiency. Additionally, stabilising interactions between the enzyme active site and non-reacting portions of the substrate (Jencks, 1975) are also thought to play an important role. Hexokinase, for example, is able to catalyse phosphoryl transfer from ATP to glucose 4×10^4-fold more rapidly than the equivalent reaction involving water, and this rate acceleration was ascribed to interactions with non-reacting parts of glucose rather than differences in the chemical reactivity of the two substrates (Koshland, 1960). More recently, the study of the contribution of distal site interactions to catalytic proficiency has focused on enzymes that bind substrates via their non-reacting phosphodianion groups (Amyes, O'Donoghue and Richard, 2001; Amyes, Richard and Tait, 2005; Amyes and Richard, 2013; Tsang, Amyes and Richard, 2008). Despite the substantially different transition states stabilised by three enzymes, glycerol 3-phosphate dehydrogenase (GPDH), orotidine 5'-monophosphate decarboxylase (OMPDC) and triose phosphate isomerase (TIM), the interaction between the enzyme and the phosphodianion groups in these systems contributes a consistent 11–13 kcal·mol^{-1} to the reduction of the activation energy barrier for their reactions (Amyes, O'Donoghue and Richard, 2001; Amyes, Richard and Tait, 2005; Tsang, Amyes and Richard, 2008; Morrow, Amyes and Richard, 2008). In each of these enzymes, a phosphodianion group is held in a positively charged distal site, where 50–80% of the intrinsic binding energy is provided through interactions with either a single arginine residue in GPDH and OMPDC, or a lysine residue in TIM.

The contribution of the distal site cation-phosphodianion interaction to enzyme catalytic proficiency can be obtained from kinetic studies (Go et al., 2010; Goryanova et al., 2013; Reyes et al., 2015a). Enhanced catalytic proficiency can be achieved

through sequestration of the substrate(s) in a low dielectric environment, coordinated extensively by a network of electrostatic interactions within a closed enzyme active site. Enzyme closure in this context involves either large conformational changes or rearrangement of flexible loops (Malabanan, Amyes and Richard, 2010; Richard et al., 2014; Kulkarni et al., 2018). In this framework, the energy derived from the cation-phosphodianion interaction is used to stabilise a closed active form (E_C:S) as the Michaelis complex (defined here as the lowest free energy enzyme-substrate complex conformation populated in the reaction coordinate), rather than to induce ground state destabilisation. However, if the utilisation of this energy is perturbed by distal site mutation, then the Michaelis complex switches in favour of an open inactive form (E_O:S). At this point, enzyme closure (E_O:S → E_C:S) becomes part of the rate limiting step of the reaction. An underlying assumption is that there is no significant communication of energy between the distal site phosphodianion group binding event and the site of the transition state for the chemical step. Simulations of TIM support this assumption, and also suggest that E_C:S is catalytically competent, regardless of the presence or absence of the substrate phosphodianion group (Kulkarni et al., 2018). Hence, the intrinsic phosphodianion binding energy is completely expressed at E_C:S, rather than being used to reduce the height of the energetic barrier of the catalytic step. Although this binding energy is consistent in magnitude across the three systems studied previously, the nature of the favored Michaelis complex is not. For GPDH and TIM (catalyzing hydride transfer and proton transfer reactions, respectively), the Michaelis complex is E_O:S (Go et al., 2010; Reyes et al., 2015a), and either a large domain reorientation or small loop rearrangements are observed upon enzyme closure (Mydy et al., 2019; Malabanan, Amyes and Richard, 2010). In contrast, for OMPDC (catalysing the decarboxylation of orotidine 5'-monophosphate via a vinyl carbanion intermediate) the Michaelis complex is E_C:S (Goryanova et al., 2013) and widespread conformational changes involving several loops are required to achieve the closed enzyme form (Miller et al., 2000). Therefore, the nature of the favored Michaelis complex does not appear to correlate with the magnitude of the conformational change required upon enzyme closure (Richard, 2019).

Phosphoryl transfer enzymes are another valuable model system with which to explore the relationship between the degree of conformational change during their catalytic cycles and the nature of the favoured Michaelis complex. These enzymes can achieve

catalytic rate constants of greater than 100 s^{-1}, even when the corresponding spontaneous non-catalysed rate constants are ~10^{-20} s^{-1} (Lad, Williams and Wolfenden, 2003). One of the most appropriate families for such investigations are the phosphomutases, which not only catalyse phosphoryl transfer between donor and acceptor groups in the proximal site, but also coordinate a non-reacting phosphodianion group of the substrate in the distal site. β-phosphoglucomutase (βPGM, EC 5.4.2.6, 25 kDa) from *Lactococcus lactis* is a well-characterised phosphomutase enzyme that catalyses the reversible isomerisation of β-glucose 1-phosphate (βG1P) to glucose 6-phosphate (G6P) via a β-glucose 1,6-bisphosphate intermediate (βG16BP) with a catalytic proficiency of 4×10^{26} M^{-1} (Figure 6.2.) (Lahiri et al., 2002a; Zhang et al., 2005; Dai et al., 2006; Baxter et al., 2006; Dai et al., 2009; Golicnik et al., 2009; Baxter et al., 2010; Griffin et al., 2012; Jin et al., 2014; Johnson et al., 2018; Wood et al., 2020; Lad, Williams and Wolfenden, 2003). Substrate-free βPGM exhibits an open conformation where the active site cleft located between the cap and core domains is exposed to bulk solvent (Zhang et al., 2005; Baxter et al., 2010; Wood et al., 2020). A cap domain rotation of 33–36º at the interdomain hinge leads to a closed transition state conformation (Zhang et al., 2005), as revealed by transition state analogue (TSA) complexes of βPGM with metallofluoride moieties and G6P (βPGM:AlF$_4$:G6P and βPGM:MgF$_3$:G6P TSA complexes) (Baxter et al., 2010; Jin et al., 2014; Johnson et al., 2018), which mimic the active site organisation of the phosphoryl transfer chemical step. The non-reacting phosphodianion group of G6P in the distal site is coordinated by the guanidinium group of residue R49 in an analogous arrangement to that present between the non-reacting phosphodianion group of the corresponding substrate and the cationic sidechains of residue R269 in GPDH, residue R235 in OMPDC and residue K12 in TIM (Reyes et al., 2015b). Additionally, a valuable property of βPGM is its amenability to analysis by a variety of NMR techniques and high-resolution X-ray crystallography (Zhang et al., 2005; Baxter et al., 2006; Baxter et al., 2008; Baxter et al., 2009; Baxter et al., 2010; Griffin et al., 2012; Jin et al., 2014; Jin, Molt and Blackburn, 2017; Johnson et al., 2018). Therefore, this suitability, together with the structural properties shared with OMPDC, GPDH and TIM allows βPGM to be used as a model system with which to tackle some of the remaining questions of how enzymes harvest the substrate binding energy to achieve high catalytic proficiency.
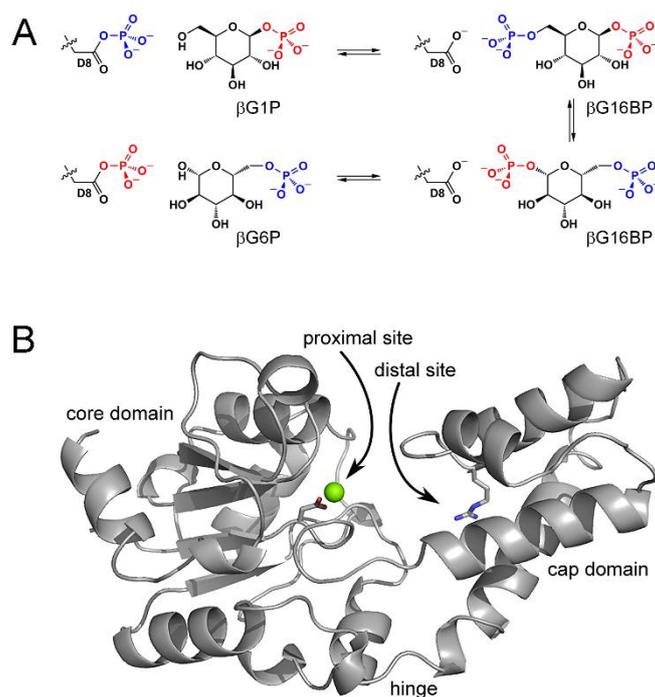
Figure 6.2 βPGM catalytic cycle and enzyme architecture. (A) βPGM catalytic cycle for the enzymatic conversion of βG1P to G6P via a βG16BP reaction intermediate. The phosphoryl transfer reaction between the phospho-enzyme (βPGM^P, phosphorylated at residue D8) and βG1P is illustrated with the transferring phosphate (blue) in the proximal site and the non-reacting 1-phosphate (red) of βG1P in the distal site. βG16BP is released to solution, which subsequently rebinds in the alternative orientation. Here, the phosphoryl transfer reaction between βPGM and βG16BP is shown with the transferring phosphate (red) in the proximal site and the non-reacting 6-phosphate (blue) of βG16BP in the distal site. G6P is released as product, together with the regeneration of βPGM^P. (B) Cartoon representation of the substrate-free βPGM$_{WT}$ crystal structure (PDB 6YDL) highlighting the architecture of the helical cap domain (T16–V87) and the α/β core domain (M1–D15, S88–K221). The proximal and distal phosphodianion binding sites are located in the cleft formed between the domains, and rotation at the hinge results in closure of the active site cleft during catalysis. Mg$_{cat}^{2+}$ (green sphere) is located in the proximal site adjacent to residue D8 (sticks) and residue R49 (sticks) in the distal site coordinates the non-reacting phosphodianion group of the substrate (or reaction intermediate).

Here we show, through combined use of site-directed mutagenesis, kinetic assays, NMR spectroscopy and X-ray crystallography that removal of the distal site cation-phosphodianion interaction in the R49K and R49A variants of βPGM (βPGM$_{R49K}$ and βPGM$_{R49A}$) reveals that the favoured Michaelis complex is E$_C$:S. This behaviour, together with the enormous catalytic proficiency reported, aligns βPGM with OMPDC, rather than with GPDH or TIM. Therefore, these results uncover a trend, whereby enzymes that catalyse intrinsically difficult chemistry utilise the substrate binding energy for stabilisation of E$_C$:S. Additionally in βPGM, such stabilisation by distal site phosphodianion group binding produces substrate inhibition, as demonstrated by the structural characterisation of a fully-closed non-productive βPGM:βG1P complex. Finally, NMR chemical shift comparisons of βPGM$_{R49K}$:AlF$_4$:G6P, βPGM$_{R49A}$:AlF$_4$:G6P, βPGM$_{R49K}$:MgF$_3$:G6P and βPGM$_{R49A}$:MgF$_3$:G6P TSA complexes, together with their βPGM$_{WT}$ counterparts, indicate that no significant communication is present between phosphodianion group binding in the distal site and the site of phosphoryl transfer. These observations justify the assumption of the model where distal site binding energy is not utilised to stabilise the transition state.

6.3 Methods

6.3.1 βPGM expression and purification

The βPGM_{R49K} and βPGM_{R49A} gene sequences were created by modifying the *pgmB* gene (encoding the βPGM_{WT} enzyme) from *Lactococcus lactis* (subspecies *lactis* IL1403) (NCBI: 1114041). The βPGM_{R49K} and βPGM_{R49A} genes were generated and cloned by GenScript into a pET22b(+) vector. The βPGM_{WT}, βPGM_{R49K}, βPGM_{R49A} and βPGM_{D170N} (Wood et al., 2021) plasmids were transformed into *Escherichia coli* BL21(DE3) cells and expressed in defined ^{15}N isotopically enriched M9 minimal media (Reed et al., 2003). Cells were grown at 37 °C with shaking until $OD_{600nm} = 0.6$, then cooled at 25 °C and induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for a further 18 h. Cells were harvested by centrifugation at 10,000 rpm for 10 min. The cell pellet was resuspended in ice-cold standard purification buffer (50 mM K^+ HEPES (pH 7.2), 5 mM $MgCl_2$, 2 mM NaN_3, 1 mM EDTA) supplemented with cOmplete™ protease inhibitor cocktail and lysed by 6 × 20 s cycles of sonication. The cell lysate was cleared by centrifugation at 20,000 rpm for 35 min at 4 °C. The soluble fraction was filtered using a 0.22 µm syringe filter and loaded onto a DEAE-Sepharose fast flow anion-exchange column connected to an ÄKTA Prime purification system, which had been washed previously with 1 M NaOH and 6 M guanidinium chloride and equilibrated with 5 column volumes of standard purification buffer. Bound proteins were eluted using a gradient of 0 to 50% standard purification buffer containing 1 M NaCl over 300 mL. Fractions containing βPGM were identified using SDS-PAGE and were concentrated to a 5–10 mL volume using centrifugation in a Vivaspin (10 kDa molecular weight cut off; Sartorius) at 4,500 rpm and 4 ⁰C (Thermo Scientific Heraeus Labofuge 400 R). The concentrated protein sample was loaded onto a pre-packed Hiload 26/600 Superdex 75 size-exclusion column connected to an ÄKTA Prime purification system, which had been washed previously with degassed 1 M NaOH and equilibrated with 3 column volumes of degassed standard purification buffer supplemented with 1 M NaCl. Following elution, the fractions containing βPGM were checked for purity and were pooled and buffer-exchanged into standard purification buffer. The final protein samples were concentrated using a Vivaspin (10 kDa molecular weight cut off; Sartorius) to a 1 mM concentration, as measured by Nanodrop One[C] (Thermo Scientific) (βPGM molecular weight = 24.2 kDa, extinction coefficient = 19,940 M^{-1} cm^{-1}), and were stored at –20 ⁰C.

6.3.2 Reagents

Unless stated otherwise, reagents were purchased from Sigma-Aldrich, GE Healthcare, Melford Laboratories or CortecNet. βG1P was prepared enzymatically from maltose using maltose phosphorylase (EC 2.4.1.8). A solution of 1 M maltose was incubated overnight with 1.5 U/mL maltose phosphorylase in 0.5 M sodium phosphate buffer (pH 7.0) at 30 °C. βG1P production was confirmed using ^{31}P NMR spectroscopy. Maltose phosphorylase (90 kDa) was removed from the solution by centrifugation using a Vivaspin (5 kDa molecular weight cut off; Sartorius) and the flow-through was used without further purification. Estimated concentrations of the components were 150 mM βG1P, 150 mM glucose, 850 mM maltose and 350 mM P_i.

6.3.3 Reaction kinetics monitored using ^{31}P NMR spectroscopy

Reaction kinetics for βPGM$_{WT}$, βPGM$_{R49K}$ and βPGM$_{R49A}$ were followed at 298 K using a Bruker 500 MHz Avance III HD spectrometer (operating at 202.48 MHz for ^{31}P) equipped with a 5-mm Prodigy BBO cryoprobe. One-dimensional ^{31}P NMR spectra recorded without proton decoupling were acquired within 1 min with 16 transients and a 2 s recycle delay to give signal-to-noise ratios for 10 mM βG1P of greater than 100:1. The equilibration of 10 mM βG1P with G6P by either 0.1 μM βPGM$_{WT}$, 0.5 μM βPGM$_{R49K}$ or 1 μM βPGM$_{R49A}$ was measured in standard kinetic buffer (200 mM K$^+$ HEPES (pH 7.2), 5 mM MgCl$_2$, 2 mM NaN$_3$) with the addition of 10% (v/v) ^{2}H$_2$O and 2 mM TSP. The reaction was initiated by and timed from the addition of 20 mM AcP and monitored by the acquisition of consecutive ^{31}P NMR experiments. Spectra were processed using TopSpin (Bruker) and normalised integral values of the G6P peak following baseline correction and 2 Hz Lorentzian apodisation were plotted against time to give kinetic profiles. The linear steady-state portion of the data was fitted using a linear least-squares fitting algorithm to derive the observed catalytic rate constant (k_{obs}).

6.3.4 NMR analysis of substrate-free βPGM

^{1}H^{15}N-TROSY spectra of substrate-free βPGM$_{WT}$, substrate-free βPGM$_{R49K}$ and substrate-free βPGM$_{R49A}$ were acquired at 298 K using a Bruker 500 MHz Avance III HD spectrometer equipped with a 5-mm QCI-F cryoprobe and z-axis gradients. Samples contained 1 mM ^{15}N-βPGM in standard NMR buffer (50 mM K$^+$ HEPES (pH 7.2), 5 mM MgCl$_2$, 2 mM NaN$_3$, with 10% (v/v) ^{2}H$_2$O and 2 mM trimethylsilyl propionate

(TSP)). Typically, $^1H^{15}N$-TROSY spectra were accumulations of 32 transients, with 256 increments and spectral widths of 32–36 ppm centred at 120 ppm in the indirect ^{15}N dimension. Experiments were processed using TopSpin (Bruker) and NMR figures were prepared using FELIX (Felix NMR, Inc.). ^1H chemical shifts were referenced relative to the internal TSP signal resonating at 0.0 ppm and ^{15}N chemical shifts were referenced indirectly using nuclei-specific gyromagnetic ratios.

6.3.5 NMR analysis of βPGM TSA complexes

$^1H^{15}N$-TROSY spectra of βPGM$_{WT}$:AlF$_4$:G6P, βPGM$_{R49K}$:AlF$_4$:G6P, βPGM$_{R49A}$:AlF$_4$:G6P, βPGM$_{WT}$:MgF$_3$:G6P, βPGM$_{R49K}$:MgF$_3$:G6P and βPGM$_{R49A}$:MgF$_3$:G6P TSA complexes were acquired at 298 K as described previously using a Bruker 500 MHz Avance III HD spectrometer equipped with a 5-mm QCI-F cryoprobe and z-axis gradients. Samples contained 1.5 mM ^{15}N-βPGM in standard NMR buffer (50 mM K$^+$ HEPES (pH 7.2), 5 mM MgCl$_2$, 2 mM NaN$_3$, with 10% (v/v) ^2H_2O and 2 mM TSP), together with 15 mM NaF, (3 mM AlCl$_3$) and 20 mM G6P. One-dimensional ^{19}F NMR spectra were acquired without proton decoupling and were processed with 10 Hz Lorentzian apodisation using TopSpin (Bruker). ^{19}F chemical shifts were referenced indirectly using nuclei-specific gyromagnetic ratios.

6.3.6 Measurement of apparent dissociation constant by ^1H NMR spectroscopy

The apparent dissociation constant (K_d) for G6P in the βPGM$_{R49K}$:AlF$_4$:G6P and βPGM$_{R49A}$:AlF$_4$:G6P TSA complexes was determined at 298 K using a Bruker Neo 800 MHz spectrometer equipped with a 5-mm TCI cryoprobe and z-axis gradients. A solution of 360–400 mM G6P in standard NMR buffer was titrated serially into separate solutions containing either 0.5 mM ^{15}N-βPGM$_{R49K}$ or 0.5 mM ^{15}N-βPGM$_{R49A}$ prepared in standard NMR buffer supplemented with 15 mM NaF and 3 mM AlCl$_3$. The titrations were monitored by the acquisition of one-dimensional ^1H NMR spectra and were processed using TopSpin (Bruker). The changing intensity of the well-resolved indole resonance of residue W24 (acting as a reporter for the closed TSA complex in slow exchange) was fitted to Equation 2.12 using a non-linear least-squares fitting algorithm corrected for dilution effects to determine K_d (G6P).

6.3.7 Reaction kinetics monitored by glucose 6-phosphate dehydrogenase coupled assay

Kinetic assays for βPGM$_{WT}$, βPGM$_{R49K}$ and βPGM$_{R49A}$ were conducted at 294 K using a FLUOstar OMEGA microplate reader and the BMG LABTECH Reader Control Software (version 5.11) (BMG Labtech) in standard kinetic buffer (200 mM K$^+$ HEPES (pH 7.2), 5 mM MgCl$_2$ and 1 mM NaN$_3$) in a 200 μl reaction volume. The rate of G6P production was measured indirectly using a glucose 6-phosphate dehydrogenase (G6PDH) coupled assay, in which G6P is oxidised and concomitant NAD$^+$ reduction is monitored by the increase in absorbance at 340 nm (NADH extinction coefficient = 6220 M^{-1} cm^{-1}). βPGM$_{WT}$, βPGM$_{R49K}$ and βPGM$_{R49A}$ concentrations were determined using a NanoDrop OneC spectrophotometer (Thermo Scientific) and diluted accordingly. Reactions were conducted in triplicate and were initiated by the addition of 20 mM AcP (10 mM AcP for βPGM$_{WT}$) to solutions containing 1 mM NAD$^+$ (0.5 mM NAD$^+$ for βPGM$_{WT}$) and 5 units mL^{-1} G6PDH, together with variable concentrations of βG1P (5, 15, 35, 50, 70, 100, 160, 230, 330 μM) and either 5 nM βPGM$_{WT}$, 60 nM βPGM$_{R49K}$ or 60 nM βPGM$_{R49A}$. The linear steady-state portion of G6P production was fitted using a linear least-squares fitting algorithm to determine the reaction velocity (v) at each βG1P concentration. Data were subsequently fitted to the standard Michaelis-Menten equation to derive k_{cat} and K_m values using an in-house python nonlinear least-squares fitting algorithm. Errors were estimated using a python bootstrap resampling protocol and are presented at one standard deviation.

6.3.8 Partitioning of contributions to transition-state stabilisation

Whilst a full Michaelis-Menten analysis, using an array of concentrations of both βG1P as the substrate and βG16BP as the phosphorylating agent, yielded a measurement of K_m (βG1P) and k_{cat} for βPGM$_{WT}$ (Wood et al., 2020), equivalent analyses are inaccessible experimentally for βPGM$_{R49K}$ and βPGM$_{R49A}$, due to the substantially weakened affinity of the distal site for phosphodianion groups. Therefore, apparent K_d (G6P) values for βPGM$_{WT}$ (apparent K_d (G6P) = 9 ± 1 μM), βPGM$_{R49K}$ (apparent K_d (G6P) = 3.0 ± 0.4 mM) and βPGM$_{R49A}$ (apparent K_d (G6P) = 18 ± 1 mM), together with k_{obs} values derived from ^{31}P NMR time course experiments for βPGM$_{WT}$ (k_{obs} = 70 ± 30 s^{-1}), βPGM$_{R49K}$ (k_{obs} = 14.8 ± 1 s^{-1}) and βPGM$_{R49A}$ (k_{obs} = 5.9 ± 0.5 s^{-1}), were used to estimate the cationic sidechain contribution to the intrinsic phosphodianion group binding energy. The free energy change in the stability of the Michaelis complex was

calculated as: $\Delta\Delta G_S = RT\cdot\ln$ (apparent K_d (βPGM_X) / apparent K_d (βPGM_{WT})) and the free energy change in the stability of the transition state was calculated as: $\Delta\Delta G^{\ddagger} = -RT\cdot\ln$ (k_{obs} (βPGM_X) / k_{obs}(βPGM_{WT})), where R is 1.987×10^{-3} kcal·mol^{-1}· K^{-1}, T = 298 K and $\beta PGM_X = \beta PGM_{R49K}$ or βPGM_{R49A}.

6.3.9 X-ray crystallography

Frozen aliquots of substrate-free βPGM in standard native buffer (50 mM K$^+$ HEPES (pH 7.2), 5 mM MgCl$_2$ and 1 mM NaN$_3$) were thawed on ice and centrifuged briefly to pellet insoluble material. Crystals of the βPGM:AlF$_4$:G6P TSA complexes were obtained from a solution of substrate-free βPGM containing 20 mM NaF, 5 mM AlCl$_3$, and 10 mM G6P. Crystals of the βPGM:MgF$_3$:G6P TSA complexes were obtained from a solution of substrate-free βPGM containing 20 mM NaF and 10 mM G6P. Crystals of the βPGM$_{D170N}$:βG1P complex were obtained from a solution of substrate-free βPGM$_{D170N}$ containing 20 mM NaF and 10 mM G6P. Crystals of the βPGM$_{WT}$:P$_i$ complex were obtained from a solution of substrate-free βPGM$_{WT}$ containing 10 mM glucose, 10 mM sodium phosphate and 15 mM NaF. Solutions were adjusted to a final protein concentration of 0.4–0.6 mM, incubated for >10 min and mixed 1:1 with precipitant (26–30% (w/v) PEG 4000, 200 mM sodium acetate and 100 mM tris-HCl (pH 7.5)). Crystals were grown at 290 K by hanging-drop vapour diffusion using a 2 µl drop suspended on a siliconised glass cover slip above a 700 µL well. Thin plate, small needle or rod shaped crystals grew typically over several days. Crystals were harvested using a mounted LithoLoop (Molecular Dimensions Ltd) and were cryo-protected in their mother liquor containing an additional 25% (v/v) ethylene glycol prior to plunging into liquid nitrogen. Diffraction data were collected at 100 K on the MX beamlines at the Diamond Light Source (DLS), Oxfordshire, United Kingdom and on beamline ID14-2 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. At the DLS, data were processed using the xia2 pipeline (Winter, 2010), whereas at the ESRF, data were processed with iMOSFLM (Battye et al., 2011). Resolution cut-offs were applied using either CC-half values or by consideration of the <I/σ(I)> and R$_{merge}$ values. Structures were determined by molecular replacement with MolRep (Vagin and Teplyakov, 1997) using previously deposited βPGM structures with the most appropriate cap and core domain relationship as a search models. Model building was carried out in COOT (Emsley et al., 2010) and either a restrained refinement with isotropic temperature factors (resolution worse than 1.5 Å) or

anisotropic temperature factors (resolutions better than 1.5 Å) was performed using REFMAC5 (Murshudov, Vagin and Dodson, 1997) in the CCP4i suite (Winn et al., 2011). Ligands were not included until the final stages of refinement to avoid biasing Fourier maps. Structure validation was carried out in COOT and MolProbity (Chen et al., 2010), superpositions were generated using PyMOL (The PyMOL Molecular Graphics System, version 1.8/2.2 Schrödinger, LLC), maps were generated using FFT (Read and Schierbeek, 1988) and domain movements were calculated using DynDom (Hayward and Berendsen, 1998).

6.4 Results

6.4.1 Structure of substrate-free βPGM$_{R49K}$ and substrate-free βPGM$_{R49A}$

Following previous work involving GPDH, OMPDC and TIM, the contribution of the cationic sidechain of R49 to the coordination of the substrate non-reacting phosphodianion group in the distal site of βPGM was studied using the βPGM$_{R49K}$ and βPGM$_{R49A}$ variants. The consequences of R49 sidechain substitution on the solution behaviour of substrate-free βPGM$_{R49K}$ and substrate-free βPGM$_{R49A}$ compared to substrate-free βPGM$_{WT}$ were examined using ^{1}H^{15}N-TROSY NMR experiments (Figure 7.1). In βPGM$_{R49K}$, the near-equivalence in backbone amide chemical shifts indicates that the substitution has only an impact in its immediate vicinity. In βPGM$_{R49A}$, small chemical shift perturbations limited to the cap domain reveal that the loss of the bulky cationic sidechain has a moderate effect on the helical packing arrangement. *Cis-trans* isomerisation of the K145–P146 peptide bond observed in βPGM$_{WT}$ (Wood et al., 2020) is also present for βPGM$_{R49K}$ and βPGM$_{R49A}$, and results in the population of two conformers in slow exchange (~70% conformer A and ~30% conformer B). These observations show that the R49 sidechain substitution has moderate but local consequences on the distal site behaviour of βPGM$_{R49K}$ and βPGM$_{R49A}$ compared to βPGM$_{WT}$. Additionally, substrate-free βPGM$_{R49K}$ and substrate-free βPGM$_{R49A}$ were crystallised and their structures were determined to 1.6 Å and 2.0 Å resolution, respectively (PDB 6HDH and PDB 6HDI; Table 7.1, Figure 7.1). Both structures overlay closely with previously deposited substrate-free βPGM$_{WT}$ structures (PDB 1ZOL (Zhang et al., 2005); PDB 2WHE (Baxter et al., 2010); Figure 7.2, Figure 7.3) and the catalytic magnesium ion (Mg$_{cat}^{2+}$) in the proximal site is coordinated analogously. Comparisons of the distal site show minimal structural perturbations to residues in the vicinity of the site of substitution in βPGM$_{R49K}$ and βPGM$_{R49A}$. The resolutions of the crystal structures are insufficient to reveal the subtle changes in helical packing of the cap domain that are observed in the solution behaviour of βPGM$_{R49A}$ (Figure 7.1). Both residue K49 in βPGM$_{R49K}$ and residue A49 in βPGM$_{R49A}$ occupy similar positions to the sidechain of residue R49 in βPGM$_{WT}$. Taken together, only a local impact is observed in the behaviour of the cap domain on R49 sidechain substitution in substrate-free βPGM$_{R49K}$ and substrate-free βPGM$_{R49A}$.

6.4.2 Structure of the βPGM$_{R49K}$ and βPGM$_{R49A}$ TSA complexes

Further investigations into the contribution of the cationic sidechain of R49 to the coordination of the substrate non-reacting phosphodianion group in a closed transition state conformation were conducted using TSA complexes of βPGM containing metallofluoride moieties and G6P. βPGM$_{WT}$, βPGM$_{R49K}$ and βPGM$_{R49A}$ were crystallised in complex with AIF$_4^-$ and G6P using conditions published previously (Baxter et al., 2006; Baxter et al., 2008; Baxter et al., 2010; Jin et al., 2014) and the structures of the resulting βPGM$_{WT}$:AIF$_4$:G6P, βPGM$_{R49K}$:AIF$_4$:G6P and βPGM$_{R49A}$:AIF$_4$:G6P TSA complexes were determined to 1.4 Å, 1.2 Å and 1.2 Å resolution, respectively (PDB 2WF6, PDB 6HDJ and PDB 6HDK; Table 7.1). When compared with the βPGM$_{WT}$:AIF$_4$:G6P TSA complex, the βPGM$_{R49K}$:AIF$_4$:G6P and βPGM$_{R49A}$:AIF$_4$:G6P TSA complexes show equivalent full domain closure, together with both near-identical domain conformations and proximal site coordination of the square-planar AIF$_4^-$ moiety (Figure 6.3 , Figure 7.2, Figure 7.3, Figure 7.4).

However, although the phosphodianion group of G6P is located in the same position in the distal site in each of the TSA complexes, its coordination differs between the βPGM$_{WT}$:AIF$_4$:G6P TSA complex and the βPGM$_{R49K}$:AIF$_4$:G6P and βPGM$_{R49A}$:AIF$_4$:G6P TSA complexes (Figure 6.3A–C). In the βPGM$_{WT}$:AIF$_4$:G6P TSA complex, the phosphodianion group is coordinated by the backbone amide group of K117 and the sidechains of S116 and N118, together with the guanidinium sidechain of residue R49 through two hydrogen bonds to separate 6-phosphate oxygen atoms of G6P. In the βPGM$_{R49K}$:AIF$_4$:G6P TSA complex, the alkylammonium sidechain of residue K49 is only able to hydrogen bond to one of these oxygen atoms, although the remaining coordination in the distal site is equivalent (Figure 6.3B). In the βPGM$_{R49A}$:AIF$_4$:G6P TSA complex, the A49 sidechain cannot substitute for either of the missing R49 sidechain hydrogen bonding interactions that coordinate the 6-phosphate oxygen atoms of G6P. Instead, the alkylammonium sidechain of residue K117 located in the core domain on the opposite face of the active site is recruited into the distal site from a solvent exposed position, thereby providing a surrogate hydrogen bonding interaction with the phosphodianion group (Figure 6.3C). A water molecule is also present that substitutes for the lost hydrogen bonding interaction between the R49 sidechain and the phosphodianion group in the βPGM$_{WT}$:AIF$_4$:G6P TSA complex. Hence, despite the local perturbation that the R49 sidechain substitution imposes on

the coordination of the distal site phosphodianion group of G6P, the repositioning of other sidechains located in the active site offers a degree of redundancy in hydrogen bonding interactions.
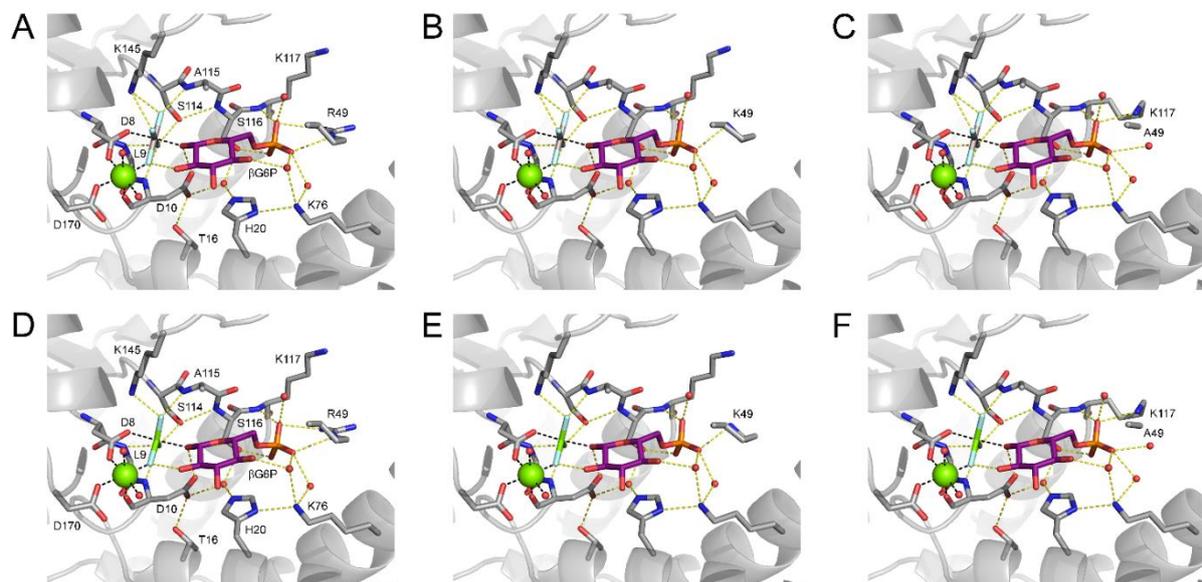
Figure 6.3 **Crystal structure comparisons of the βPGM:AlF₄:G6P and βPGM:MgF₃:G6P TSA complexes.** Active site details of (A) βPGM_{WT}:AlF₄:G6P complex (PDB 2WF6), (B) βPGM_{R49K}:AlF₄:G6P complex (PDB 6HDJ), (C) βPGM_{R49A}:AlF₄:G6P complex (PDB 6HDK), (D) βPGM_{WT}:MgF₃:G6P complex (PDB 2WF5), (E) βPGM_{R49K}:MgF₃:G6P complex (PDB 6HDL) and (F) βPGM_{R49A}:MgF₃:G6P complex (PDB 6HDM). Selected residues (sticks) and ligands (sticks) are illustrated, together with the square-planar AlF₄⁻ moiety (dark grey and light blue), the trigonal-planar MgF₃⁻ moiety (green and light blue), βG6P (purple carbon atoms), structural waters (red spheres) and Mg_{cat}^{2+} (green sphere). Yellow dashes indicate hydrogen bonds and black dashes show metal ion coordination. The sidechain of residue N118, which coordinates one of the 6-phosphate oxygen atoms of G6P equivalently in the TSA complexes, has been omitted for clarity.

Structural investigations were extended to include TSA complexes of βPGM_{R49K} and βPGM_{R49A} containing a trigonal-planar MgF₃⁻ moiety. MgF₃⁻ complexes are more expanded and less stable than their AlF₄⁻ counterparts, owing to the instability of MgF₃⁻ in solution (Baxter. et al., 2006). However, the trigonal-planar MgF₃⁻ moiety is near-isosteric and isoelectronic with PO₃⁻ in the proximal site and therefore is a closer mimic of the transition state for the chemical step (Baxter et al., 2008; Baxter et al., 2010; Jin et al., 2016). βPGM_{R49K} and βPGM_{R49A} were crystallised in complex with MgF₃⁻ and G6P using conditions published previously (Baxter et al., 2010; Jin et al., 2014; Wood et al., 2020) and the structures of the resulting βPGM_{R49K}:MgF₃:G6P and βPGM_{R49A}:MgF₃:G6P TSA complexes were determined to 1.2 Å and 1.3 Å resolution,

respectively (PDB 6HDL and PDB 6HDM; Table 7.1). When compared with the βPGM$_{WT}$:MgF$_3$:G6P TSA complex (1.3 Å resolution; PDB 2WF5; Baxter et al., 2010), the fully closed βPGM$_{R49K}$:MgF$_3$:G6P and βPGM$_{R49A}$:MgF$_3$:G6P TSA complexes show a near-identical correspondence in domain conformation and proximal site coordination of the trigonal-planar MgF$_3^-$ moiety (Figure 6.3D–F, Figure 7.2, Figure 7.3, Figure 7.4). Additionally, in the absence of the guanidinium sidechain of residue R49, equivalent recruitment of the alkylammonium sidechain of residue K49 occurs in both the βPGM$_{R49K}$:AlF$_4$:G6P and βPGM$_{R49K}$:MgF$_3$:G6P TSA complexes. Furthermore, the redundancy present within the distal site allows identical recruitment of the alkylammonium sidechain of residue K117 in the βPGM$_{R49A}$:AlF$_4$:G6P and βPGM$_{R49A}$:MgF$_3$:G6P TSA complexes to coordinate the phosphodianion group of G6P (Figure 6.3, Figure 7.4).

6.4.3 Solution behaviour of the βPGM$_{R49K}$ and βPGM$_{R49A}$ TSA complexes

The βPGM$_{R49K}$:AlF$_4$:G6P, βPGM$_{R49A}$:AlF$_4$:G6P, βPGM$_{R49K}$:MgF$_3$:G6P and βPGM$_{R49A}$:MgF$_3$:G6P TSA complexes were further investigated using NMR spectroscopy to examine their solution properties. The TSA complexes readily self-assemble in solution from mixtures containing 0.5 mM βPGM, 5 mM MgCl$_2$, 15 mM NaF, (3 mM AlCl$_3$) and 20 mM G6P in K$^+$ HEPES buffer (pH 7.2). Since AlF$_4^-$ is populated free in solution from its constituent ions (Bodor et al., 2000), a βPGM:AlF$_4$ complex readily forms in the absence of G6P, which represents a TSA of phospho-enzyme (βPGM^P, phosphorylated at residue D8) hydrolysis (Baxter et al., 2010; Wang et al., 2002; Baxter et al., 2008). Therefore, the apparent dissociation constant (K_d) of G6P was determined by titration into separate βPGM$_{R49K}$:AlF$_4$ and βPGM$_{R49A}$:AlF$_4$ complexes and the formation of the βPGM$_{R49K}$:AlF$_4$:G6P and βPGM$_{R49A}$:AlF$_4$:G6P TSA complexes was monitored using one-dimensional ^{1}H NMR spectra. The changing intensity of the well-resolved indole resonance of residue W24 (acting as a reporter for the closed TSA complex in slow exchange) was fitted to determine the apparent K_d (G6P) for the βPGM$_{R49K}$:AlF$_4$:G6P TSA complex (3.0 ± 0.4 mM) and the βPGM$_{R49A}$:AlF$_4$:G6P TSA complex (18 ± 1 mM) (Figure 7.5). For the βPGM$_{WT}$:AlF$_4$:G6P TSA complex, an apparent K_d (G6P) = 9 ± 1 µM was determined using isothermal titration calorimetry (Baxter et al., 2010). However, an equivalent approach to determine the apparent K_d (G6P) in the βPGM$_{R49K}$:MgF$_3$:G6P and βPGM$_{R49A}$:MgF$_3$:G6P TSA complexes is more complicated to interpret as the

formation constant for MgF_3^- is very low and $\beta PGM_{R49K}:MgF_3$ and $\beta PGM_{R49A}:MgF_3$ complexes are not observed (Baxter et al., 2008). The corresponding increase in apparent K_d (G6P) of 330-fold and 2000-fold for the $\beta PGM_{R49K}:AlF_4:G6P$ and $\beta PGM_{R49A}:AlF_4:G6P$ TSA complexes, respectively, indicate that the coordination redundancy operating in each distal site is far from optimal, and impairs TSA complex formation.

Any communication relayed between distal site phosphodianion group coordination and the proximal site within the closed TSA complexes should be reflected in ^1H and ^{15}N chemical shift perturbations of protein NMR resonances. Structural modifications arising from a single amino acid substitution generally result in weighted chemical shift changes of 1–2 ppm for backbone amide groups within 5 Å, as the local electronic environment is perturbed (Baxter et al., 2017; Cruz-Navarrete et al., 2019). Significantly larger weighted chemical shift changes report on more pronounced alterations in protein conformation (Wood et al., 2020). Additionally, ^{19}F chemical shifts are strongly perturbed by the electronic environment in the vicinity of the fluorine nuclei and therefore the presence of metallofluoride moieties in the proximal site provides a highly sensitive measurement of the extent of inter-site communication in the TSA complexes. For example, ^{19}F chemical shift changes of less than 1.7 ppm are observed for the fluorine nuclei when comparing $\beta PGM_{WT}:MgF_3:G6P$ and $\beta PGM_{WT}:MgF_3:$glucose 6-phosphonate TSA complexes, where the methylene group of the non-hydrolysable G6P analogue results in changes to the electrostatic distribution within the distal site (Baxter et al., 2010). In contrast, substantially larger ^{19}F chemical shift changes (1.7–18.1 ppm) are observed when the coordination of the MgF_3^- moiety is substantially perturbed, for example when G6P is substituted by a non-native hexose monophosphate (2-deoxyG6P or α-galactose 1-phosphate (αGal1P)) in the corresponding TSA complexes (Baxter et al., 2009).

One-dimensional ^{19}F spectra of the $\beta PGM_{R49K}:AlF_4:G6P$ and $\beta PGM_{R49A}:AlF_4:G6P$ TSA complexes revealed four protein bound ^{19}F resonances, which were readily assigned according to their chemical shift ranges and their solvent induced isotope shifts (Baxter et al., 2008; Baxter et al., 2010; Jin et al., 2014) (Figure 6.4A, C, Table 6.1). When compared with the $\beta PGM_{WT}:AlF_4:G6P$ TSA complex, the observed changes in chemical shift of the $\beta PGM_{R49K}:AlF_4:G6P$ and $\beta PGM_{R49A}:AlF_4:G6P$ TSA complexes showed a slight chemical shift change to a lower frequency (upfield shift) for F2 (–0.3

ppm and –0.8 ppm, respectively) and F3 (–0.5 ppm and –0.7 ppm, respectively), a slight shift to a higher frequency (downfield shift) for F1 (+0.1 ppm and +0.4 ppm, respectively), and no change for F4. Equivalent spectra for the βPGM_{R49K}:MgF_3:G6P and $\beta PGM_{R49A}MgF_3$:G6P TSA complexes showed three protein bound ^{19}F resonances, which were readily assigned using the βPGM_{WT}:MgF_3:G6P TSA complex (Baxter et al., 2008; Baxter et al., 2010; Jin et al., 2014) (Figure 6.4B, D, Table 6.1). Comparisons of ^{19}F frequencies revealed a similar shift to a lower frequency for F2 (–0.4 ppm and –1.3 ppm, respectively), whereas F3 (–0.2 ppm and +0.1 ppm, respectively) and F1 (–0.2 ppm and +0.3 ppm, respectively) showed different minor responses for the two variants. Notably, all of the observed chemical shift changes are small (< 1.7 ppm). It is likely that these result from subtle modifications in the chemical environment of the fluorine nuclei (Figure 6.3, Figure 6.4A, B) due to small differences in the positioning of G6P and proximal site residues when the coordination of the phosphodianion group in the distal site is perturbed.

Table 6.1 **[19]F chemical shifts (ppm) observed for AlF_4^- and MgF_3^- moieties present in the βPGM:AlF_4:G6P and βPGM:MgF_3:G6P TSA complexes.**

TSA complex	F1	F2	F3	F4
βPGM_{WT}:AlF_4:G6P	–144.0	–137.0	–130.6	–140.7
βPGM_{R49K}:AlF_4:G6P	–143.9	–137.3	–131.1	–140.8
βPGM_{R49A}:AlF_4:G6P	–143.6	–137.8	–131.3	–140.7
βPGM_{WT}:MgF_3:G6P	–159.0	–147.0	–151.9	–
βPGM_{R49K}:MgF_3:G6P	–159.2	–147.4	–152.1	–
βPGM_{R49A}:MgF_3:G6P	–158.7	–148.3	–151.8	–
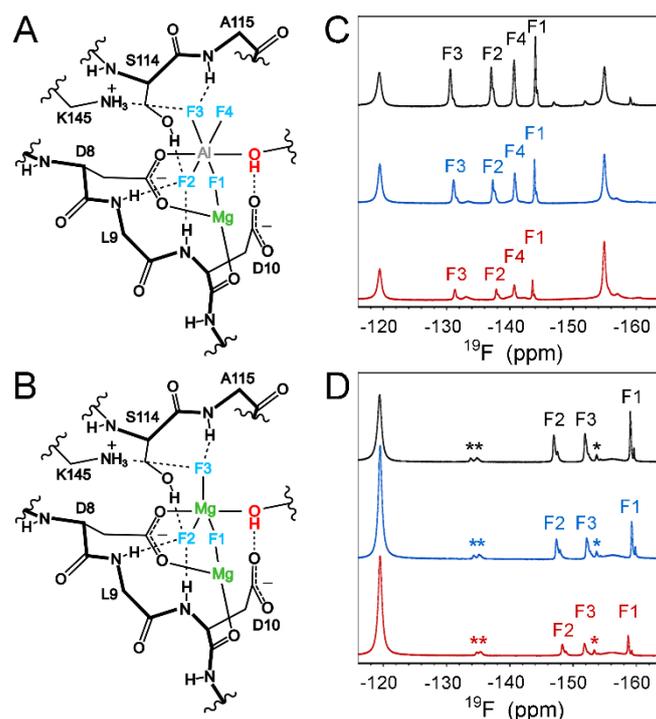
Figure 6.4 Active site coordination and ^{19}F NMR spectra of the AlF_4^- and MgF_3^- moieties present in the βPGM:AlF_4:G6P and βPGM:MgF_3:G6P TSA complexes. (A–B) Schematic representation of (A) the square-planar AlF_4^- moiety within the βPGM:AlF_4:G6P TSA complexes and (B) the trigonal-planar MgF_3^- moiety within the βPGM:MgF_3:G6P TSA complexes, showing coordination by proximal site residues, the 1-hydroxyl group of βG6P and Mg_{cat}^{2+}. Fluorine atoms have been labelled in accordance with IUPAC recommendations. (C–D) ^{19}F NMR spectra for (C) βPGM$_{WT}$:AlF_4:G6P complex (black), βPGM$_{R49K}$:AlF_4:G6P complex (blue) and βPGM$_{R49A}$:AlF_4:G6P complex (red), and (D) βPGM$_{WT}$:MgF_3:G6P complex (black), βPGM$_{R49K}$:MgF_3:G6P complex (blue) and βPGM$_{R49A}$:MgF_3:G6P complex (red), acquired in standard NMR buffer containing 1 mM βPGM, 15 mM NaF, (3 mM $AlCl_3$) and 20 mM G6P. Fluorine resonances corresponding to the AlF_4^- and MgF_3^- moieties have been labelled accordingly. Small shoulders situated upfield (right) of the main resonances result from primary solvent induced isotope shifts arising from 10% v/v ^2H_2O present in the samples. Resonances indicated by asterisks correspond to an alternative conformation of the βPGM:MgF_3:G6P TSA complexes. Free F^- resonates at –119 ppm and free AlF_x species resonate at –155 ppm.

Additionally, backbone amide group chemical shift assignments were determined for the βPGM$_{R49K}$:AlF_4:G6P, βPGM$_{R49A}$:AlF_4:G6P, βPGM$_{R49K}$:MgF_3:G6P and βPGM$_{R49A}$:MgF_3:G6P TSA complexes by comparison with their βPGM$_{WT}$ TSA counterparts. Pairwise weighted chemical shift perturbations relative to the βPGM$_{WT}$:AlF_4:G6P and βPGM$_{WT}$:MgF_3:G6P TSA complexes display broadly similar distributions across the four comparisons (Figure 6.5). Residues that comprise the two interdomain hinges (D15–T16 and V87–S88) show only small chemical shift changes indicating that the degree of domain closure is consistent. The magnitude of the

chemical shift changes that occur in the substrate specificity loop (K45–S52) (Lahiri et al., 2004) and a cap domain α-helix (A73–N78) mirror those observed in substrate-free βPGM$_{R49K}$ and βPGM$_{R49A}$ (Figure 7.1). In the closed TSA complexes, further propagation of the effect of R49 sidechain substitution is reflected in the D137–P148 loop due to the close proximity of the cap and core domains. Additionally, small (0.1–0.4 ppm) chemical shift changes are observed in the S114–N118 loop interconnecting the proximal and distal sites (Figure 6.3; Figure 6.5). Residues S114 and A115 coordinate the AlF$_4^-$ and MgF$_3^-$ moieties, residue S116 forms key hydrogen bond interactions with both S114 and one of the 6-phosphate oxygen atoms of G6P and an additional coordination of the phosphodianion group in the distal site is mediated by residues K117 and N118 (Figure 6.3A, D). In particular, the local effects of differential coordination of the phosphodianion group on R49 sidechain substitution is evident through the behaviour of the backbone amide group of K117, owing to its hydrogen bond with one of the 6-phosphate oxygen atoms of G6P (Figure 6.3; Figure 6.5E, F). In the βPGM:MgF$_3$:G6P TSA complexes, there is some further propagation through the MgF$_3^-$ moiety to the backbone amide groups of L9 and D10, together with residues coordinating Mg$_{cat}^{2+}$ (Figure 6.3; Figure 6.5C, D). Taken together, the small magnitude of all of these changes in ^{1}H, ^{15}N and ^{19}F chemical shifts indicate that no substantial communication exists between the coordination of the non-reacting phosphodianion group in the distal site and the site of phosphoryl transfer in a fully-closed, near-transition state conformation.
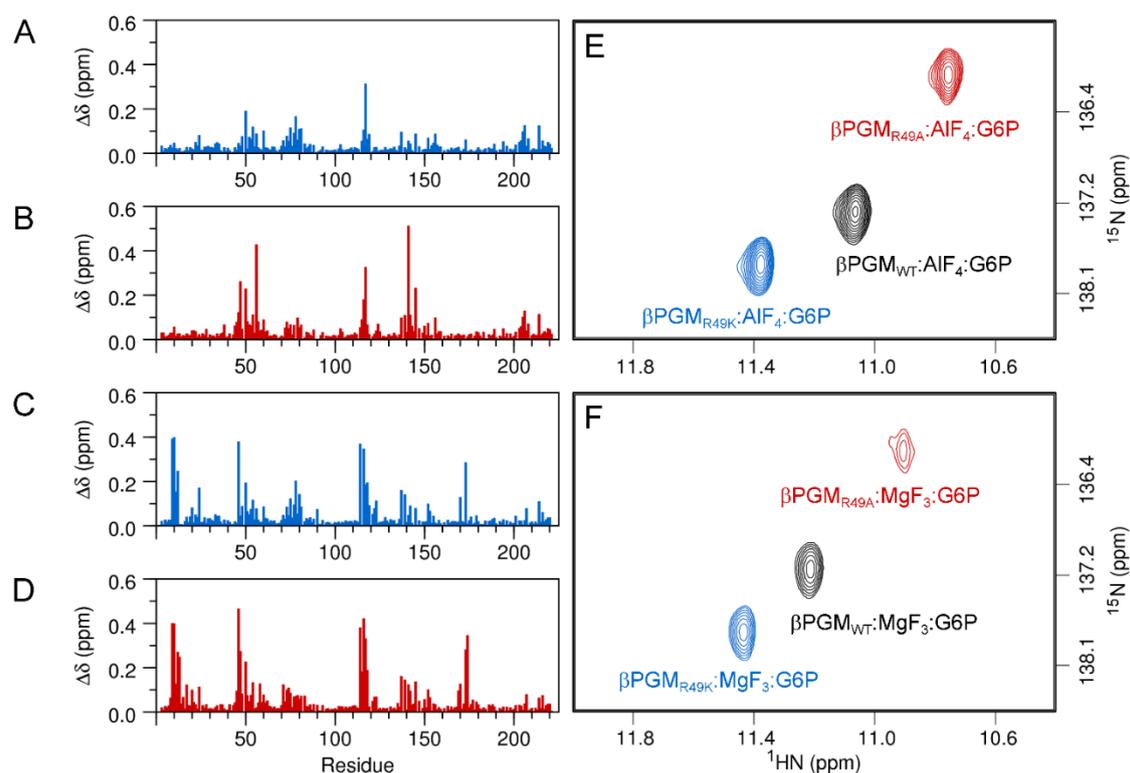
Figure 6.5 Residue specific backbone amide group chemical shift perturbations involving the βPGM:AlF$_4$:G6P and βPGM:MgF$_3$:G6P TSA complexes. (A–D) Histograms of weighted chemical shift differences, calculated as $\Delta\delta = [(\delta_{HN-X} - \delta_{HN-Y})^2 + (0.13\times(\delta_{N-X} - \delta_{N-Y}))^2]^{1/2}$, where X and Y are the two species being compared. (A) Chemical shift differences between βPGM$_{R49K}$:AlF$_4$:G6P complex and βPGM$_{WT}$:AlF$_4$:G6P complex. (B) Chemical shift differences between βPGM$_{R49A}$:AlF$_4$:G6P complex and βPGM$_{WT}$:AlF$_4$:G6P complex. (C) Chemical shift differences between βPGM$_{R49K}$:MgF$_3$:G6P complex and βPGM$_{WT}$:MgF$_3$:G6P complex. (D) Chemical shift differences between βPGM$_{R49A}$:MgF$_3$:G6P complex and βPGM$_{WT}$:MgF$_3$:G6P complex. (E–F) Overlays of a section of ^{1}H^{15}N-TROSY spectra for the βPGM:AlF$_4$:G6P and βPGM:MgF$_3$:G6P TSA complexes highlighting the behaviour of residue K117. (E) βPGM$_{WT}$:AlF$_4$:G6P complex (black), βPGM$_{R49K}$:AlF$_4$:G6P complex (blue) and βPGM$_{R49A}$:AlF$_4$:G6P complex (red). (F) βPGM$_{WT}$:MgF$_3$:G6P complex (black), βPGM$_{R49K}$:MgF$_3$:G6P complex (blue) and βPGM$_{R49A}$:MgF$_3$:G6P complex (red). Samples contained 1.5 mM ^{15}N-βPGM in 50 mM K$^+$ HEPES buffer (pH 7.2), 5 mM MgCl$_2$, 2 mM NaN$_3$, 10% (v/v) ^{2}H$_2$O, 2 mM TSP, together with 15 mM NaF, (3 mM AlCl$_3$) and 20 mM G6P.

6.4.4 Kinetic activity of βPGM$_{R49K}$ and βPGM$_{R49A}$

The consequences of R49 sidechain substitution on enzyme activity were investigated further using kinetic assays. The βPGM$_{R49K}$ and βPGM$_{R49A}$ catalysed conversion of 10 mM βG1P to G6P was monitored using one-dimensional ^{31}P NMR time course experiments with the addition of 20 mM acetyl phosphate (AcP) as a phosphorylating agent to initiate the reaction. Normalised G6P peak integral values were plotted as a function of time and resulted in similar kinetic profiles, which are broadly comparable to the βPGM$_{WT}$ time course (Figure 6.6A). Subsequent fitting of the steady-state linear segments of the kinetic profiles yielded observed catalytic rate constants for βPGM$_{R49K}$ (k_{obs} = 14.8 ± 1 s^{-1}) and βPGM$_{R49A}$ (k_{obs} = 5.9 ± 0.5 s^{-1}). These values represent a 5-fold and 12-fold reduction in catalytic rate constant compared to βPGM$_{WT}$ (k_{obs} = 70 ± 30 s^{-1}; Johnson et al., 2018). The trend in the reduced observed catalytic rate constants for βPGM$_{R49K}$ and βPGM$_{R49A}$ is consistent with the increases in the apparent K_d (G6P) values for the βPGM$_{R49K}$:AlF$_4$:G6P and βPGM$_{R49A}$:AlF$_4$:G6P TSA complexes, implying that enzyme catalytic activity is significantly affected by differential coordination of the distal site phosphodianion group.

A previously reported kinetic characterisation of βPGM$_{WT}$ activity identified the presence of a lag phase prior to the attainment of steady-state kinetic behaviour (Golicnik et al., 2009), which is the result from two independent kinetic components. The first component is an allomorphic effect operating over a short timeframe (< 5 min), where the full rate of catalysis is retarded until the concentration of the βG16BP intermediate is sufficiently elevated to phosphorylate βPGM$_{WT}$ efficiently (Wood et al., 2020). The second component is due to substrate inhibition operating over a longer timeframe (5–15 min), where βG1P associates with substrate-free βPGM$_{WT}$ (K_i = 1510 ± 100 μM) in a non-productive binding mode (Wood et al., 2020). For βPGM$_{R49K}$ and βPGM$_{R49A}$, the allomorphic component of the lag phase persists in the early parts of the kinetic profiles, while differences in the βG1P inhibition component is more difficult to distinguish as the observed catalytic rate constants are smaller. However, the ^{31}P NMR spectra acquired to monitor βPGM$_{R49K}$ and βPGM$_{R49A}$ catalysis show the presence of the βG16BP intermediate building to measurable concentrations in the reaction sample, whereas equivalent experiments recorded using βPGM$_{WT}$ indicate that the steady-state concentration of βG16BP generated is too low to be detected (Figure 6.5B–D). These observations demonstrate that turnover of the reaction

intermediate is compromised by the perturbation of the distal site phosphodianion group coordination.
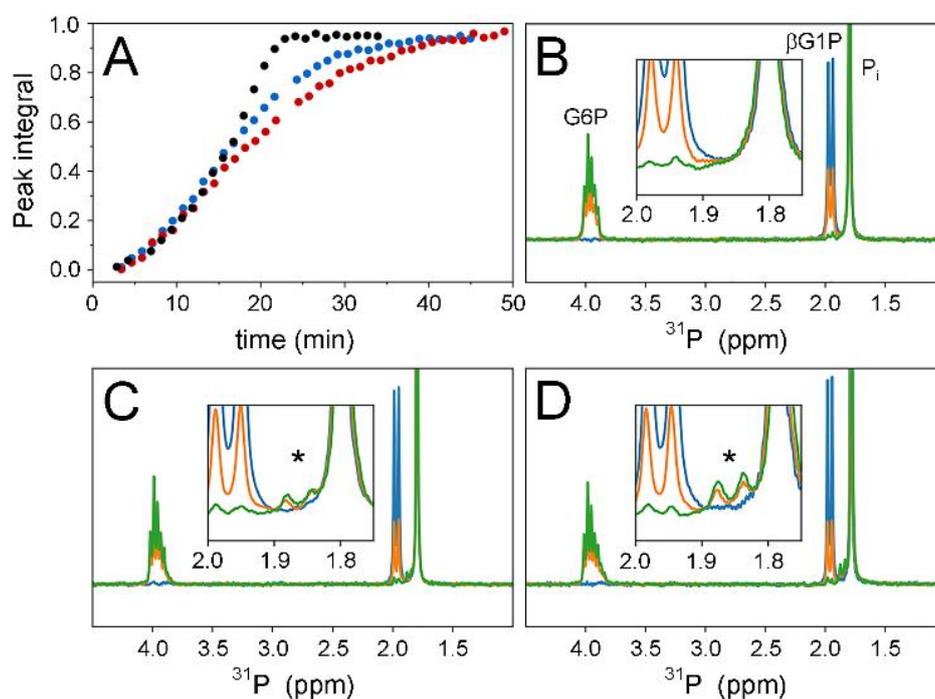
Figure 6.6 Activity of βPGM_WT, βPGM_R49K and βPGM_R49A monitored using ^{31}P NMR spectroscopy. (A) Reaction kinetics for the equilibration of 10 mM βG1P with G6P in standard kinetic buffer catalysed by 0.25 µM βPGM_WT (black circles), 0.5 µM βPGM_R49K (blue circles) and 1.0 µM βPGM_R49A (red circles). The reactions were initiated by and timed from the addition of 20 mM AcP. Normalised integral values for the G6P peak are plotted as a function of time. (B–D) Overlays of ^{31}P NMR spectra from the beginning (blue), midpoint (orange) and end (green) of the kinetic profiles for (B) βPGM_WT, (C) βPGM_R49K and (D) βPGM_R49A. Corresponding changes in βG1P and G6P peak intensities are observed as the reactions progress. Inlays highlight the formation of up to ~1 mM βG16BP reaction intermediate (black asterisks; 1-phosphate doublet of βG16BP) during the course of the reactions catalysed by βPGM_R49K and βPGM_R49A, whereas for the βPGM_WT reaction, βG16BP accumulation is not observed.

Further kinetic experiments were conducted for βPGM_R49K and βPGM_R49A to investigate the dependence of the steady-state reaction velocity on βG1P concentration. Here, a glucose 6-phosphate dehydrogenase coupled assay was used to monitor the conversion of βG1P to G6P with AcP present as the phosphorylating agent (Johnson et al., 2018). As for βPGM_WT, the kinetic profiles for βPGM_R49K and βPGM_R49A display an initial allomorphic lag phase (Wood et al., 2020), whereas the βG1P inhibition component acting over longer timeframes prior to steady-state kinetics is much less prominent (Figure 7.6A–C). Unfortunately, the weak substrate affinity of both βPGM_R49K and βPGM_R49A prevented the determination of reliable kinetic

parameters over the experimentally accessible βG1P concentration range (Figure 7.6D–F). However, a linear fit to the initial data points of each Michaelis-Menten plot allowed the k_{cat}/K_m ratio to be derived for βPGM$_{WT}$ (k_{cat}/K_m = 0.29 s^{-1}·μM^{-1}), βPGM$_{R49K}$ (k_{cat}/K_m = 0.05 s^{-1}·μM^{-1}) and βPGM$_{R49A}$ (k_{cat}/K_m = 0.02 s^{-1}·μM^{-1}) (Figure 7.6D–F). These values represent a 6-fold and 15-fold reduction in the k_{cat}/K_m ratio compared to βPGM$_{WT}$ under the same conditions, which mirrors the reduction in observed catalytic rate constant determined using ^{31}P NMR time course experiments. Taken together, the kinetics results obtained from the ^{31}P NMR time course experiments and the coupled assays indicate that perturbation of the distal site phosphodianion group coordination manifests in reduced observed catalytic rate constants (k_{obs}), together with impaired binding of βG16BP and βG1P in a productive and a non-productive binding mode (βG1P inhibition).

6.4.5 βPGM$_{D170N}$ binds βG1P in a closed non-productive complex

To further examine the contribution of the distal site phosphodianion group coordination to the observed βG1P inhibition component of the lag phase, crystallisation trials of the inhibited βPGM:βG1P complex were attempted. Since βG1P readily equilibrates with G6P in solution in the presence of βPGM$_{WT}$, the non-hydrolysable β-glucose 1-fluorophosphonate mimic (Jin et al., 2014) was used in co-crystallisation experiments, but all trials were unsuccessful. Therefore, the partially inactivated D170N variant (βPGM$_{D170N}$) (Wood et al., 2021) was used, where perturbation of the Mg$_{cat}^{2+}$ site was achieved through an anionic to neutral sidechain substitution (Figure 6.3). βPGM$_{D170N}$ showed a significantly reduced catalytic activity, a decrease in Mg$_{cat}^{2+}$ affinity, together with an increase in βG1P affinity and a similar level of βG1P inhibition (k_{obs} = 3.0 × 10^{-3} s^{-1}, apparent K_m (Mg^{2+}) = 690 ± 110 μM, apparent K_m (βG1P) = 6.9 ± 1.0 μM and apparent K_i (βG1P) = 1540 ± 170 μM (Wood et al., 2021) compared to βPGM$_{WT}$ (k_{cat} = 382 ± 12 s^{-1}, K_m (Mg^{2+}) = 180 ± 40 μM, K_m (βG1P) = 91 ± 4 μM and K_i (βG1P) = 1510 ± 100 μM (Wood et al., 2020)). These kinetic parameters indicate that the sidechain substitution in βPGM$_{D170N}$ primarily perturbs Mg$_{cat}^{2+}$ binding in the proximal site resulting in a reduction in catalytic activity. However, coordination of βG1P, both in a productive and non-productive binding mode, are only modestly affected. Comparison of ^{1}H^{15}N-TROSY NMR spectra indicated that substrate-free βPGM$_{D170N}$ has similar solution properties and an overall protein fold to substrate-free βPGM$_{WT}$, including the slow-exchange behaviour that

arises from *cis-trans* proline isomerisation at the K145–P146 peptide bond (Wood et al., 2020). Substrate-free βPGM_{D170N} was crystallised and the structure was determined to 1.4 Å resolution (PDB 6HDF; Table 7.1, Figure 6.7). The open domain arrangement closely resembles other substrate-free βPGM structures (Figure 7.2, Figure 7.3). However, in both monomers of the asymmetric unit, a Na^+ ion is poorly coordinated instead of Mg_{cat}^{2+} in the proximal site via the sidechain carboxamide group of N170. Therefore, βPGM_{D170N} is a suitable candidate for crystallisation trials of the inhibited βPGM:βG1P complex.
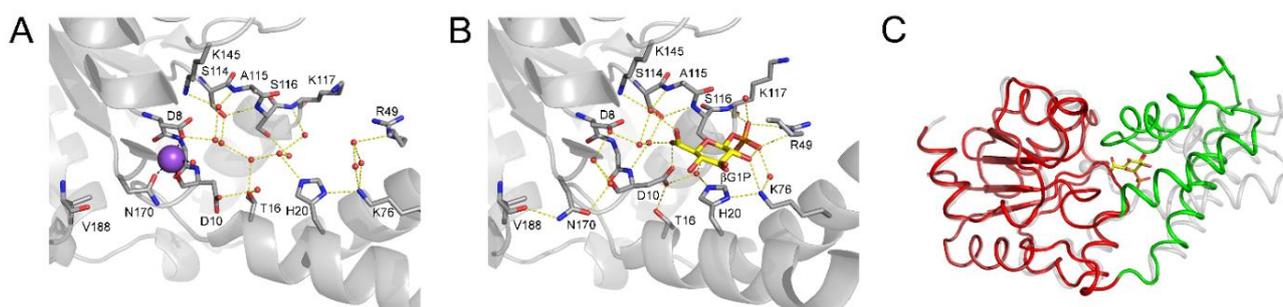
Figure 6.7 Crystal structure comparisons of substrate-free βPGM_{D170N} and the βPGM_{D170N}:βG1P complex. (A) Active site details of substrate-free βPGM_{D170N} (PDB 6HDF), with selected residues (sticks) and structural waters (red spheres) shown and a Na^+ atom (purple sphere) occupying the Mg_{cat}^{2+} site. (B) Active site details of the βPGM_{D170N}:βG1P complex (PDB 6HDG), with selected residues (sticks), structural waters (red spheres) and βG1P (gold carbon atoms) illustrated. The 6-phosphate group of βG1P in the proximal site has two arrangements resolved for the C5–C6 bond. Yellow dashes indicate hydrogen bonds and black dashes show metal ion coordination. (C) Superposition of substrate-free βPGM_{D170N} (PDB 6HDF) and the βPGM_{D170N}:βG1P complex (PDB 6HDG) on the core domain showing the extent of domain closure. The protein backbone of the βPGM_{D170N}:βG1P complex is depicted as a ribbon, with the core (red) and cap (green) domains indicated and βG1P shown as sticks (gold carbon atoms). The protein backbone of substrate-free βPGM_{D170N} is displayed as a pale grey ribbon, with the Na^+ atom shown as a grey sphere. The sidechain of residue N118, which coordinates one of the 1-phosphate oxygen atoms of βG1P, has been omitted for clarity.

Initially, crystallisation trials involving βPGM_{D170N} along with MgF_3^- and G6P were prepared to obtain a structure of the βPGM_{D170N}:MgF_3:G6P TSA complex. The resulting structure was determined to 1.2 Å resolution (PDB 6HDG; Table 7.1, Figure 6.7B, Figure 7.7A), but further analysis revealed a βPGM_{D170N}:βG1P complex. The presence of βG1P in the crystallisation buffer is a result of βPGM_{D170N} activity (Wood et al., 2021), which is a process that has been reported previously for βPGM_{WT} in crystallisation experiments (Griffin et al., 2012; Lahiri et al., 2003). The trigonal-planar MgF_3^- moiety mimicking the transferring phosphoryl group in the proximal site was

absent, which is a surprising observation given the presence of 20 mM NaF in the crystallisation buffer and the high affinity of βPGM for metallofluoride moieties as monitored by ^{19}F NMR spectroscopy (Baxter et al., 2010; Jin et al., 2014; Jin, Molt and Blackburn, 2017). Inspection of the electron density map indicates that neither a Mg^{2+} ion nor a Na^+ ion is coordinated in the Mg_{cat}^{2+} site, despite the inclusion of 5 mM Mg^{2+} and ~200 mM Na^+ ions in the crystallisation buffer. Instead, the sidechain of N170 is rotated 103º about χ_1 such that the carboxamide group forms a hydrogen bond with the backbone carbonyl group of V188, rather than being positioned to coordinate a cation in the Mg_{cat}^{2+} site (Figure 6.7B, Figure 7.7A). The 1-phosphate group of βG1P is coordinated in the distal site by the backbone amide group of K117, the sidechain hydroxyl group of S116, the sidechain carboxamide group of N118 and the guanidinium group of R49, in an analogous arrangement to that present in the $βPGM_{WT}:MgF_3:βG1CP$ TSA complex (PDB 4C4R; Jin et al., 2014). Also, a comparable extensive hydrogen bond network involving residues of the active site, coordinates three hexose ring hydroxyl groups of βG1P directly, rather than being mediated by water molecules as observed in equivalent $βPGM:MgF_3:G6P$ TSA complexes (Jin et al., 2014). In the proximal site, the 6-hydroxyl group of βG1P has two arrangements resolved for the C5–C6 bond, which differ in their rotation by 140º. This arrangement facilitates hydrogen bonding separately with two of the three water molecules that now occupy the location of the missing trigonal-planar MgF_3^- moiety (Figure 6.7B, Figure 7.7B). Furthermore, such proximity of the C6–O6 bond of βG1P to the site of phosphoryl transfer allows alignment with the Oδ1 carboxylate atom of residue D8 (nucleophile) and engagement of residue D10 (general acid-base) in the active site, along with coordination of residue T16 in a manner associated with full domain closure (Johnson et al., 2018). Therefore, this structure represents a ground state complex with near-transition state architecture (Figure 7.2, Figure 7.3; Figure 6.7C), which serves as an excellent model for an inhibited $βPGM_{WT}:βG1P$ complex. The population of such a stable complex is consistent with the βG1P inhibition component of the lag phase observed in kinetic experiments.

6.4.6 $βPGM_{WT}$ coordinates a phosphate anion in the distal site

An approach to obtain an alternative $βPGM_{WT}$-substrate complex was attempted by the pairwise combination of 10 mM sodium phosphate and 10 mM glucose in the crystallisation buffer (Baxter et al., 2010). Only when the two ligands were included in

both the mother-liquor and the cryoprotectant was an open $\beta PGM_{WT}:P_i$ complex solved to 1.8 Å resolution (PDB 6H93; Table 7.1, Figure 7.2, Figure 7.3). Both monomers of the asymmetric unit reveal that the phosphate anion is coordinated in the distal site by the guanidinium group of R49 and the alkylammonium sidechains of K76 (via a water molecule) and K117 (Figure 7.7C, D), but analysis of the domain arrangements shows that chain B is slightly more closed. These residues occupy identical locations to those present in substrate-free βPGM and their Cα atom positions are *ca.* 3 Å more separated than their equivalent positions in the fully closed TSA complexes. Notably, there was neither any evidence of glucose present in the complex nor of a phosphate anion coordinated in the proximal site. Therefore, the $\beta PGM_{WT}:P_i$ complex presents a binding mode for phosphate anion coordination in the distal site, which is independent of a covalently attached hexose group and it offers a plausible mechanism for the phosphate anion inhibition of βPGM_{WT} activity reported previously (Johnson et al., 2018). Furthermore, the observation of a phosphate anion and not glucose in the active site of the open $\beta PGM_{WT}:P_i$ complex is consistent with the distal site cation-phosphodianion coordination dominating the initial interaction between substrate ($\beta G1P$ or G6P) and βPGM_{WT}^P or between the $\beta G16BP$ intermediate and substrate-free βPGM_{WT}, with consequential stabilisation of the closed enzyme.

6.5 Discussion

Sidechain substitution of the guanidinium group of R49 in either βPGM_{R49K} or βPGM_{R49A} impairs βG1P, G6P and βG16BP binding and leads to the partial alleviation of βG1P inhibition. Such changes are also mirrored by modest reductions in their observed catalytic rate constants compared to βPGM_{WT}. These substitutions induce an alternative coordination of the non-reacting phosphodianion group via the recruitment of neighbouring alkylammonium sidechains in the TSA complexes involving G6P, however the proximal site architecture, expulsion of water from the active site and degree of domain closure are equivalent to βPGM_{WT} TSA complexes. Hence, distal site phosphodianion group interactions are sufficient to stabilise a fully-closed, near-transition state conformation.

The free energy contribution of the distal site cation-phosphodianion interaction to enzyme closure can be obtained by measuring the change in the stability of both the Michaelis complex ($\Delta\Delta G_S$) and the transition state ($\Delta\Delta G^{\ddagger}$) on perturbation of the key cationic sidechain. When comparing the $\Delta\Delta G_S$ and $\Delta\Delta G^{\ddagger}$ components of the distal site cation-phosphodianion interaction energy, one of two scenarios are observed that reveal the impact of the perturbation on the catalytic cycle: 1) a dominant $\Delta\Delta G_S$ component indicates that E_C:S becomes destabilized and the Michaelis complex switches from E_C:S to E_O:S, with enzyme closure (E_O:S $\rightarrow$ E_C:S) becoming part of the rate-limiting step of the reaction, or 2) a dominant $\Delta\Delta G^{\ddagger}$ component indicates that E_O:S remains as the Michaelis complex (Go et al., 2010; Goryanova et al., 2013; Reyes et al., 2015a). In βPGM, apparent K_d (G6P) and k_{obs} values are used to estimate the $\Delta\Delta G_S$ and $\Delta\Delta G^{\ddagger}$ components of the impact on enzyme closure following R49 sidechain substitution. The 4.3 kcal·mol^{-1} and 6.0 kcal·mol^{-1} interaction energy derived using the kinetic parameters for βPGM_{R49K} and βPGM_{R49A}, respectively, are partitioned into a larger $\Delta\Delta G_S$ component ($\Delta\Delta G_S$ = 3.4 kcal·mol^{-1} and $\Delta\Delta G_S$ = 4.5 kcal·mol^{-1}, respectively) and a smaller $\Delta\Delta G^{\ddagger}$ component ($\Delta\Delta G^{\ddagger}_{kcat}$ = 0.9 kcal·mol^{-1} and $\Delta\Delta G^{\ddagger}_{kcat}$ = 1.5 kcal·mol^{-1}, respectively) (Figure 6.8). Such a partitioning implies that the Michaelis complex of βPGM_{WT} is E_C:S, but switches to E_O:S in βPGM_{R49K} and βPGM_{R49A}. However, due to the redundancy of coordinating cationic sidechains within the distal site, the impact of the R49 sidechain substitution on E_C:S stabilisation derived from the analysis of βPGM_{R49A} may be underestimated.
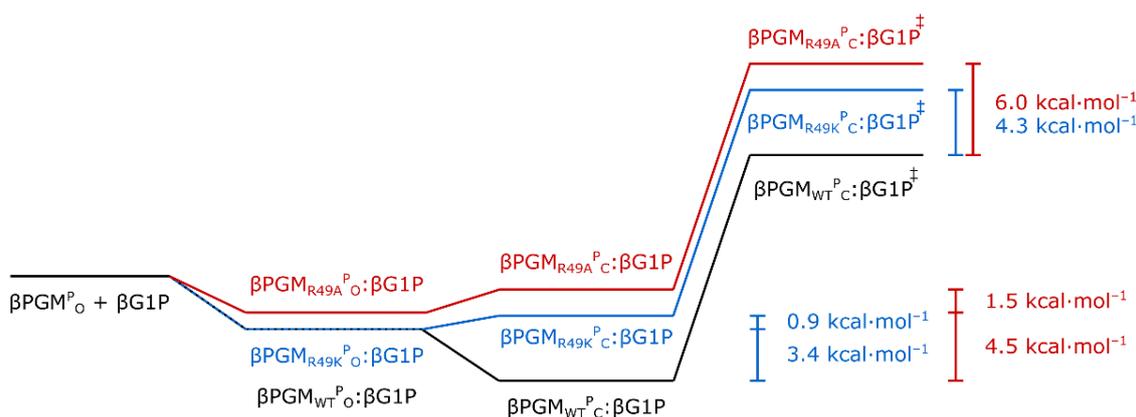
Figure 6.8 Free energy reaction profiles for βPGM illustrating the effect of R49 sidechain substitution on the kinetic parameters. The apparent dissociation constants together with the observed catalytic rate constants were used to estimate the cationic sidechain contribution to the intrinsic phosphodianion binding energy and its role in the stabilisation of the transition state. The 4.3 kcal·mol^{-1} and 6.0 kcal·mol^{-1} contributions derived using the kinetic parameters for βPGM$_{R49K}$ and βPGM$_{R49A}$, respectively, are partitioned into a larger ΔΔG$_S$ component (3.4 kcal·mol^{-1} and 4.5 kcal·mol^{-1}, respectively) and a smaller ΔΔG‡ component (0.9 kcal·mol^{-1} and 1.5 kcal·mol^{-1}, respectively). The dominant ΔΔG$_S$ component implies that the energy derived from binding the non-reacting phosphodianion group of βG1P in the distal site of the open phosphoenzyme (βPGM^P_O) is utilised primarily to trigger a shift in the equilibrium from an open βPGM^P_O:βG1P complex to a catalytically competent closed βPGM^P_C:βG1P complex. The energy difference between the βPGM^P_C:βG1P complex and the βPGM^P_C:βG1P‡ transition state is not significantly affected by each of the distal site sidechain substitutions as demonstrated by the minimal inter-site communication observed in near-transition state complexes. The change in stability of the respective βPGM^P_C:βG1P‡ transition states (4.3 kcal·mol^{-1} and 6.0 kcal·mol^{-1}) results from the differential stability of the corresponding closed βPGM^P_C:βG1P complexes. The energy of the βPGM$_{WT}$$^P_O$:βG1P complex is estimated to be similar to that of the βPGM$_{R49K}$$^P_O$:βG1P complex, since both retain a cationic charge in the distal site.

The dominant ΔΔG$_S$ component in βPGM mirrors that reported for OMPDC (ΔΔG$_S$ = 4.0 kcal·mol^{-1} and ΔΔG‡ = 1.6 kcal·mol^{-1}) (Goryanova et al., 2013). In contrast, both GPDH (ΔΔG$_S$ = 2.8 kcal·mol^{-1} and ΔΔG$^{\ddagger}_{kcat}$ = 6.3 kcal·mol^{-1}) and TIM (ΔΔG$_S$ = 2.3 kcal·mol^{-1} and ΔΔG$^{\ddagger}_{kcat}$ = 5.6 kcal·mol^{-1}) display dominant ΔΔG‡ components (Go et al., 2010; Reyes et al., 2015a). Hence, for both βPGM and OMPDC, the favored Michaelis complex is E$_C$:S and for both GPDH and TIM the favored Michaelis complex is E$_O$:S. Furthermore, the nature of the favored Michaelis complex does not correlate with the complexity of the conformational change required upon enzyme closure, since βPGM, OMPDC and GPDH all display large non-H atom RMSD values (> 2.0 Å) between open and closed enzyme forms (Figure 6.7C, Figure 6.9, Figure 7.2). Moreover, OMPDC displays a dominant ΔΔG‡ component for the proton-deuterium

exchange reaction involving the non-native substrate 5-fluorouridine 5'-monophosphate ($\Delta\Delta G_S$ = 2.2 kcal·mol^{-1} and $\Delta\Delta G^{\ddagger}$ = 5.0 kcal·mol^{-1}), implying that for this reaction (OMPDC*; Figure 6.9), the Michaelis complex is E_O:S (Goryanova et al., 2013). This observation suggests that nature of the favored Michaelis complex is instead dependent on the reaction catalyzed by the enzyme. OMPDC has a catalytic proficiency of 3×10^{10} M^{-1} for the proton-deuterium exchange reaction (Tsang et al., 2012), which is comparable to 7×10^{10} M^{-1} for the hydride transfer reaction between NADH and dihydroxyacetone phosphate in GPDH (Reyes, Amyes and Richard, 2018) and 2×10^{12} M^{-1} for the proton transfer isomerisation reaction between dihydroxyacetone phosphate and (R)-glyceraldehyde 3-phosphate in TIM (Go et al., 2010; Hall and Knowles, 1975). In contrast, OMPDC possesses a catalytic proficiency of 4×10^{22} M^{-1} for the decarboxylation of orotidine 5'-monophosphate (Goryanova et al., 2013; Miller and Wolfenden, 2002; Richard, Amyes and Reyes, 2018) and βPGM has a catalytic proficiency of 4×10^{26} M^{-1} for monoester phosphoryl transfer (Lad, Williams and Wolfenden, 2003). Therefore, the nature of the favored Michaelis complex correlates with the intrinsic difficulty of the enzyme catalyzed reaction (Figure 6.9). In conclusion, the analysis described here for βPGM, together with the data for GPDH, TIM and OMPDC, uncovers a trend, whereby the enzyme catalytic proficiency correlates with the utilization of the substrate binding energy for stabilization of E_C:S.

An underlying assumption of this framework, supported by simulations (Kulkarni et al., 2018), is that the only consequence of cationic sidechain substitution in an enzyme distal site is the destabilisation of E_C:S. Hence, once E_C:S has been attained, the organisation of catalytic groups within the solvent depleted active site is sufficient for catalysis to occur, implying that there is no significant communication between the phosphodianion group binding site and the site of the chemical step. Experimental evidence to support such minimal inter-site communication is observed in βPGM through only modest ^{19}F chemical shift changes (< 1.3 ppm) of the AlF$_4^-$ and MgF$_3^-$ moieties present in the βPGM$_{R49K}$ and βPGM$_{R49A}$ TSA complexes compared to their βPGM$_{WT}$ counterparts. Additionally, the magnitude of the weighted backbone amide chemical shift changes observed are not consistent with a pronounced alteration in active site architecture. Therefore in βPGM, no substantial inter-site communication exists between the proximal and distal sites in a fully closed, near-transition state conformation.
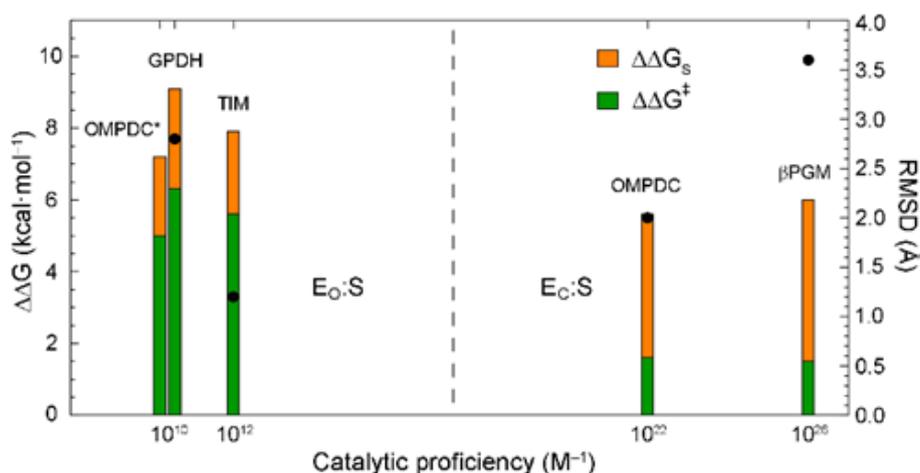
Figure 6.9 The relationship between the partitioning of the cationic sidechain contribution to the intrinsic phosphodianion binding energy and catalytic proficiency of the enzyme. For glycerol 3-phosphate dehydrogenase (GPDH), triose phosphate isomerase (TIM), orotidine 5'-monophosphate decarboxylase (OMPDC) and βPGM, perturbations of the active site responsible for coordinating the non-reacting phosphodianion group of the substrate allow the contributions to the intrinsic binding energy to be partitioned into $\Delta\Delta G_S$ components (orange bars) and $\Delta\Delta G^{\ddagger}_{kcat}$ components (green bars). For OMPDC*, $\Delta\Delta G_S$ = 2.2 kcal·mol^{-1} and $\Delta\Delta G^{\ddagger}$ = 5.0 kcal·mol^{-1} (Goryanova et al., 2013). For GPDH, $\Delta\Delta G_S$ = 2.8 kcal·mol^{-1} and $\Delta\Delta G^{\ddagger}_{kcat}$ = 6.3 kcal·mol^{-1} (Reyes et al., 2015a). For TIM, $\Delta\Delta G_S$ = 2.3 kcal·mol^{-1} and $\Delta\Delta G^{\ddagger}_{kcat}$ = 5.6 kcal·mol^{-1} (Go et al., 2010). For OMPDC, $\Delta\Delta G_S$ = 4.0 kcal·mol^{-1} and $\Delta\Delta G^{\ddagger}_{kcat}$ = 1.6 kcal·mol^{-1} (Goryanova et al., 2013). For βPGM (using βPGM$_{WT}$ and βPGM$_{R49A}$ kinetic parameters), $\Delta\Delta G_S$ = 4.5 kcal·mol^{-1} and $\Delta\Delta G^{\ddagger}_{kcat}$ = 1.5 kcal·mol^{-1}. The catalytic proficiency of an enzyme is calculated as $(k_{ex}/K_d)/k_{non}$, $(k_{cat}/K_m)/k_{non}$ or $(k_{obs}/K_d)/k_{non}$, where k_{non} is the rate constant for the corresponding spontaneous non-catalysed reaction. For OMPDC*, $(k_{ex}/K_d)/k_{non}$ = 3 × 10^{10} M^{-1} for the proton-deuterium exchange reaction involving 5-fluorouridine 5'-monophosphate (Tsang et al., 2012). For GPDH, $(k_{cat}/K_m)/k_{non}$ = 7 × 10^{10} M^{-1} for the hydride transfer reaction between NADH and dihydroxyacetone phosphate (Reyes, Amyes and Richard., 2018). For TIM, $(k_{cat}/K_m)/k_{non}$ = 2 × 10^{12} M^{-1} for the proton transfer isomerisation reaction between dihydroxyacetone phosphate and (R)-glyceraldehyde 3-phosphate (Go et al., 2010; Hall and Knowles, 1975). For OMPDC, $(k_{cat}/K_m)/k_{non}$ = 4 × 10^{22} M^{-1} for the decarboxylation of orotidine 5'-monophosphate (Goryanova et al., 2013; Miller and Wolfenden, 2002; Richard et al., 2018). For βPGM, using k_{obs} = 70 s^{-1}, apparent K_d = 9 μM and k_{non} = 2.0 × 10^{-20} s^{-1} for the spontaneous non-catalysed rate constant for phosphomonoester dianion hydrolysis (Lad, Williams and Wolfenden., 2003), $(k_{obs}/K_d)/k_{non}$ = 4 × 10^{26} M^{-1} for the conversion of βG1P to G6P via a βG16BP reaction intermediate. The extent of the conformational change upon substrate binding is reported as pairwise non-H RMSD values derived from the structures of open and closed enzymes (black circles). For GPDH, RMSD = 2.8 Å (PDB 6E8Z chain A and PDB 6E90 chain A). For TIM, RMSD = 1.3 Å (PDB 3TIM chain A and PDB 1IIH chain B). For OMPDC, RMSD = 2.0 Å (PDB 1DQW and PDB 1DQX). For βPGM, RMSD = 3.6 Å (PDB 2WFA and PDB 2WF8).

The minimal extent of inter-site communication enables the kinetic and structural consequences of distal site perturbations to be deconvoluted from those elicited by proximal site perturbations. In substrate-free βPGM_{D170N}, both Mg_{cat}^{2+} binding and catalytic activity are impaired, whilst the K_m (βG1P) and K_i (βG1P) are only modestly affected (Wood et al., 2021). Structurally, βG1P binding in a non-productive mode enables the ground state βPGM_{D170N}:βG1P complex to adopt a fully closed near-transition state architecture. This observation is consistent with the βG1P-dependent lag phase operating in βPGM (Golicnik et al., 2009; Johnson et al., 2018; Wood et al., 2020), which is partially alleviated in kinetic assays involving βPGM_{R49K} and βPGM_{R49A}. Additionally, αGal1P is another hexose 1-phosphate that behaves as a competitive inhibitor of βPGM (K_i = 30 µM; Tremblay et al., 2005). Although αGal1P is a poor surrogate for βG1P, owing to differences in stereochemistry at both the C1 and C4 positions, the βPGM_{WT}:αGal1P complex can adopt a similar fully closed near-transition state architecture (Figure 7.7B) (PDB 1Z4O and PDB 1Z4N; Tremblay et al., 2005). Therefore, coordination of the hexose 1-phosphate phosphodianion group in the distal site leads to enzyme closure, whereas a phosphate anion (P_i) in isolation cannot stabilise a fully closed complex. Additionally, the βPGM_{WT}:P_i complex suggests that the residue sidechains comprising the distal site are pre-organised to provide the initial mode of phosphodianion group interaction (Figure 7.7C, D). In summary, binding of the phosphodianion group of either the substrates, the reaction intermediate or non-native hexose monophosphates in the distal site stabilises E_C:S. Furthermore, these results suggest that βPGM is poised to utilise the phosphodianion group binding energy to stabilise E_C:S, but at the risk of introducing substrate inhibition.

In conclusion, the results presented establish a structural model of how enzymes utilise substrate binding energy to achieve significant catalytic proficiencies necessary for efficient enzyme catalysis. Moreover, for enzymes that catalyse intrinsically difficult reactions, this substrate binding energy is fully expressed in the stabilisation of the catalytically competent closed complex prior to the achievement of the transition state. However, this catalytic proficiency mechanism risks introducing substrate inhibition to catalysis.

6.6 Accession codes

Data supporting the findings of this manuscript are available from the corresponding author upon reasonable request. The atomic coordinates and structure factors have been deposited in the Protein Data Bank (www.rcsb.org) with the following codes: substrate-free βPGM_{D170N} (PDB 6HDF), βPGM_{D170N}:βG1P complex (PDB 6HDG), βPGM_{WT}:P_i complex (PDB 6H93), substrate-free βPGM_{R49K} (PDB 6HDH), substrate-free βPGM_{R49A} (PDB 6HDI), βPGM_{WT}:AlF_4:G6P TSA complex (PDB 2WF6), βPGM_{R49K}:AlF_4:G6P TSA complex (PDB 6HDJ), βPGM_{R49A}:AlF_4:G6P TSA complex (PDB 6HDK), βPGM_{R49K}:MgF_3:G6P TSA complex (PDB 6HDL) and βPGM_{R49A}:MgF_3:G6P TSA complex (PDB 6HDM).

6.7 Acknowledgements

6.8 Author contributions

A.J.R., F.A.C.N., H.P.W., N.J.B., J.P.W. designed research; A.J.R., H.P.W., N.V. produced isotopically enriched protein; H.P.W. performed coupled assay kinetic experiments; H.P.W. produced and purified βG1P; A.J.R., F.A.C.N., H.P.W., N.V., A.M.H., M.J.C. acquired NMR experiments; A.J.R., F.A.C.N., N.J.B. analysed NMR data; A.J.R., H.P.W., C.B. performed and analysed X-ray crystallography experiments; F.A.C.N. interpreted the data; F.A.C.N., H.P.W., N.J.B., J.P.W. wrote the paper with help from all authors.

6.9 Competing interests

The authors declare no competing interest.

7 Supplementary Information

SUPPLEMENTARY INFORMATION

Enzymes that catalyse intrinsically difficult reactions harvest substrate binding energy for stabilisation of the closed state but at the risk of introducing substrate inhibition

Angus J. Robertson,[1,#] F. Aaron Cruz-Navarrete,[1,#] Henry P. Wood,[1] Nikita Vekaria,[2] Andrea M. Hounslow,[1] Claudine Bisson,[1,4] Matthew J. Cliff,[2] Nicola J. Baxter,[1,2] and Jonathan P. Waltho[*,1,2]

[1] Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield, S10 2TN, United Kingdom

[2] Manchester Institute of Biotechnology and School of Chemistry, The University of Manchester, Manchester, M1 7DN, United Kingdom

[#]A.J.R. and F.A.C.N. contributed equally to this work

[*]To whom correspondence may be addressed: Prof. Jonathan P. Waltho, Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield, S10 2TN, +44 114 22717, j.waltho@sheffield.ac.uk, ORCID 0000-0002-7402-5492

Figure 7.1 Solution behaviour of substrate-free βPGM. (A–B) Pairwise overlays of $^1H^{15}N$-TROSY spectra for (A) substrate-free βPGM$_{WT}$ (black) and substrate-free βPGM$_{R49K}$ (blue) and (B) substrate-free βPGM$_{WT}$ (black) and substrate-free βPGM$_{R49A}$ (red), acquired in standard NMR buffer. Backbone amide peaks for βPGM$_{R49K}$ and βPGM$_{R49A}$ that shift their positions relative to βPGM$_{WT}$ are labelled. There is high degree of correspondence between βPGM$_{WT}$ and βPGM$_{R49K}$, indicating that the R49K substitution does not have a significant impact on the protein fold (residues V47 and N77 are with 5 Å of K49). In marked contrast, small but widespread differences in peak positions between βPGM$_{WT}$ and βPGM$_{R49A}$ show that the R49A substitution has a moderate effect on the solution properties of the helical cap domain (T16–V87). *Cis-trans* isomerisation of the K145-P146 peptide bond which is observed in βPGM$_{WT}$ (Wood et al., 2020) is also present for βPGM$_{R49K}$ and βPGM$_{R49A}$, and results in the population of two conformers in slow exchange (~70% conformer A and ~30% conformer B). Additionally, ca. six peaks are present in βPGM$_{R49A}$ that are absent in βPGM$_{WT}$ due to backbone conformational exchange on the millisecond timescale (Wood et al., 2020). This observation indicates that residue A49 in βPGM$_{R49A}$ abolishes the intermediate exchange dynamic that residue R49 propagates in βPGM$_{WT}$. (C–D) Chemical shift perturbations for substrate-free βPGM$_{R49K}$ and substrate-free βPGM$_{R49A}$ with respect to substrate-free βPGM$_{WT}$ calculated as $\Delta\delta = [(\delta_{HN-X} - \delta_{HN-Y})^2 + (0.13\times(\delta_{N-X} - \delta_{N-Y}))^2]^{1/2}$, where X and Y are the two species being compared. (C) Crystal structure of βPGM$_{R49K}$ (PDB 6HDH, chain A) showing residues of the cap domain with 0.00 ppm < $\Delta\delta$ ≤ 0.11 ppm coloured in shades of blue for the βPGM$_{WT}$ and βPGM$_{R49K}$ pairwise comparison. Mg$_{cat}^{2+}$ (green sphere) and residue K49 (green sticks) are highlighted. (D) Crystal structure of βPGM$_{R49A}$ (PDB 6HDI, chain A) showing residues of the cap domain with 0.00 ppm < $\Delta\delta$ ≤ 0.16 ppm coloured in shades of red for the βPGM$_{WT}$ and βPGM$_{R49A}$ pairwise comparison. Mg$_{cat}^{2+}$ (green sphere) and residue A49 (green sticks) are highlighted.

	1ZOL_A	2WHE_A	6HDH_A	6HDI_A	6H93_A	6HDF_B	6HDF_A	6H93_B	6HDH_B	6HDI_B	2WF9_A	5OK1_A	5OK0_A	2WF8_A	6HDM_A	6HDK_A	2WF5_A	2WF6_A	6HDJ_A	6HDG_A	6HDL_A
1ZOL_A	0.0	0.0	8.1	9.3	9.1	12.7	14.9	17.0	17.1	18.4	25.3	33.7	37.3	33.5	32.7	32.7	32.8	32.8	33.3	33.5	33.5
2WHE_A	0.0	0.0	9.3	10.3	10.0	14.4	16.6	19.1	19.1	20.5	27.4	35.9	39.5	35.6	35.0	35.0	35.0	35.1	35.5	35.9	35.8
6HDH_A	8.1	9.3	0.0	0.0	0.0	6.5	8.5	11.6	11.0	12.5	24.3	30.7	34.1	29.7	29.0	29.0	29.0	29.0	29.5	29.6	29.7
6HDI_A	9.3	10.3	0.0	0.0	0.0	5.7	7.6	11.0	10.2	11.7	24.0	30.3	33.6	29.1	28.4	28.5	28.5	28.5	29.0	29.0	29.2
6H93_A	9.1	10.0	0.0	0.0	0.0	6.2	8.2	11.7	10.7	12.4	24.9	31.1	34.4	29.8	29.2	29.3	29.3	29.3	29.8	29.8	30.0
6HDF_B	12.7	14.4	6.5	5.7	6.2	0.0	1.7	5.5	4.6	6.0	20.8	25.9	29.0	21.5	20.2	20.4	20.6	20.1	20.9	21.4	20.5
6HDF_A	14.9	16.6	8.5	7.6	8.2	1.7	0.0	4.9	4.9	5.5	20.1	24.5	27.5	22.8	21.6	21.6	22.3	22.3	22.7	22.0	22.7
6H93_B	17.0	19.1	11.6	11.0	11.7	5.5	4.9	0.0	2.4	0.0	15.5	19.8	22.8	17.0	17.6	17.4	18.9	18.9	19.2	19.2	20.4
6HDH_B	17.1	19.1	11.0	10.2	10.7	4.6	4.9	2.4	0.0	1.6	17.5	21.5	24.4	21.7	17.9	21.3	17.9	20.1	18.0	19.3	20.5
6HDI_B	18.4	20.5	12.5	11.7	12.4	6.0	5.5	0.0	1.6	0.0	16.1	19.6	22.5	19.1	17.2	16.1	18.7	18.8	17.6	17.7	18.0
2WF9_A	25.3	27.4	24.3	24.0	24.9	20.8	20.1	15.5	17.5	16.1	0.0	10.0	13.5	17.4	17.2	16.9	17.3	17.5	17.4	16.9	17.6
5OK1_A	33.7	35.9	30.7	30.3	31.1	25.9	24.5	19.8	21.5	19.6	10.0	0.0	3.6	13.1	13.3	12.9	13.2	13.4	13.0	12.8	13.2
5OK0_A	37.3	39.5	34.1	33.6	34.4	29.0	27.5	22.8	24.4	22.5	13.5	3.6	0.0	13.6	14.1	13.7	14.0	14.2	13.7	13.4	13.9
2WF8_A	33.5	35.6	29.7	29.1	29.8	21.5	22.8	17.0	21.7	19.1	17.4	13.1	13.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6HDM_A	32.7	35.0	29.0	28.4	29.2	20.2	21.6	17.6	17.9	17.2	17.2	13.3	14.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6HDK_A	32.7	35.0	29.0	28.5	29.3	20.4	21.6	17.4	21.3	16.1	16.9	12.9	13.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2WF5_A	32.8	35.0	29.0	28.5	29.3	20.6	22.3	18.9	17.9	18.7	17.3	13.2	14.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2WF6_A	32.8	35.1	29.0	28.5	29.3	20.1	22.3	18.9	20.1	18.8	17.5	13.4	14.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6HDJ_A	33.3	35.5	29.5	29.0	29.8	20.9	22.7	19.2	18.0	17.6	17.4	13.0	13.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6HDG_A	33.5	35.9	29.6	29.0	29.8	21.4	22.0	19.2	19.3	17.7	16.9	12.8	13.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6HDL_A	33.5	35.8	29.7	29.2	30.0	20.5	22.7	20.4	20.5	18.0	17.6	13.2	13.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Angle [degrees]

Open 0 | Open I | Open II | NAC | TSA

Figure 7.2 The closure angle (°) describing cap domain movement through rotation at the interdomain hinge between pairs of βPGM crystal structures determined using DynDom. Comparisons where no dynamic domains were found by the algorithm are denoted with a rotation angle of 0.0° with an upper bound of 0.2°. Crystal structures of substrate-free βPGM and βPGM complexes, together with their corresponding PDB identification codes are listed as follows: substrate-free βPGM$_{WT}$ (PDB 1ZOL; Zhang et al., 2005), substrate-free βPGM$_{WT}$ (PDB 2WHE; Baxter et al., 2010), substrate-free βPGM$_{R49K}$ (PDB 6HDH), substrate-free βPGM$_{R49A}$ (PDB 6HDI), βPGM$_{WT}$:P$_i$ complex (PDB 6H93), substrate-free βPGM$_{D170N}$ (PDB 6HDF), βPGM$_{WT}$:BeF$_3$:G6P complex (PDB 2WF9; Griffin et al., 2012), βPGM$_{D10N}$:βG16BP complex (PDB 5OK1; Johnson et al., 2018), βPGM$_{D10N}$:βG16BP complex (PDB 5OK0; Johnson et al., 2018), βPGM$_{WT}$:BeF$_3$:βG1P complex (PDB 2WF8; Griffin et al., 2012), βPGM$_{R49A}$:MgF$_3$:G6P complex (PDB 6HDM), βPGM$_{R49A}$:AlF$_4$:G6P complex (PDB 6HDK), βPGM$_{WT}$:MgF$_3$:G6P complex (PDB 2WF5; Baxter et al., 2010), βPGM$_{WT}$:AlF$_4$:G6P complex (PDB 2WF6), βPGM$_{R49K}$:AlF$_4$:G6P complex (PDB 6HDJ), βPGM$_{D170N}$:βG1P complex (PDB 6HDG) and βPGM$_{R49K}$:MgF$_3$:G6P complex (PDB 6HDL). PDB identification codes containing suffixes _A and _B denote chain A and chain B, respectively for monomers of the asymmetric unit. Crystal structures have been categorised as open with three clusters of interdomain closure angle, near attack complexes (NAC) or transition state analogue (TSA) complexes and are indicated by bars.

Figure 7.3 Non-H atom RMSD values (Å) for cap domain and core domain alignments between pairs of βPGM crystal structures determined using PyMOL (The PyMOL Molecular Graphics System, version 1.8/2.2 Schrödinger, LLC). The upper right pseudo-triangular matrix indicates pairwise RMSD values for cap domain (T16–V87) alignments, while the lower left pseudo-triangular matrix indicates pairwise RMSD values for core domain (M1–D15, S88–K221) alignments. A greater level of perturbation is observed for residues of the cap domain in pairwise comparisons involving structures with different closure angles. Crystal structures of substrate-free βPGM and βPGM complexes, together with their corresponding PDB identification codes are listed as follows: substrate-free $βPGM_{WT}$ (PDB 1ZOL; Zhang et al., 2005), substrate-free $βPGM_{WT}$ (PDB 2WHE; Baxter et al., 2010), substrate-free $βPGM_{R49K}$ (PDB 6HDH), substrate-free $βPGM_{R49A}$ (PDB 6HDI), $βPGM_{WT}$:P_i complex (PDB 6H93), substrate-free $βPGM_{D170N}$ (PDB 6HDF), $βPGM_{WT}$:BeF_3:G6P complex (PDB 2WF9; Griffin et al., 2012), $βPGM_{D10N}$:βG16BP complex (PDB 5OK1; Johnson et al., 2018), $βPGM_{D10N}$:βG16BP complex (PDB 5OK0; Johnson et al., 2018), $βPGM_{WT}$:BeF_3:βG1P complex (PDB 2WF8; Griffin et al., 2012), $βPGM_{R49A}$:MgF_3:G6P complex (PDB 6HDM), $βPGM_{R49A}$:AlF_4:G6P complex (PDB 6HDK), $βPGM_{WT}$:MgF_3:G6P complex (PDB 2WF5; Baxter et al., 2010), $βPGM_{WT}$:AlF_4:G6P complex (PDB 2WF6), $βPGM_{R49K}$:AlF_4:G6P complex (PDB 6HDJ), $βPGM_{D170N}$:βG1P complex (PDB 6HDG) and $βPGM_{R49K}$:MgF_3:G6P complex (PDB 6HDL). PDB identification codes containing suffixes _A and _B denote chain A and chain B, respectively for monomers of the asymmetric unit. Crystal structures have been categorised as open with three clusters of interdomain closure angle, near attack complexes (NAC) or transition state analogue (TSA) complexes and are indicated by bars.

Figure 7.4 Difference density and active site details of the βPGM:AlF₄:G6P and βPGM:MgF₃:G6P TSA complexes. The active sites of (A) βPGM_{R49K}:AlF₄:G6P complex (PDB 6HDJ), (B) βPGM_{R49A}:AlF₄:G6P complex (PDB 6HDK), (C) βPGM_{R49K}:MgF₃:G6P complex (PDB 6HDL) and (D) βPGM_{R49A}:MgF₃:G6P complex (PDB 6HDM). Selected residues (sticks) and ligands (sticks) are illustrated, with the square-planar AlF_4^- moiety (dark grey and light blue), the trigonal-planar MgF_3^- moiety (green and light blue), βG6P (purple carbon atoms), structural waters (red spheres) and Mg_{cat}^{2+} (green sphere). Yellow dashes indicate hydrogen bonds and black dashes show metal ion coordination. Difference density (Fo – Fc; green mesh) is contoured at 3σ and was generated following ligand omission from the final structures. The sidechain of residue N118, which coordinates one of the 6-phosphate oxygen atoms of G6P equivalently in the TSA complexes, has been omitted for clarity.
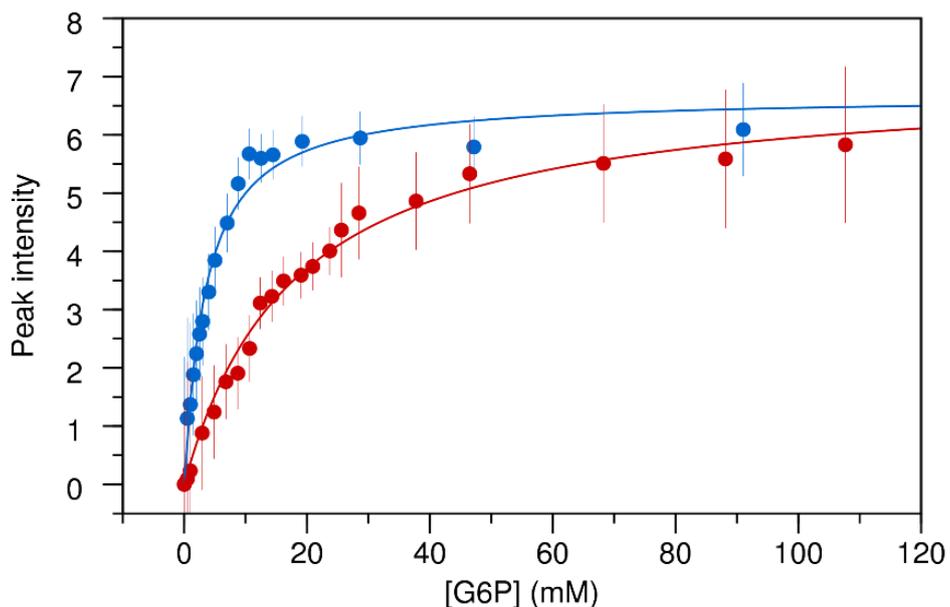
Figure 7.5 Determination of the apparent dissociation constant (K_d) for G6P in the βPGM$_{R49K}$:AlF$_4$:G6P and βPGM$_{R49A}$:AlF$_4$:G6P TSA complexes monitored using one-dimensional ^{1}H NMR spectroscopy. A solution of 360–400 mM G6P was titrated serially into βPGM$_{R49K}$ (blue circles) or βPGM$_{R49A}$ (red circles) prepared in standard NMR buffer supplemented with 15 mM NaF and 3 mM AlCl$_3$. The changing intensity of the well-resolved indole resonance of residue W24 (acting as a reporter for the closed TSA complex in slow exchange) was fitted to determine apparent dissociation constants for the βPGM$_{R49K}$:AlF$_4$:G6P TSA complex (apparent K_d = 3.0 ± 0.4 mM) and the βPGM$_{R49A}$:AlF$_4$:G6P TSA complex (apparent K_d = 18 ± 1 mM). Vertical error bars indicate estimated errors in the measurement of peak intensities.
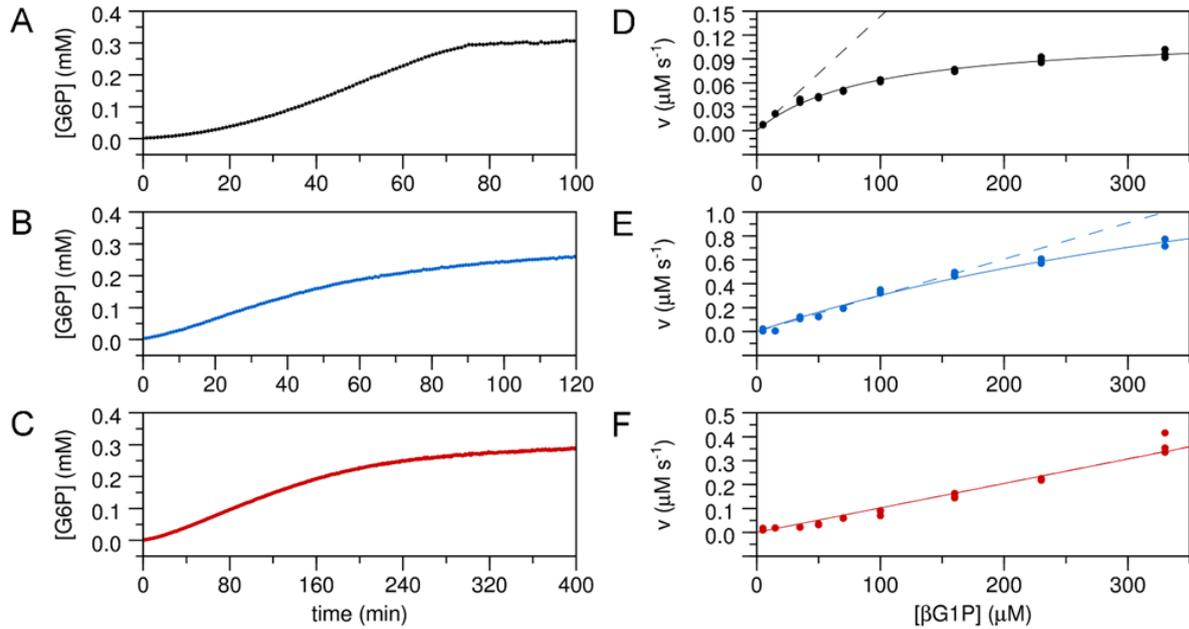
Figure 7.6 Reaction kinetics for the conversion of βG1P to G6P catalysed by βPGM$_{WT}$, βPGM$_{R49K}$ and βPGM$_{R49A}$. The rate of G6P production was measured indirectly using a glucose 6-phosphate dehydrogenase coupled assay, in which G6P is oxidized and concomitant NAD$^+$ reduction is monitored by the increase in absorbance at 340 nm. (A–C) Kinetic profiles showing the time-dependent conversion of 330 μM βG1P to G6P in the presence of 20 mM AcP (10 mM AcP for βPGM$_{WT}$) in standard kinetic buffer for (A) 5 nM βPGM$_{WT}$, (B) 60 nM βPGM$_{R49K}$ and (C) 60 nM βPGM$_{R49A}$. (D–F) Michaelis-Menten plots showing the dependence of the steady-state reaction velocity (v) on βG1P concentration (5, 15, 35, 50, 70, 100, 160, 230, 330 μM) for (D) 5 nM βPGM$_{WT}$ (n=3), (E) 60 nM βPGM$_{R49K}$ (n=3) and (F) 60 nM βPGM$_{R49A}$ (n=3). Data in each plot were fitted to the standard Michaelis-Menten equation to derive apparent k$_{cat}$ and apparent K$_m$ (βG1P) values and the line of best fit is shown (solid lines). For βPGM$_{WT}$, the βG1P concentration range used provided reliable fitted parameters (apparent k$_{cat}$ = 24.5 ± 0.7 s^{-1} and apparent K$_m$ (βG1P) = 92 ± 6 μM). However for βPGM$_{R49K}$, a weak βG1P affinity resulted in fitted parameters with large associated errors (apparent k$_{cat}$ = 35 ± 5 s^{-1} and apparent K$_m$ (βG1P) = 600 ± 100 μM). For βPGM$_{R49A}$, a linear dependence of steady-state reaction velocity on βG1P concentration precluded the derivation of fitted parameters over the accessible βG1P concentration range. Therefore, the initial data points of each Michaelis-Menten plot were fitted to a linear equation to derive the apparent k$_{cat}$/K$_m$ ratio for βPGM$_{WT}$ (apparent k$_{cat}$/K$_m$ = 0.29 s^{-1}·μM^{-1} for [βG1P] = 5–15 μM), βPGM$_{R49K}$ (apparent k$_{cat}$/K$_m$ = 0.05 s^{-1}·μM^{-1} for [βG1P] = 5–100 μM) and βPGM$_{R49A}$ (apparent k$_{cat}$/K$_m$ = 0.02 s^{-1}·μM^{-1} for [βG1P] = 5–330 μM) and the line of best fit is shown (dashed lines). For βPGM$_{R49A}$, the fitted lines derived using the standard Michaelis-Menten equation and a linear equation are overlapped.
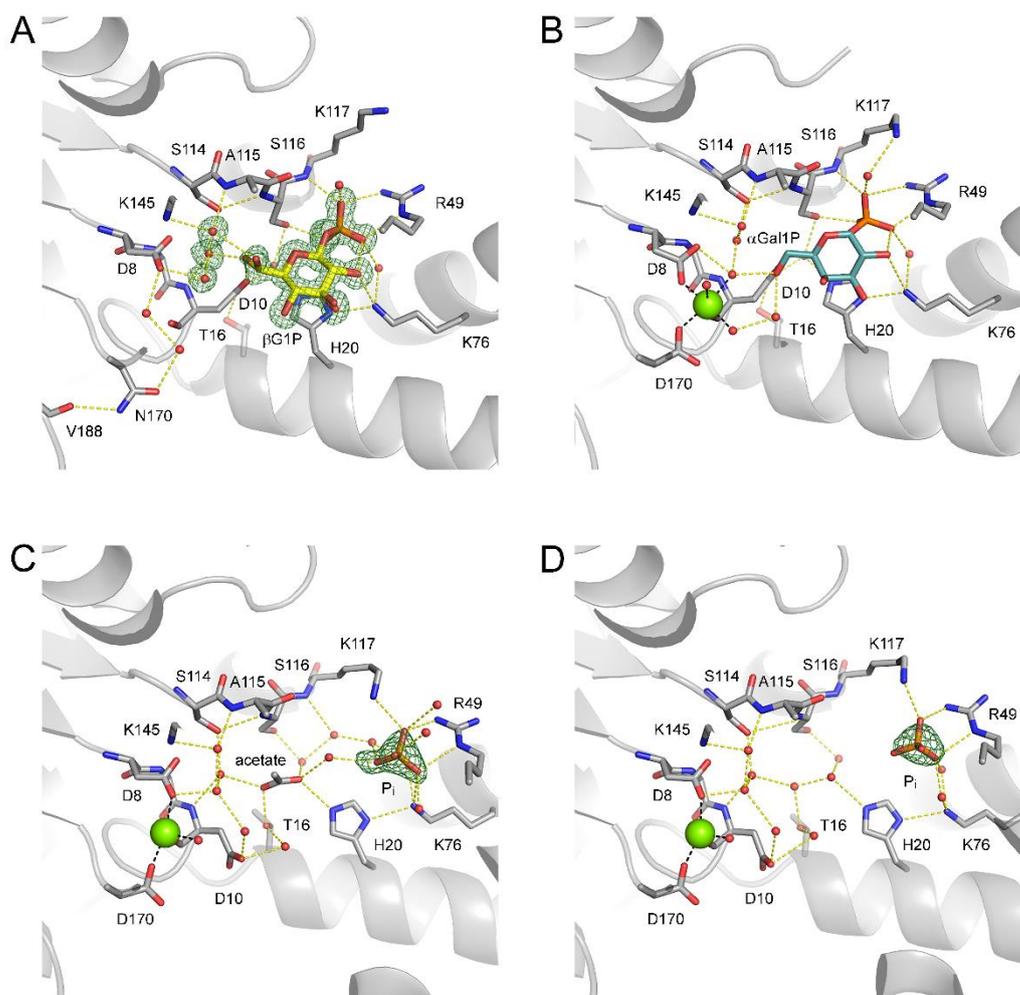
Figure 7.7   Difference density and active site details of the βPGM$_{D170N}$:βG1P complex, the βPGM$_{WT}$:αGal1P complex and the βPGM$_{WT}$:P$_i$ complex. The active site of (A) βPGM$_{D170N}$:βG1P complex (PDB 6HDG), (B) βPGM$_{WT}$:αGal1P complex (PDB 1Z4O, chain A; Tremblay et al., 2005), (C) βPGM$_{WT}$:P$_i$ complex (PDB 6H93, chain A) and (D) βPGM$_{WT}$:P$_i$ complex (PDB 6H93, chain B). Selected residues (sticks) and ligands (sticks) are illustrated, together with βG1P (gold carbon atoms), αGal1P (sky blue carbon atoms), P$_i$ occupying the distal site, structural waters (red spheres) and Mg$_{cat}^{2+}$ (green sphere). Yellow dashes indicate hydrogen bonds and black dashes show metal ion coordination. Difference density (Fo – Fc; green mesh) is contoured at 3σ and was generated following ligand omission from the final structures. The 6-phosphate group of βG1P in the proximal site has two arrangements resolved for the C5–C6 bond. The sidechain of residue N118, which coordinates one of the 1-phosphate oxygen atoms of both βG1P and αGal1P, has been omitted for clarity.

Table 7.1 X-ray data collection, data processing and refinement statistics.

Data collection and data processing statistics					
Complex	βPGM$_{D170N}$	βPGM$_{D170N}$:βG1P	βPGM$_{WT}$:P$_i$	*βPGM$_{R49K}$*	*βPGM$_{R49A}$*
PDB code	PDB 6HDF	PDB 6HDG	PDB 6H93	PDB 6HDH	PDB 6HDI
Wavelength (Å)	0.92819	0.92819	0.97950	0.97624	0.97625
Beamline, Facility	i04-1, DLS	i04-1, DLS	i04, DLS	i03, DLS	i03, DLS
Space group	P2$_1$	P2$_1$2$_1$2$_1$	P2$_1$	P2$_1$	P2$_1$
Cell dimensions:					
a, b, c, (Å)	38.77, 119.31, 53.17	37.56, 55.08, 105.84	38.35, 117.14, 53.19	38.13, 117.14, 53.01	38.20, 116.90, 53.17
α, β, γ (º)	90.0, 94.8, 90.0	90.0, 90.0, 90.0	90.0, 99.1, 90.0	90.0, 97.4, 90.0	90.0, 98.1, 90.0
Resolution (Å) [1]	52.98-1.40 (1.42-1.40)	48.86-1.15 (1.17-1.15)	32.05-1.77 (1.80-1.77)	31.36-1.62 (1.65-1.62)	52.64-2.03 (2.07-2.03)
Rmerge [1,2]	0.091 (1.243)	0.126 (1.986)	0.085 (1.426)	0.052 (0.770)	0.131 (0.829)
Rpim [1,3]	0.055 (0.701)	0.037 (0.634)	0.035 (0.583)	0.030 (0.487)	0.077 (0.488)
CC-half [1]	0.997 (0.462)	0.999 (0.562)	0.999 (0.467)	0.998 (0.491)	0.992 (0.525)
<I/σ(I)> [1]	8.2 (1.0)	10.8 (1.2)	12.5 (1.2)	12.7 (1.5)	7.5 (1.7)
Completeness (%) [1]	95.4 (95.1)	100.0 (100.0)	100.0 (100.0)	96.1 (92.7)	99.6 (99.2)
Multiplicity [1]	3.8 (4.0)	12.6 (10.6)	6.8 (6.9)	3.7 (3.4)	3.7 (3.8)
Total reflections	342050	991156	308459	210878	111221
Unique reflections	90031	78880	45050	56242	29676
Molecular replacement model	PDB 2WHE	PDB 2WF5	PDB 2WHE	PDB 2WHE	PDB 2WHE
Refinement statistics					
Complex	βPGM$_{D170N}$	βPGM$_{D170N}$:βG1P	βPGM$_{WT}$:P$_i$	βPGM$_{R49K}$	βPGM$_{R49A}$
PDB code	PDB 6HDF	PDB 6HDG	PDB 6H93	PDB 6HDH	PDB 6HDI
R (%) [4] / Rfree (%) [5]	17.2 / 22.3	14.9 / 17.8	17.9 / 23.3	18.2 / 21.8	20.5 / 27.3
Number of atoms:					
Protein [6]	1737, 1701	1772	1697, 1689	1686, 1678	1692, 1693
Ligands [7]	0	48	10	0	0
Metal ions [8]	2	1	2	2	2
Water	291	241	282	210	243
Protein residues [6]	219, 219	219	219, 219	218, 218	220, 221
RMS deviations:					
Bonds (Å)	0.01	0.01	0.12	0.01	0.01
Angles (º)	1.50	1.51	1.49	1.50	1.51
Average B factors(A^2):					
Main chain [6]	18.7, 18.1	12.5	26.7, 29.9	27.4	26.4, 28.3
Side chains [6]	23.5, 23.0	15.8	32.0, 35.1	33.4	31.5, 33.1
Ligands [7]	–	13.6	63.9	–	–
Metal ions [8]	20.4	13.5	30.5	27.7	23.3
Water	26.5	24.2	36.9	36.6	35.3
Ramachandran analysis:					
Favoured / allowed (%)	98.2	97.8	98.4	98.6	97.7
Disallowed (%)	0.0	0.0	0.0	0.0	0.0
MolProbity score (percentile)	0.76 (100th)	0.86 (100th)	0.73 (100th)	0.97 (100th)	1.16 (100th)

[1] Values for the higher resolution shell are in parentheses

$$[2]\ R_{merge} = \frac{\sum_{hkl}\sum_i |I_i - I_m|}{\sum_{hkl}\sum_i I_i}$$

$$[3]\ R_{pim} = \frac{\sum_{hkl}\sqrt{\frac{1}{n-1}}\sum_{i=1}^{n}|I_i - I_m|}{\sum_{hkl}\sum_i I_i}$$

where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively

$$[4]\ R = \frac{\sum_{hkl}||F_{obs}| - k|F_{calc}||}{\sum_{hkl}|F_{obs}|}$$

where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively

Table 7.1 X-ray data collection, data processing and refinement statistics (continued).

Data collection and data processing statistics					
Complex	βPGM$_{WT}$:AlF$_4$:G6P	βPGM$_{R49K}$:AlF$_4$:G6P	βPGM$_{R49A}$:AlF$_4$:G6P	βPGM$_{R49K}$:MgF$_3$:G6P	βPGM$_{R49A}$:MgF$_3$:G6P
PDB code	PDB 2WF6	PDB 6HDJ	PDB 6HDK	PDB 6HDL	PDB 6HDM
Wavelength (Å)	0.933	0.97625	0.97625	0.97629	0.97625
Beamline, Facility	ID14-2, ESRF	i03, DLS	i03, DLS	i03, DLS	i03, DLS
Space group	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$
Cell dimensions:					
a, b, c, (Å)	37.80, 54.50, 105.00	104.21, 37.22, 54.22	37.23, 54.29, 104.24	37.55, 54.30, 104.20	37.30, 54.34, 104.62
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å) [1]	20.00-1.40 (1.44-1.40)	48.10-1.16 (1.18-1.16)	54.29-1.24 (1.26-1.24)	37.55-1.16 (1.16-1.18)	54.34-1.30 (1.32-1.30)
Rmerge [1,2]	0.1 (0.37)	0.084 (1.082)	0.099 (1.019)	0.068 (1.345)	0.052 (0.263)
Rpim [1,3]	–	0.033 (0.460)	0.040 (0.480)	0.027 (0.591)	0.022 (0.138)
CC-half [1]	–	0.999 (0.554)	0.999 (0.530)	0.999 (0.515)	0.999 (0.944)
<I/σ(I)> [1]	7.3 (2.2)	11.2 (1.5)	10.1 (1.4)	14.2 (1.3)	21.1 (6.2)
Completeness (%) [1]	98.5 (99.4)	95.4 (88.5)	100.0 (97.9)	98.5 (92.3)	99.7 (95.3)
Multiplicity [1]	–	7.3 (6.2)	7.0 (5.3)	7.1 (5.9)	6.8 (4.5)
Total reflections	–	515051	424367	518578	361965
Unique reflections	43021	70516	60728	73452	53048
Molecular replacement model	PDB 2WF5	PDB 2WF6	PDB 2WF6	PDB 2WF5	PDB 2WF5

Data Refinement					
Complex	βPGM$_{WT}$:AlF$_4$:G6P	βPGM$_{R49K}$:AlF$_4$:G6P	βPGM$_{R49A}$:AlF$_4$:G6P	βPGM$_{R49K}$:MgF$_3$:G6P	βPGM$_{R49A}$:MgF$_3$:G6P
PDB code	PDB 2WF6	PDB 6HDJ	PDB 6HDK	PDB 6HDL	PDB 6HDM
R (%) [4] / *Rfree* (%) [5]	16.1 / 19.1	14.3 / 16.6	13.6 / 16.7	13.2 / 16.4	12.6 / 14.8
Protein [6]	1680	1739	1706	1774	1802
Ligands [7]	21	21	21	20	20
Metal ions [8]	2	1	3	2	2
Water	253	179	212	278	213
Protein residues [6]	218	219	219	219	219
RMS deviations:					
Bonds (Å)	0.01	0.01	0.01	0.01	0.01
Angles (°)	1.40	1.50	1.50	1.50	1.47
Average B factors(A^2)					
Main chain [6]	13.5	14.1	13.7	14.0	13.3
Side chains [6]	15.2	17.6	17.7	16.9	16.5
Ligands [7]	11.0	9.1, 8.8	9.0, 9.5	11.6, 10.9	10.0, 9.1
Metal ions [8]	12.4	8.6	23.1, 8.9	15.6, 9.3	14.4, 8.4
Water	23.4	26.1	28.4	27.5	24.1
Ramachandran analysis					
Favoured / allowed (%)	96.8	98.2	97.7	97.4	98.7
Disallowed (%)	0.0	0.0	0.0	0.0	0.0
MolProbity score (percentile)	1.42(99th)	0.76 (100th)	0.82 (100th)	1.07 (98th)	1.19 (95th)

[5] $R_{free} = \dfrac{\sum_{hkl \subset T} \lVert F_{obs} \rvert - k \lvert F_{calc} \rVert}{\sum_{hkl \subset T} \lvert F_{obs} \rvert}$

where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively and T is the test set of data omitted from refinement (5% in this case)

[6] For structures where there are two monomers in the asymmetric unit, the values for chain A and chain B are given, respectively

[7] Only relevant ligands are presented; other ligands (e.g. ethylene glycol and acetate, etc) have been omitted

[8] Generally, Mg^{2+} ions were only observed in the crystals, however in some cases Na$^+$ ions were also noted. Where this was the case, B-factors are listed for Na$^+$ ions and Mg^{2+} ions, respectively.

8 Conclusions and Future Directions

8.1 The conformational landscape of βPGM

This thesis has focused on the conformational landscape of βPGM and its role in catalysis and post-translational regulation. First, it was noted that substrate-free βPGM$_{WT}$ shows both extensive conformational intermediate exchange dynamics in residues surrounding the active site (Baxter et al., 2006), and slow multi-second exchange dynamics in 52% of the assigned residues (Wood et al., 2020). The βPGM$_{P146A}$ variant reduced the intermediate exchange dynamics (see section 3.3.3) (Cruz-Navarrete et al., 2019) and abolished the slow conformational exchange process (see section 4.4.3) (Wood et al., 2020). Interestingly, the K145–P146 peptide bond adopts a *cis* conformation in all the reported βPGM$_{WT}$ crystal structures and in conformer A in solution. In contrast, the K145–A146 peptide bond in βPGM$_{P146A}$ displays a *trans* conformation in the substrate-free βPGM$_{P146A}$ crystal structure and its solution conformation is closer to conformer B of βPGM$_{WT}$. Hence, the *cis-trans* proline isomerisation of the K145–P146 peptide bond is involved in both the intermediate and slow exchange dynamics of βPGM$_{WT}$.

The intermediate exchange dynamics, i.e., on the millisecond timescale, observed in βPGM is present in residues that have significant roles in catalysis and substrate binding (Figure 5.4). These millisecond dynamics are likely to play a role during the βPGM catalytic cycle, as both processes occur on the same timescale (βPGM$_{WT}$ single turnover time is ~3 ms; see section 4.4.4) (Wood et al., 2020). For instance, βPGM may be transiently populating conformations that are involved in the rate-limiting step of the catalytic cycle. Possible processes that can limit the catalytic rate are product release, rearrangement of the Mg$^{2+}_{cat}$ binding site, domain closure and secondary structural changes (Kern et al., 2007). Further investigation into the nature of the intermediate exchange dynamics in the substrate-free conformers and several ground state analogues and TSA complexes is needed to identify their relationship with the rate-limiting step in the βPGM catalytic cycle

The multi-second exchange process seen in substrate-free βPGM$_{WT}$ is caused by the *cis-trans* proline isomerisation of the K145–P146 peptide bond. The most pronounced difference between conformer A (*cis*) and conformer B (*trans*) structures is the position of the K145 sidechain relative to the active site (Figure 4.4). This positively charged

alkylammonium group provides charge balance during formation of the transition state of phosphoryl transfer. In conformer B, the failure of the alkylammonium group to engage in the active site implies that this conformer represents an inactive form of βPGM. The adoption of a *cis* K145–A146 peptide bond in the βPGM$_{P146A}$:MgF$_3$:G6P TSA complex further supports the fact that conformer A is the active form and is the only form competent for phosphoryl transfer in the catalytic cycle. Therefore, *cis-trans* K145–X146 peptide bond isomerisation works as a switch that allows βPGM to modulate its activity.

8.2 Allomorphy vs Allostery and Allokairy

Mechanisms of fine control of enzyme activity in monomeric enzymes require the existence of at least two conformations with different activities. In allostery, the conformational change between the two conformations is facilitated by the binding of a ligand at a site remote from the active site to modulate enzyme activity (Changeux, 2012). In contrast, allokairy, also known as kinetic cooperativity, is a kinetic control mechanism where the enzyme activity is modulated by the near-equivalence of the catalytic rate and the exchange rate (Whittington et al., 2015; Hilser, Anderson and Motlagh, 2015). In this thesis, a fine control mechanism related to, but distinct from, allostery and allokairy was described for βPGM (see chapter 4) (Wood et al., 2020). Allomorphy requires the existence of a conformational exchange process between two forms with differing activities that are part of the catalytic cycle. An allomorphic full activator, acting as a substrate (in the active site), can bias the enzyme population towards the faster catalytic pathway by stabilising the most active form. Conversely, a partial activator is unable to shift the enzyme population effectively to the most active form, and hence the slow conformational change becomes the rate-limiting step. In βPGM, the effective stabilisation of the *cis-trans* K145–A146 peptide bond isomerisation by the reaction intermediate βG16BP, but not F16BP or AcP (see section 4.4.5) indicates that the conformational landscape of βPGM is finely tuned to effectively switch the enzyme population to either the inactive from or active form depending on the relative carbohydrate concentrations in *L. lactis*. The detailed structural mechanism with which βPGM is able to discriminate between allomorphic full and partial activators requires further investigation.

βPGM displays a lag phase when either F16BP or AcP are used as activators (Figure 4.5). Enzymes that exhibit a burst or a lag phase in their kinetic profiles are known as hysteretic enzymes. Although several theoretical models have been proposed to explain hysteresis (Ainslie, Neet and Shill, 1972; Ricard, Meunier and Buc, 1974), before this work only the allokairy mechanism in glucokinase has been described at the structural level (Whittington et al., 2015). Allomorphy is potentially present in other hysteretic enzymes. For example, several phosphomutases also display lag phases when different activators are used but show linear kinetics when their reaction intermediate is present (see section 4.5). The study of the structural basis of their hysteretic behaviour will allow us to broaden our understanding of allomorphy as a general mechanism of post-translational control of activity in monomeric enzymes.

8.3 How substrates influence the conformational landscape of βPGM?

The last part of this thesis focused on how βPGM utilises its substrate binding energy to promote catalysis (see chapter 6). Enzymes can invest this binding energy into ground state destabilisation or transition-state stabilisation (Andrews, Fenn and Herschlag, 2013; Warshel, Arieh et al., 2006). However, βPGM and other enzymes that bind substrates with non-reacting phosphodianion groups use this energy to facilitate the transition towards the closed active form (see section 6.5). The closed conformation in βPGM sequesters the reacting phosphodianion group into the *proximal* site away from the solvent and allows the formation of favourable electrostatic interactions with active site residues (Griffin et al., 2012; Johnson et al., 2018). However, the presence of a *proximal* phosphodianion group is not required for βPGM to close, as βG1P binding in a non-productive mode, i.e., with the phosphodianion group bound in the *distal* site, also stabilises a closed conformation with near-transition state architecture (see section 6.4.5) (Figure 6.6; Figure 7.7). Therefore, βPGM is poised to utilise the substrate binding energy to populate otherwise high-energy conformations. Moreover, from the observed relationship between the increased utilisation of substrate binding energy and high catalytic proficiency across GPDH, TIM, OMPDC and βPGM, it can be inferred that efficient adoption of a catalytically competent conformation is one of the available mechanisms that enzymes exploit to catalyse intrinsically difficult reactions (see section 6.5). This knowledge can be leveraged in the design of enzymes for difficult reactions not found in Nature.

During its catalytic cycle, βPGM crosses over two high-energy barriers: 1) the conformational change from the inactive conformer B to the active conformer A and 2) the stabilisation of the closed active form. The binding of βG16BP accelerates the first process first (see chapter 4) and the energy required for the second process is derived from the binding of βG1P and G6P (see chapter 6). Given that only a *distal* phosphodianion group is required to promote closure and βG16BP contains two phosphodianion groups, it can be hypothesised that adoption of a closed conformation and *cis-trans* isomerisation of the K145–X146 peptide bond in the *proximal* site are coupled processes. This hypothesis is supported by the observation that all reported crystal structures of βPGM in a closed conformation contain a *cis* K145–X146 peptide bond, including the βPGM-hexose monophosphate complexes (see sections 4.4.3, 6.4.2 and 6.4.5) (Wood et al., 2020). It also explains why AcP is unable to stabilise conformer A, as AcP is too small to occupy both the *proximal* and *distal* sites simultaneously. Yet, F16BP is unable to bias the βPGM population towards the fastest kinetic pathway as effectively as βG16BP, even though both molecules contain two phosphodianion groups in a similar arrangement. To reconcile these two seemingly contradictory observations it is necessary to remember that allomorphy, and hysteresis in general, is a kinetic phenomenon. Binding of βG16BP to βPGM not only changes the energy minimum of its conformational landscape, but also reduces the energetic barriers between the low-energy states. In contrast, binding of F16BP is not enough to lower the energetic barrier of *cis-trans* proline isomerisation, but it does stabilise the active conformation. Hence, the lowest-energy conformation of a βPGM$_{WT}$:F16BP complex (if observable) would display a *cis* K145–P146 peptide bond. Furthermore, a βPGM$_{P146A}$:F16BP complex would show a *trans* K145–A146 peptide bond, due to the lower stability of the *cis* peptide bond in this variant (see sections 4.4.3 and 4.4.5) (Figure 4.4) (Wood et al., 2020). The detailed structural mechanism that allows βPGM to couple substrate binding, enzyme closure and *cis-trans* proline isomerisation is an exciting area of research that requires further investigation.

8.4 Is *cis-trans* proline isomerisation the rate-limiting step in the catalytic cycle of βPGM?

Although the differences between allostery and allomorphy are apparent, the distinction between allokairy and allomorphy is less evident (see section 4.5, 8.2) (Figure 4.7). An alternative hypothesis that was considered during this thesis is that

allomorphy and allokairy are extreme examples of the same temporal phenomenon. This phenomenon is the substrate-dependent exchange process between two conformations with different activities. This exchange process has been thoroughly studied in glucokinase, where the exchange rate and the catalytic rate have near-equivalence (Kamata et al., 2004; Larion and Miller, 2012; Larion et al., 2012; Larion et al., 2015; Whittington et al., 2015). In consequence, the glucokinase rate vs substrate concentration curve is sigmoidal (Figure 4.7). In this alternative scenario, allomorphy would be the result of the presence of a conformational change significantly slower than the catalytic cycle. This slow exchange process would display a lag phase in the kinetic profile, i.e., the transition from the inactive to the active conformation would be observed during the experiment. However, it would not result in a sigmoidal curve but in two distinct classic Michaelis-Menten curves, due to the possibility of deconvoluting the rate of conformational change from the true catalytic rate. Therefore in this scenario, both allomorphy and allokairy are extreme examples of a continuum of behaviours where the substrate-dependent conformational landscape of an enzyme regulates its activity.

In opposition to this alternative hypothesis is the observation of the B^P species in both βPGM_{WT} and βPGM_{P146A} (see section 4.4.6) (Figure 4.2; Figure 5.3). For B^P to be generated, conformer B needs to bind the activator and catalyse phosphoryl transfer, albeit at a lower rate. In chapter 4 it was concluded that conformer B is an inactive form, but this was only proven for the step $\beta PGM:\beta G16BP \rightarrow \beta PGM^P:G6P$ (see section 4.4.5; 4.4.6) (Figure 4.2; Figure 4.4; Figure 5.3). It is still unknown which conformer is better poised to catalyse phosphoryl transfer when F16BP or AcP are bound. Hence, it was proposed that two different kinetic pathways exist in βPGM: a fast one that goes through conformer A (and A^P), and a slow one that involves the conformer B to conformer A (or B^P to A^P) transition.

An implicit difference between the two hypotheses is the placement of the *cis-trans* proline isomerisation event in the βPGM catalytic cycle. In the alternative hypothesis, *cis-trans* proline isomerisation happens outside the catalytic cycle (the off-pathway mechanism). In the mechanism proposed in chapter 4, *cis-trans* proline isomerisation is part of the cycle (the on-pathway mechanism), i.e., peptide bond isomerisation occurs every catalytic turnover. Evidence supporting the on-pathway mechanism is the lack of a lag phase when $\beta G16BP$ is used as an activator (see section 4.4.4)

(Figure 4.5), implying that βG16BP binding accelerates peptide bond isomerisation in the conformer B population within the dead time of the coupled assay experiments. In contrast, a lag phase with a time constant similar to the exchange time (between 1 s and 5 min) would be observed if the off-pathway mechanism were present in βPGM. Further evidence is derived from the comparison of activity and stability parameters between βPGM_{WT} and βPGM_{P146A} and *cis-trans* peptide bond isomerisation rates of wild-type and P39A variant of RNase T1 refolding (Table 8.1; Figure 8.1) (Odefey, Mayr and Schmid, 1995). To my knowledge, *cis-trans* isomerisation rates of both prolyl and alanyl peptide bonds in proteins have only been determined in RNase T1 by stopped-flow experiments. The P39A substitution in RNase T1 destabilises the *cis* Y38–X39 peptide bond by 2.4 kcal mol^{-1} (Table 8.1a; Figure 8.1). Similarly, the P146A substitution in βPGM destabilises the *cis* K145–X146 peptide bond by at least 2.2 kcal mol^{-1} (Table 8.1b; Figure 8.1). Furthermore, βPGM_{P146A} displays a 1.8 kcal mol^{-1} free energy change in both k_{cat} and K_m (βG16BP) values (Table 8.1b), closely mirroring both the RNase T1 and βPGM stability changes upon proline to alanine substitution. This correspondence between free energy differences is not consistent with the off-pathway mechanism, as the proline isomerisation step would be outside the catalytic cycle.

Two different strategies can be employed to determine experimentally if *cis-trans* proline isomerisation occurs every turnover in βPGM. First, a comparison between the k_{cat} values and the K145–X146 peptide bond isomerisation rates can be made for a series βPGM variants where position 146 is substituted separately with every amino acid. If peptide bond isomerisation is rate-limiting, a positive correlation between both kinetic parameters will observed. Second, through the incorporation of fluorinated proline residues in βPGM, the K145–P146 peptide bond isomerisation can be monitored during catalysis using relaxation-dispersion ^{19}F NMR spectroscopy. If the on-pathway mechanism is present, the measured exchange rate of the ^{19}F nuclei will mirror the independently determined k_{cat} of βPGM.

Table 8.1 Changes in free energy due to proline to alanine substitution in RNase T1 refolding and βPGM activity.

Table 8.1a Comparison between the *cis-trans* isomerisation free energy barriers of Tyr-Pro and Tyr-Ala peptide bonds in RNase T1 refolding (Odefey, Mayr and Schmid, 1995).

	WT	P39A	$\Delta\Delta G_{eq}$ or $\Delta\Delta G^{\ddagger}$ (P39A − WT)
ΔG_{eq} (*trans* → *cis*) [1]	1.4 kcal mol^{-1}	3.8 kcal mol^{-1}	2.4 kcal mol^{-1}
$\Delta G^{\ddagger}_{trans \rightarrow cis}$ [2]	21.1 kcal mol^{-1}	21.1 kcal mol^{-1}	0 kcal mol^{-1}
$\Delta G^{\ddagger}_{cis \rightarrow trans}$ [2]	19.7 kcal mol^{-1}	17.3 kcal mol^{-1}	−2.4 kcal mol^{-1}

Table 8.1b Free energy comparison between βPGM$_{WT}$ and βPGM$_{P146A}$ activity and stability parameters (this study).

	βPGM$_{WT}$	βPGM$_{P146A}$	$\Delta\Delta G^{\ddagger}_{kcat}$ or $\Delta\Delta G_{S}$ (βPGM$_{P146A}$ − βPGM$_{WT}$)
ΔG_{eq} (*trans* → *cis*) [1]	−0.5 kcal mol^{-1}	1.7 kcal mol^{-1} [3]	2.2 kcal mol^{-1}
$\Delta G^{\ddagger}_{kcat}$ [2]	13.9 kcal mol^{-1}	15.7 kcal mol^{-1}	1.8 kcal mol^{-1}
ΔG_{S} (βG16BP) [1]	−6.9 kcal mol^{-1}	−5.1 kcal mol^{-1}	1.8 kcal mol^{-1}
ΔG_{S} (βG1P) [1]	−5.5 kcal mol^{-1}	−5.2 kcal mol^{-1}	0.3 kcal mol^{-1}
ΔG_{i} (βG1P) [1]	−3.8 kcal mol^{-1}	−4.1 kcal mol^{-1}	−0.3 kcal mol^{-1}

[1] ΔG_{eq}, ΔG_{S} and ΔG_{i} were calculated using equation 1.1 from K_{eq}, K_{m} and K_{i} values, respectively.

[2] $\Delta G^{\ddagger}_{cis \rightarrow trans}$, $\Delta G^{\ddagger}_{trans \rightarrow cis}$ and $\Delta G^{\ddagger}_{kcat}$ were calculated using equation 1.3 from $k_{cis \rightarrow trans}$, $k_{trans \rightarrow cis}$ and k_{cat} values, respectively.

[3] Estimated using a $^{1}H^{15}N$–TROSY signal-to-noise value of at least 20 ($K_{eq} \leq 0.05$).
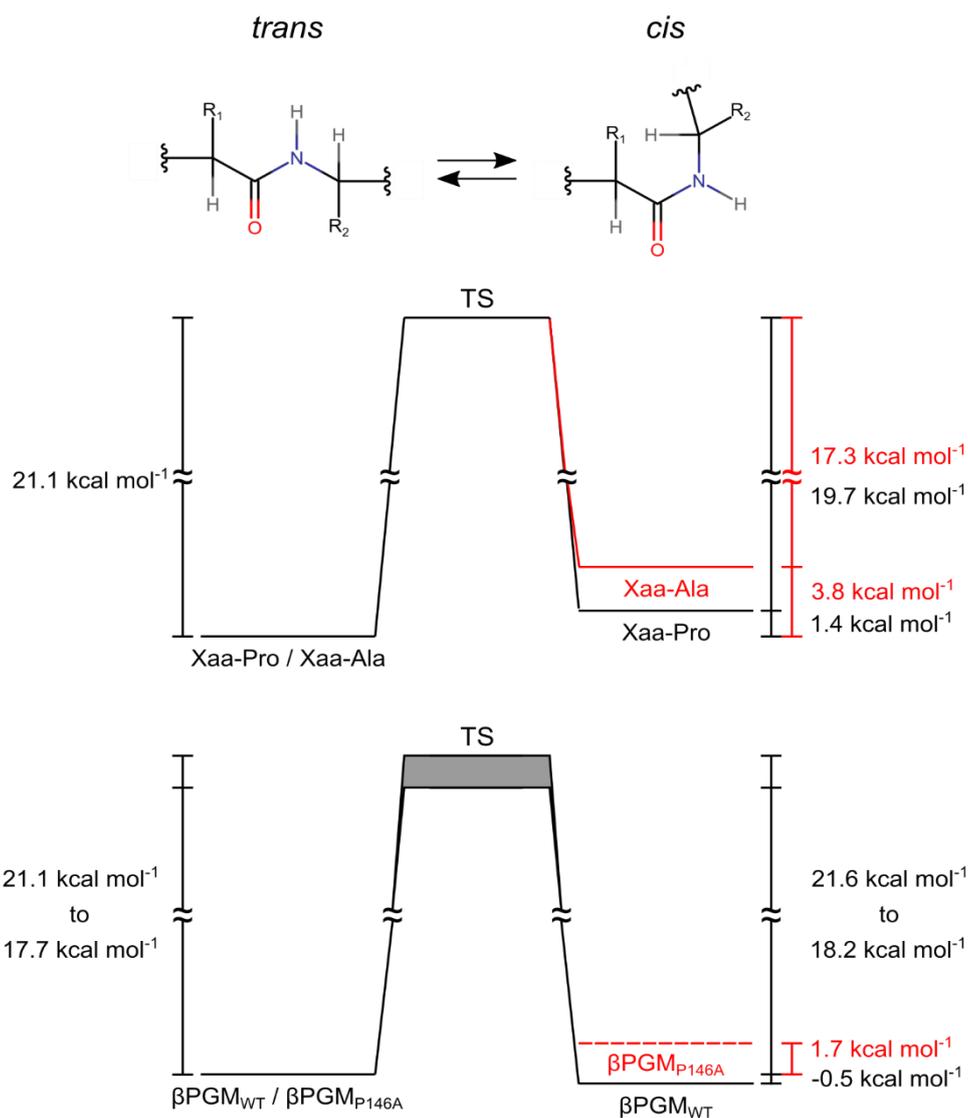
Figure 8.1 Correspondence between free energy reaction profiles of *cis-trans* isomerisation of Xaa-Pro and Xaa-Ala peptide bonds in RNase T1 and βPGM. (Top) Structural scheme of *cis-trans* peptide bond isomerisation. (Middle) Reaction profile of *cis-trans* peptide bond isomerisation in RNase T1 refolding (Odefey, Mayr and Schmid, 1995). The *trans* → *cis* isomerisation energy barriers of both Xaa-Pro and Xaa-Ala peptide bonds are 21.1 kcal mol^{-1}. The *cis* → *trans* isomerisation energy barrier of the Xaa-Pro peptide bond is 19.7 kcal mol^{-1}, while the same energy barrier for isomerisation of the Xaa-Ala peptide bond is 17.3 kcal mol^{-1}. The resulting equilibrium free energies are 1.4 and 3.8 kcal mol^{-1} for Xaa-Pro and Xaa-Ala peptide bonds respectively. (Bottom) Reaction profile of *cis-trans* peptide bond isomerisation in the βPGM catalytic cycle. The *trans* → *cis* (17.7 to 21.1 kcal mol^{-1}) and *cis* → *trans* (18.2 to 21.6 kcal mol^{-1}) isomerisation energy barriers in βPGM$_{WT}$ were estimated considering the interconversion rate between conformer B and conformer A of 0.003 to 1.0 s^{-1} and the measured equilibrium free energy of –0.5 kcal mol^{-1}. The shaded area is the estimated position of the transition state (TS). The equilibrium free energy of *cis-trans* alanine isomerisation in βPGM$_{P146A}$ of 1.7 kcal mol^{-1} is a lower bound estimate using a ^{1}H^{15}N–TROSY signal-to-noise value of 20 (red dashes). The difference in free energy between a *cis* Xaa-Ala and Xaa-Pro peptide bond in βPGM can be estimated as 2.2 kcal mol^{-1} (1.7 kcal mol^{-1} – (–0.5 kcal mol^{-1})). The equilibrium free energy values were calculated using equation 1.1. The free energy barriers were calculated using equation 1.3.

8.5 Future directions

i.   Investigation of the intermediate exchange dynamics in the substrate-free conformers and several GSA and TSA complexes is needed to identify its relationship with the rate-limiting step present in the βPGM catalytic cycle.

ii. The detailed structural mechanism with which βPGM is able to discriminate between allomorphic full and partial activators requires further investigation.

iii. Confirmation of the presence of allomorphic control in other hysteretic enzymes will broaden our understanding of allomorphy as a general mechanism of post-translational control of activity in monomeric enzymes.

iv. The efficient adoption of a catalytically competent conformation facilitated by substrate binding can be leveraged in the design of enzymes for difficult reactions not found in Nature.

v.   Identification of the detailed structural mechanism that allows βPGM to couple substrate binding, enzyme closure and *cis-trans* proline isomerisation is an exciting area of research that requires further investigation.

8.6 Conclusion

In conclusion, the conformational landscape of βPGM directly influences its catalysis and modulates its activity. It allows βPGM to redirect substrate binding energy efficiently to achieve an optimal chemical environment for catalysis. Its investigation uncovered a hitherto hidden post-translational control mechanism of enzyme activity that is likely to be a prevalent feature of Nature's toolbox. Further impact of this thesis will be strongest in the fields of 1) drug discovery and 2) synthetic biology. New drug targets may be exploitable with allomorphic inhibitors that stabilise low activity conformers. Allomorphy also provides an attractive structural design template for engineering proline switches in synthetic enzymes. The understanding of the conformational landscape of enzymes can be employed in the design of synthetic enzymes to generate novel chemicals or accelerate production of current molecules. Furthermore, distal sites can be incorporated into synthetic enzymes to fully leverage the substrate binding energy to catalyse difficult reactions not found in Nature.

9 References

Agarwal, P.K. (2005). Role of protein dynamics in reaction rate enhancement by enzymes. *Journal of the American Chemical Society.* **127**(43), pp.15248-15256.

Agarwal, P.K., Geist, A. and Gorin, A. (2004). Protein dynamics and enzymatic catalysis: Investigating the peptidyl-prolyl cis-trans isomerization activity of cyclophilin A. *Biochemistry.* **43**(33), pp.10605-10618.

Ainslie, G.R., Neet, K.E. and Shill, J.P. (1972). Transients and cooperativity - slow transition model for relating transients and cooperative kinetics of enzymes. *Journal of Biological Chemistry.* **247**(21), pp.7088-&.

Allen, K.N. and Dunaway-Mariano, D. (2004). Phosphoryl group transfer: evolution of a catalytic scaffold. *Trends in Biochemical Sciences.* **29**(9), pp.495-503.

Amyes, T.L., O'Donoghue, A.C. and Richard, J.P. (2001). Contribution of phosphate intrinsic binding energy to the enzymatic rate acceleration for triosephosphate isomerase. *Journal of the American Chemical Society.* **123**(45), pp.11325-11326.

Amyes, T.L. and Richard, J.P. (2013). Specificity in Transition State Binding: The Pauling Model Revisited. *Biochemistry.* **52**(12), pp.2021-2035.

Amyes, T.L., Richard, J.P. and Tait, J.J. (2005). Activation of orotidine 5 '-monophosphate decarboxylase by phosphite dianion: The whole substrate is the sum of two parts. *Journal of the American Chemical Society.* **127**(45), pp.15708-15709.

Andersson, U., Levander, F. and Radstrom, P. (2001). Trehalose-6-phosphate phosphorylase is part of a novel metabolic pathway for trehalose utilization in Lactococcus lactis. *Journal of Biological Chemistry.* **276**(46), pp.42707-42713.

Andersson, U., Molenaar, D., Radstrom, P. and de Vos, W.M. (2005). Unity in organisation and regulation of catabolic operons in Lactobacillus plantarum, Lactococcus lactis and Listeria monocytogenes. *Systematic and Applied Microbiology.* **28**(3), pp.187-195.

Andrews, L.D., Fenn, T.D. and Herschlag, D. (2013). Ground State Destabilization by Anionic Nucleophiles Contributes to the Activity of Phosphoryl Transfer Enzymes. *Plos Biology.* **11**(7), p18.

Barrozo, A., Liao, Q.H., Esguerra, M., Marloie, G., Florian, J., Williams, N.H. and Kamerlin, S.C.L. (2018). Computer simulations of the catalytic mechanism of wild-type and mutant beta-phosphoglucomutase. *Organic & Biomolecular Chemistry.* **16**(12), pp.2060-2073.

Battye, T.G.G., Kontogiannis, L., Johnson, O., Powell, H.R. and Leslie, A.G.W. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D-Biological Crystallography.* **67**, pp.271-281.

Baxter, N.J., Blackburn, G.M., Marston, J.P., Hounslow, A.M., Cliff, M.J., Bermel, W., Williams, N.H., Hollfelder, F., Wemmer, D.E. and Waltho, J.P. (2008). Anionic charge is prioritized over geometry in aluminum and magnesium fluoride transition state analogs of phosphoryl transfer enzymes. *Journal of the American Chemical Society.* **130**(12), pp.3952-3958.

Baxter, N.J., Bowler, M.W., Alizadeh, T., Cliff, M.J., Hounslow, A.M., Wu, B., Berkowitz, D.B., Williams, N.H., Blackburn, G.M. and Waltho, J.P. (2010). Atomic details of near-transition state conformers for enzyme phosphoryl transfer revealed by MgF3- rather than by phosphoranes. *Proceedings of the National Academy of Sciences of the United States of America.* **107**(10), pp.4555-4560.

Baxter, N.J., Hounslow, A.M., Bowler, M.W., Williams, N.H., Blackburn, G.M. and Waltho, J.P. (2009). MgF3- and alpha-Galactose 1-Phosphate in the Active Site of beta-Phosphoglucomutase Form a Transition State Analogue of Phosphoryl Transfer. *Journal of the American Chemical Society.* **131**(45), pp.16334-+.

Baxter, N.J., Olguin, L.F., Golicnik, M., Feng, G., Hounslow, A.M., Bermel, W., Blackburn, G.M., Hollfelder, F., Waltho, J.P. and Williams, N.H. (2006). A Trojan horse transition state analogue generated by MgF3- formation in an enzyme active site. *Proceedings of the National Academy of Sciences of the United States of America.* **103**(40), pp.14732-14737.

Baxter, N.J., Roetzer, A., Liebig, H.D., Sedelnikova, S.E., Hounslow, A.M., Skern, T. and Waltho, J.P. (2006). Structure and dynamics of coxsackievirus B4 2A proteinase, an enyzme involved in the etiology of heart disease. *Journal of Virology.* **80**(3), pp.1451-1462.

Baxter, N.J., Zacharchenko, T., Barsukov, I.L. and Williamson, M.P. (2017). Pressure-Dependent Chemical Shifts in the R3 Domain of Talin Show that It Is Thermodynamically Poised for Binding to Either Vinculin or RIAM. *Structure.* **25**(12), pp.1856-+.

Benkovic, S.J. and Hammes-Schiffer, S. (2003). A perspective on enzyme catalysis. *Science.* **301**(5637), pp.1196-1202.

Bennett, B.D., Kimball, E.H., Gao, M., Osterhout, R., Van Dien, S.J. and Rabinowitz, J.D. (2009). Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nature Chemical Biology.* **5**(8), pp.593-599.

Berjanskii, M.V. and Wishart, D.S. (2005). A Simple Method To Predict Protein Flexibility Using Secondary Chemical Shifts. *J. Am. Chem. Soc.* **127**(43), pp.14970-14971.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. (2000). The Protein Data Bank. *Nucleic Acids Res.* **28**(1), pp.235-242.

Bhabha, G., Biel, J.T. and Fraser, J.S. (2015). Keep on Moving: Discovering and Perturbing the Conformational Dynamics of Enzymes. *Accounts of Chemical Research.* **48**(2), pp.423-430.

Bodor, A., Toth, I., Banyai, I., Szabo, Z. and Hefter, G.T. (2000). F-19 NMR study of the equilibria and dynamics of the Al3+/F- system. *Inorganic Chemistry.* **39**(12), pp.2530-2537.

Boehr, D.D., McElheny, D., Dyson, H.J. and Wright, P.E. (2006). The dynamic energy landscape of dihydrofolate reductase catalysis. *Science.* **313**(5793), pp.1638-1642.

Boonstra, J. and Verkleij, A.J. (2004). Regulation of enzyme activity *in vivo* is determined by its cellular localization. *Advances in Enzyme Regulation, Vol 44.* **44**, pp.61-73.

Bowler, M.W., Cliff, M.J., Waltho, J.P. and Blackburn, G.M. (2010). Why did Nature select phosphate for its dominant roles in biology? *New Journal of Chemistry.* **34**(5), pp.784-794.

Bruice, T.C. (1976). Some pertinent aspects of mechanism as determined with small molecules. *Annual Review of Biochemistry.* **45**, pp.331-373.

Burschowsky, D., van Eerde, A., Okvist, M., Kienhofer, A., Kast, P., Hilvert, D. and Krengel, U. (2014). Electrostatic transition state stabilization rather than reactant destabilization provides the chemical basis for efficient chorismate mutase catalysis. *Proceedings of the National Academy of Sciences of the United States of America.* **111**(49), pp.17516-17521.

Carey, P.R. (2006). Spectroscopic Characterization of Distortion in Enzyme Complexes. *Chem. Rev.* **106**(8), pp.3043-3054.

Castellana, M., Wilson, M.Z., Xu, Y., Joshi, P., Cristea, I.M., Rabinowitz, J.D., Gitai, Z. and Wingreen, N.S. (2014). Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nature Biotechnology.* **32**(10), pp.1011-+.

Cavanagh, J. (2007). *Protein NMR spectroscopy : principles and practice.* 2nd ed. Amsterdam; Boston: Academic Press.

Chae, Y.K. and Markley, J.L. (2000). Functional recombinant rabbit muscle phosphoglucomutase from Escherichia coli. *Protein Expression and Purification.* **20**(1), pp.124-127.

Chalissery, J., Banerjee, S., Bandey, I. and Sen, R. (2007). Transcription termination defective mutants of Rho: role of different functions of rho in releasing RNA from the elongation complex. *Journal of Molecular Biology.* **371**(4), pp.855-872.

Changeux, J.-P. (2012). Allostery and the Monod-Wyman-Changeux Model After 50 Years. *Annu Rev Biophys.* **41**(1), pp.103-133.

Changeux, J.-P. (2013). 50 years of allosteric interactions: the twists and turns of the models. *Nature Reviews Molecular Cell Biology.* **14**(12), pp.819-829.

Chaptal, V., Vincent, F., Gueguen-Chaignon, V., Monedero, V., Poncet, S., Deutscher, J., Nessler, S. and Morera, S. (2007). Structural analysis of the bacterial HPr kinase/phosphorylase V267F mutant gives insights into the allosteric regulation mechanism of this bifunctional enzyme. *Journal of Biological Chemistry.* **282**(48), pp.34952-34957.

Chen, V.B., Arendall, W.B., III, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D-Structural Biology.* **66**, pp.12-21.

Ciechanover, A. (2005). Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nature Reviews Molecular Cell Biology.* **6**(1), pp.79-86.

Cliff, M.J., Bowler, M.W., Varga, A., Marston, J.P., Szabo, J., Hounslow, A.M., Baxter, N.J., Blackburn, G.M., Vas, M. and Waltho, J.P. (2010). Transition State Analogue Structures of Human Phosphoglycerate Kinase Establish the Importance of Charge Balance in Catalysis. *Journal of the American Chemical Society.* **132**(18), pp.6507-6516.

Codreanu, S.G., Ladner, J.E., Xiao, G.Y., Stourman, N.V., Hachey, D.L., Gilliland, G.L. and Armstrong, R.N. (2002). Local protein dynamics and catalysis: Detection of segmental motion

associated with rate-limiting product release by a glutathione transferase. *Biochemistry.* **41**(51), pp.15161-15172.

Corazza, A., Rennella, E., Schanda, P., Mimmi, M.C., Cutuil, T., Raimondi, S., Giorgetti, S., Fogolari, F., Viglino, P., Frydman, L., Gal, M., Bellotti, V., Brutscher, B. and Esposito, G. (2010). Native-unlike Long-lived Intermediates along the Folding Pathway of the Amyloidogenic Protein beta(2)-Microglobulin Revealed by Real-time Two-dimensional NMR. *Journal of Biological Chemistry.* **285**(8), pp.5827-5835.

Cornish-Bowden, A. (1995). Fundamentals of enzyme kinetics. *Fundamentals of enzyme kinetics.* pxiii+343p.

Crean, R.M., Gardner, J.M. and Kamerlin, S.C.L. (2020). Harnessing Conformational Plasticity to Generate Designer Enzymes. *Journal of the American Chemical Society.* **142**(26), pp.11324-11342.

Crick, F. (1970). Central Dogma of Molecular Biology. *Nature.* **227**(5258), pp.561-563.

Cruz-Navarrete, F.A., Baxter, N.J., Wood, H.P., Hounslow, A.M. and Waltho, J.P. (2019). H-1, N-15 and C-13 backbone resonance assignments of the P146A variant of beta-phosphoglucomutase from Lactococcus lactis in its substrate-free form. *Biomolecular Nmr Assignments.* **13**(2), pp.349-356.

Dai, J., Finci, L., Zhang, C., Lahiri, S., Zhang, G., Peisach, E., Allen, K.N. and Dunaway-Mariano, D. (2009). Analysis of the Structural Determinants Underlying Discrimination between Substrate and Solvent in beta-Phosphoglucomutase Catalysis. *Biochemistry.* **48**(9), pp.1984-1995.

Dai, J., Wang, L., Allen, K.N., Radstrom, P. and Dunaway-Mariano, D. (2006). Conformational cycling in beta-phosphoglucomutase catalysis: Reorientation of the beta-D-glucose 1,6-(bis) phosphate intermediate. *Biochemistry.* **45**(25), pp.7818-7824.

De Dios, A.C., Pearson, J.G. and Oldfield, E. (1993). Secondary and Tertiary Structural Effects on Protein NMR Chemical Shifts: An ab Initio Approach. *Science.* **260**(5113), pp.1491-1496.

de Graffenried, C.L. and Bertozzi, C.R. (2004). The roles of enzyme localisation and complex formation in glycan assembly within the Golgi apparatus. *Current Opinion in Cell Biology.* **16**(4), pp.356-363.

Dill, K.A., Ozkan, S.B., Shell, M.S. and Weikl, T.R. (2008). The protein folding problem. *Annu Rev Biophys.* **37**(1), pp.289-316.

Doshi, U., McGowan, L.C., Ladani, S.T. and Hamelberg, D. (2012). Resolving the complex role of enzyme conformational dynamics in catalytic function. *Proceedings of the National Academy of Sciences of the United States of America.* **109**(15), pp.5699-5704.

Elsasser, B., Dohmeier-Fischer, S. and Fels, G. (2012). Theoretical investigation of the enzymatic phosphoryl transfer of beta-phosphoglucomutase: revisiting both steps of the catalytic cycle. *Journal of Molecular Modeling.* **18**(7), pp.3169-3179.

Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D-Biological Crystallography.* **66**, pp.486-501.

Evans, P.R. and Murshudov, G.N. (2013). How good are my data and what is the resolution? *Acta Crystallographica Section D-Biological Crystallography.* **69**, pp.1204-1214.

Eyring, H. (1935). The Activated Complex in Chemical Reactions. *The Journal of chemical physics.* **3**(2), pp.107-115.

Felli, I.C. and Brutscher, B. (2009). Recent Advances in Solution NMR: Fast Methods and Heteronuclear Direct Detection. *Chemphyschem.* **10**(9-10), pp.1356-1368.

Finlay, C.C., Maus, S., Beggan, C.D., Bondar, T.N., Chambodut, A., Chernova, T.A., Chulliat, A., Golovkov, V.P., Hamilton, B., Hamoudi, M., Holme, R., Hulot, G., Kuang, W., Langlais, B., Lesur, V., Lowes, F.J., Lühr, H., Macmillan, S., Mandea, M., McLean, S., Manoj, C., Menvielle, M., Michaelis, I., Olsen, N., Rauberg, J., Rother, M., Sabaka, T.J., Tangborn, A., Tøffner-Clausen, L., Thébault, E., Thomson, A.W.P., Wardinski, I., Wei, Z. and Zvereva, T.I. (2010). International Geomagnetic Reference Field: the eleventh generation. *Geophys. J. Int.* **183**(3), pp.1216-1230.

Flynn, P.F., Mattiello, D.L., Hill, H.D.W. and Wand, A.J. (2000). Optimal Use of Cryogenic Probe Technology in NMR Studies of Proteins. *J. Am. Chem. Soc.* **122**(19), pp.4823-4824.

Foster, M.P., McElroy, C.A. and Amero, C.D. (2007). Solution NMR of large molecules and assemblies. *Biochemistry.* **46**(2), pp.331-340.

Franco, R., Favier, A., Schanda, P. and Brutscher, B. (2017). Optimized fast mixing device for real-time NMR applications. *J Magn Reson.* **281**, pp.125-129.

Frauenfelder, H., Sligar, S.G. and Wolynes, P.G. (1991). The Energy Landscapes and Motions of Proteins. *Science.* **254**(5038), pp.1598-1603.

Frieden, C. (1979). Slow transitions and hysteretic behavior in enzymes. *Annual Review of Biochemistry.* **48**, pp.471-489.

Garcia-Viloca, M., Gao, J., Karplus, M. and Truhlar, D.G. (2004). How Enzymes Work: Analysis by Modern Rate Theory and Computer Simulations. *Science.* **303**(5655), pp.186-195.

Gardner, J.M., Biler, M., Risso, V.A., Sanchez-Ruiz, J.M. and Kamerlin, S.C.L. (2020). Manipulating Conformational Dynamics To Repurpose Ancient Proteins for Modern Catalytic Functions. *Acs Catalysis.* **10**(9), pp.4863-4870.

Gardner, K.H. and Kay, L.E. (1998). The use of H-2, C-13, N-15 multidimensional NMR to study the structure and dynamics of proteins. *Annual Review of Biophysics and Biomolecular Structure.* **27**, pp.357-406.

Go, M.K., Koudelka, A., Amyes, T.L. and Richard, J.P. (2010). Role of Lys-12 in Catalysis by Triosephosphate Isomerase: A Two-Part Substrate Approach. *Biochemistry.* **49**(25), pp.5377-5389.

Goel, A., Santos, F., de Vos, W.M., Teusink, B. and Molenaar, D. (2012). Standardized Assay Medium To Measure Lactococcus lactis Enzyme Activities while Mimicking Intracellular Conditions. *Applied and Environmental Microbiology.* **78**(1), pp.134-143.

Golicnik, M., Olguin, L.F., Feng, G., Baxter, N.J., Waltho, J.P., Williams, N.H. and Hollfelder, F. (2009). Kinetic Analysis of beta-Phosphoglucomutase and Its Inhibition by Magnesium Fluoride. *Journal of the American Chemical Society.* **131**(4), pp.1575-1588.

Goryanova, B., Goldman, L.M., Amyes, T.L., Gerlt, J.A. and Richard, J.P. (2013). Role of a Guanidinium Cation-Phosphodianion Pair in Stabilizing the Vinyl Carbanion Intermediate of Orotidine 5 '-Phosphate Decarboxylase-Catalyzed Reactions. *Biochemistry.* **52**(42), pp.7500-7511.

Grathwohl, C. and Wüthrich, K. (1981). Nmr studies of the rates of proline cis-trans isomerization in oligopeptides. *Biopolymers.* **20**(12), pp.2623-2633.

Griffin, J.L., Bowler, M.W., Baxter, N.J., Leigh, K.N., Dannatt, H.R.W., Hounslow, A.M., Blackburn, G.M., Webster, C.E., Cliff, M.J. and Waltho, J.P. (2012). Near attack conformers dominate beta-phosphoglucomutase complexes where geometry and charge distribution reflect those of substrate. *Proceedings of the National Academy of Sciences of the United States of America.* **109**(18), pp.6910-6915.

Hall, A. and Knowles, J.R. (1975). Uncatalyzed rates of enolization of dihydroxyacetone phosphate and of glyceraldehyde-3-phosphate in neutral aqueous-solution - quantitative assessment of effectiveness of an enzyme catalyst. *Biochemistry.* **14**(19), pp.4348-4352.

Hammes-Schiffer, S. and Benkovic, S.J. (2006). Relating protein motion to catalysis. *Annual Review of Biochemistry.* **75**, pp.519-541.

Harper, S.M., Neil, L.C., Day, I.J., Hore, P.J. and Gardner, K.H. (2004). Conformational changes in a photosensory LOV domain monitored by time-resolved NMR spectroscopy. *Journal of the American Chemical Society.* **126**(11), pp.3390-3391.

Hayward, S. and Berendsen, H.J.C. (1998). Systematic analysis of domain motions in proteins from conformational change: New results on citrate synthase and T4 lysozyme. *Proteins, structure, function, and bioinformatics.* **30**(2), pp.144-154.

Henzler-Wildman, K. and Kern, D. (2007). Dynamic personalities of proteins. *Nature.* **450**(7172), pp.964-972.

Hilser, V.J., Anderson, J.A. and Motlagh, H.N. (2015). Allostery vs. "allokairy". *Proceedings of the National Academy of Sciences of the United States of America.* **112**(37), pp.11430-11431.

Holland, D.J., Bostock, M.J., Gladden, L.F. and Nietlispach, D. (2011). Fast Multidimensional NMR Spectroscopy Using Compressed Sensing. *Angew. Chem. Int. Ed.* **50**(29), pp.6548-6551.

Hong, B.S., Senisterra, G., Rabeh, W.M., Vedadi, M., Leonardi, R., Zhang, Y.M., Rock, C.O., Jackowski, S. and Park, H.W. (2007). Crystal structures of human pantothenate kinases - Insights into allosteric regulation and mutations linked to a neurodegeneration disordew. *Journal of Biological Chemistry.* **282**(38), pp.27984-27993.

Huang, X.Y., Holden, H.M. and Raushel, F.M. (2001). Channeling of substrates and intermediates in enzyme-catalyzed reactions. *Annual Review of Biochemistry.* **70**, pp.149-180.

Huang, Y. and Breitwieser, G.E. (2007). Rescue of calcium-sensing receptor mutants by allosteric modulators reveals a conformational checkpoint in receptor biogenesis. *Journal of Biological Chemistry.* **282**(13), pp.9517-9525.

Hur, S. and Bruice, T.C. (2002). The mechanism of catalysis of the chorismate to prephenate reaction by the Escherichia coli mutase enzyme. *Proceedings of the National Academy of Sciences of the United States of America.* **99**(3), pp.1176-1181.

Hur, S. and Bruice, T.C. (2003). The near attack conformation approach to the study of the chorismate to prephenate reaction. *Proceedings of the National Academy of Sciences of the United States of America.* **100**(21), pp.12015-12020.

Hyberts, S.G., Milbradt, A.G., Wagner, A.B., Arthanari, H. and Wagner, G. (2012). Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *Journal of Biomolecular Nmr.* **52**(4), pp.315-327.

Hyberts, S.G., Robson, S.A. and Wagner, G. (2013). Exploring signal-to-noise ratio and sensitivity in non-uniformly sampled multi-dimensional NMR spectra. *Journal of Biomolecular Nmr.* **55**(2), pp.167-178.

Hyberts, S.G., Takeuchi, K. and Wagner, G. (2010). Poisson-Gap Sampling and Forward Maximum Entropy Reconstruction for Enhancing the Resolution and Sensitivity of Protein NMR Data. *J. Am. Chem. Soc.* **132**(7), pp.2145-2147.

Iommarini, L., Ghelli, A., Gasparre, G. and Porcelli, A.M. (2017). Mitochondrial metabolism and energy sensing in tumor progression. *Biochimica Et Biophysica Acta-Bioenergetics.* **1858**(8), pp.582-590.

Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in synthesis of proteins. *Journal of Molecular Biology.* **3**(3), pp.318-+.

James, L.C. and Tawfik, D.S. (2003). Conformational diversity and protein evolution - a 60-year-old hypothesis revisited. *Trends in Biochemical Sciences.* **28**(7), pp.361-368.

Jencks, W.P. (1975). Binding-energy, specificity, and enzymic catalysis - circe effect. *Advances in Enzymology and Related Areas of Molecular Biology.* **43**, pp.219-410.

Jin, Y., Bhattasali, D., Pellegrini, E., Forget, S.M., Baxter, N.J., Cliff, M.J., Bowler, M.W., Jakeman, D.L., Blackburn, G.M. and Waltho, J.P. (2014). alpha-Fluorophosphonates reveal how a phosphomutase conserves transition state conformation over hexose recognition in its two-step reaction. *Proceedings of the National Academy of Sciences of the United States of America.* **111**(34), pp.12384-12389.

Jin, Y., Molt, R.W. and Blackburn, G.M. (2017). Metal Fluorides: Tools for Structural and Computational Analysis of Phosphoryl Transfer Enzymes. *Topics in Current Chemistry.* **375**(2).

Jin, Y., Molt, R.W., Jr., Waltho, J.P., Richards, N.G.J. and Blackburn, G.M. (2016). 19FNMR and DFT Analysis Reveal Structural and Electronic Transition State Features for RhoA-Catalyzed GTP Hydrolysis. *Angewandte Chemie - International Edition.* **55**(10), pp.3318-3322.

Johnson, L.A., Robertson, A.J., Baxter, N.J., Trevitt, C.R., Bisson, C., Jin, Y., Wood, H.P., Hounslow, A.M., Cliff, M.J., Blackburn, G.M., Bowler, M.W. and Waltho, J.P. (2018). van der Waals Contact between Nucleophile and Transferring Phosphorus Is Insufficient To Achieve Enzyme Transition-State Architecture. *Acs Catalysis.* **8**(9), pp.8140-8153.

Kamata, K., Mitsuya, M., Nishimura, T., Eiki, J.-i. and Nagata, Y. (2004). Structural Basis for Allosteric Regulation of the Monomeric Allosteric Enzyme Human Glucokinase. *Structure.* **12**(3), pp.429-438.

Kamerlin, S.C.L. and Warshel, A. (2010). At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis? *Proteins-Structure Function and Bioinformatics.* **78**(6), pp.1339-1375.

Kern, D., Karplus, M., Thai, V., Lei, M., Kerns, S.J. and Henzler-Wildman, K.A. (2007). A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature.* **450**(7171), pp.913-916.

Kleckner, I.R. and Foster, M.P. (2011). An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica Et Biophysica Acta-Proteins and Proteomics.* **1814**(8), pp.942-968.

Klein, A.H., Shulla, A., Reimann, S.A., Keating, D.H. and Wolfe, A.J. (2007). The intracellular concentration of acetyl phosphate in Escherichia coli is sufficient for direct phosphorylation of two-component response regulators. *Journal of Bacteriology.* **189**(15), pp.5574-5581.

Klyuyeva, A., Tuganova, A. and Popov, K.M. (2008). Allosteric coupling in pyruvate dehydrogenase kinase 2. *Biochemistry.* **47**(32), pp.8358-8366.

Koshland, D.E. (1960). The active site and enzyme action. *Advances in Enzymology and Related Subjects of Biochemistry.* **22**, pp.45-97.

Koshland, D.E., Nemethy, G. and Filmer, D. (1966). Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry.* **5**(1), pp.365-385.

Krebs, E.G. and Beavo, J.A. (1979). Phosphorylation-dephosphorylation of enzymes. *Annual Review of Biochemistry.* **48**, pp.923-959.

Kukic, P., Farrell, D., McIntosh, L.P., Garcia-Moreno, E.B., Jensen, K.S., Toleikis, Z., Teilum, K. and Nielsen, J.E. (2013). Protein Dielectric Constants Determined from NMR Chemical Shift Perturbations. *Journal of the American Chemical Society.* **135**(45), pp.16968-16976.

Kulkarni, Y.S., Liao, Q.H., Bylehn, F., Amyes, T.L., Richard, J.P. and Kamerlin, S.C.L. (2018). Role of Ligand-Driven Conformational Changes in Enzyme Catalysis: Modeling the Reactivity of the Catalytic Cage of Triosephosphate Isomerase. *Journal of the American Chemical Society.* **140**(11), pp.3854-3857.

Kwan, A.H., Mobli, M., Gooley, P.R., King, G.F. and Mackay, J.P. (2011). Macromolecular NMR spectroscopy for the non-spectroscopist. *FEBS J.* **278**(5), pp.687-703.

Lad, C., Williams, N.H. and Wolfenden, R. (2003). The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases. *Proceedings of the National Academy of Sciences of the United States of America.* **100**(10), pp.5607-5610.

Lahiri, S.D., Zhang, G.F., Dai, J.Y., Dunaway-Mariano, D. and Allen, K.N. (2004). Analysis of the substrate specificity loop of the HAD superfamily cap domain. *Biochemistry.* **43**(10), pp.2812-2820.

Lahiri, S.D., Zhang, G.F., Dunaway-Mariano, D. and Allen, K.N. (2002a). Caught in the act: The structure of phosphorylated beta-phosphoglucomutase from Lactococcus lactis. *Biochemistry.* **41**(26), pp.8351-8359.

Lahiri, S.D., Zhang, G.F., Dunaway-Mariano, D. and Allen, K.N. (2003). The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction. *Science.* **299**(5615), pp.2067-2071.

Lahiri, S.D., Zhang, G.F., Radstrom, P., Dunaway-Mariano, D. and Allen, K.N. (2002b). Crystallization and preliminary X-ray diffraction studies of beta-phosphoglucomutase from Lactococcus lactus. *Acta Crystallographica Section D-Biological Crystallography.* **58**, pp.324-326.

Larion, M., Hansen, A.L., Zhang, F.L., Bruschweiler-Li, L., Tugarinov, V., Miller, B.G. and Bruschweiler, R. (2015). Kinetic Cooperativity in Human Pancreatic Glucokinase Originates from Millisecond Dynamics of the Small Domain. *Angewandte Chemie-International Edition.* **54**(28), pp.8129-8132.

Larion, M. and Miller, B.G. (2012). Homotropic allosteric regulation in monomeric mammalian glucokinase. *Arch Biochem Biophys.* **519**(2), pp.103-111.

Larion, M., Salinas, R.K., Bruschweiler-Li, L., Miller, B.G. and Brueschweiler, R. (2012). Order-Disorder Transitions Govern Kinetic Cooperativity and Allostery of Monomeric Human Glucokinase. *Plos Biology.* **10**(12).

Lee, J. and Goodey, N.M. (2011). Catalytic Contributions from Remote Regions of Enzyme Structure. *Chemical Reviews.* **111**(12), pp.7595-7624.

Levander, F., Andersson, U. and Radstrom, P. (2001). Physiological role of beta-phosphoglucomutase in Lactococcus lactic. *Applied and Environmental Microbiology.* **67**(10), pp.4546-4553.

Levitt, M.H. (2008). *Spin dynamics : basics of nuclear magnetic resonance.* 2nd ed. Chichester, England ; Hoboken, NJ: John Wiley & Sons.

Lewis, M. (2005). The lac repressor. *Comptes Rendus Biologies.* **328**(6), pp.521-548.

Liang, Z.X., Lee, T., Resing, K.A., Ahn, N.G. and Klinman, J.P. (2004). Thermal-activated protein mobility and its correlation with catalysis in thermophilic alcohol dehydrogenase. *Proceedings of the National Academy of Sciences of the United States of America.* **101**(26), pp.9556-9561.

Lightstone, F.C. and Bruice, T.C. (1996). Ground state conformations and entropic and enthalpic factors in the efficiency of intramolecular and enzymatic reactions .1. Cyclic anhydride formation by substituted glutarates, succinate, and 3,6-endoxo-Delta(4)-tetrahydrophthalate monophenyl esters. *Journal of the American Chemical Society.* **118**(11), pp.2595-2605.

Liu, X.Y., Lu, S.Y., Song, K., Shen, Q.C., Ni, D., Li, Q., He, X.H., Zhang, H., Wang, Q., Chen, Y.Y., Li, X.Y., Wu, J., Sheng, C.Q., Chen, G.Q., Liu, Y.Q., Lu, X.F. and Zhang, J. (2020). Unraveling allosteric landscapes of allosterome with ASD. *Nucleic Acids Research.* **48**(D1), pp.D394-D401.

Lu, K.P., Finn, G., Lee, T.H. and Nicholson, L.K. (2007). Prolyl cis-trans isomerization as a molecular timer. *Nature Chemical Biology.* **3**(10), pp.619-629.

Malabanan, M.M., Amyes, T.L. and Richard, J.P. (2010). A role for flexible loops in enzyme catalysis. *Current Opinion in Structural Biology.* **20**(6), pp.702-710.

Marcos, E., Field, M.J. and Crehuet, R. (2010). Pentacoordinated phosphorus revisited by high-level QM/MM calculations. *Proteins-Structure Function and Bioinformatics.* **78**(11), pp.2405-2411.

McNicholas, S., Potterton, E., Wilson, K.S. and Noble, M.E.M. (2011). Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallographica Section D-Structural Biology.* **67**, pp.386-394.

Midelfort, C.F., Gupta, R.K. and Rose, I.A. (1976). Fructose 1,6-bisphosphate - isomeric composition, kinetics and substrate-specificity for aldolases. *Biochemistry.* **15**(10), pp.2178-2185.

Miller, B.G., Hassell, A.M., Wolfenden, R., Milburn, M.V. and Short, S.A. (2000). Anatomy of a proficient enzyme: The structure of orotidine 5 '-monophosphate decarboxylase in the presence and absence of a potential transition state analog. *Proceedings of the National Academy of Sciences of the United States of America.* **97**(5), pp.2011-2016.

Miller, B.G. and Wolfenden, R. (2002). Catalytic proficiency: The unusual case of OMP decarboxylase. *Annual Review of Biochemistry.* **71**, pp.847-885.

Miller, G.P., Wahnon, D.C. and Benkovic, S.J. (2001). Interloop contacts modulate ligand cycling during catalysis by Escherichia coli dihydrofolate reductase. *Biochemistry.* **40**(4), pp.867-875.

Mittermaier, A.K. and Kay, L.E. (2009). Observing biological dynamics at atomic resolution using NMR. *Trends in Biochemical Sciences.* **34**(12), pp.601-611.

Monod, J., Changeux, J.P. and Jacob, F. (1963). Allosteric proteins and cellular control systems. *Journal of Molecular Biology.* **6**(4), pp.306-&.

Monod, J., Wyman, J. and Changeux, J.P. (1965). On nature of allosteric transitions - a plausible model. *Journal of Molecular Biology.* **12**(1), pp.88-&.

Morrow, J.R., Amyes, T.L. and Richard, J.P. (2008). Phosphate Binding Energy and Catalysis by Small and Large Molecules. *Acc. Chem. Res.* **41**(4), pp.539-548.

Motlagh, H.N., Wrabl, J.O., Li, J. and Hilser, V.J. (2014). The ensemble nature of allostery. *Nature.* **508**(7496), pp.331-339.

Moult, J. and Melamud, E. (2000). From fold to function. *Current Opinion in Structural Biology.* **10**(3), pp.384-389.

Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D-Structural Biology.* **53**, pp.240-255.

Mydy, L.S., Cristobal, J.R., Katigbak, R.D., Bauer, P., Reyes, A.C., Kamerlin, S.C.L., Richard, J.P. and Gulick, A.M. (2019). Human Glycerol 3-Phosphate Dehydrogenase: X-ray Crystal Structures That Guide the Interpretation of Mutagenesis Studies. *Biochemistry.* **58**(8), pp.1061-1073.

Nashine, V.C., Hammes-Schiffer, S. and Benkovic, S.J. (2010). Coupled motions in enzyme catalysis. *Current Opinion in Chemical Biology.* **14**(5), pp.644-651.

Neves, A.R., Pool, W.A., Kok, J., Kuipers, O.P. and Santos, H. (2005). Overview on sugar metabolism and its control in Lactococcus lactis - The input from *in vivo* NMR. *Fems Microbiology Reviews.* **29**(3), pp.531-554.

Odefey, C., Mayr, L.M. and Schmid, F.X. (1995). Non-prolyl cis-trans peptide-bond isomerization as a rate-determining step in protein unfolding and refolding. *Journal of Molecular Biology.* **245**(1), pp.69-78.

Oesterhelt, C., Schnarrenberger, C. and Gross, W. (1997). The reaction mechanism of phosphomannomutase in plants. *Febs Letters.* **401**(1), pp.35-37.

Orengo, C.A., Todd, A.E. and Thornton, J.M. (1999). From protein structure to function. *Curr Opin Struct Biol.* **9**(3), pp.374-382.

Pardee, A.B. (2006). Regulatory molecular biology. *Cell Cycle.* **5**(8), pp.846-852.

Pisliakov, A.V., Cao, J., Kamerlin, S.C.L. and Warshel, A. (2009). Enzyme Millisecond Conformational Dynamics Do Not Catalyze the Chemical Step. *Proc Natl Acad Sci U S A.* **106**(41), pp.17359-17364.

Preiswerk, N., Beck, T., Schulz, J.D., Milovnik, P., Mayer, C., Siegel, J.B., Baker, D. and Hilvert, D. (2014). Impact of scaffold rigidity on the design and evolution of an artificial Diels-Alderase. *Proceedings of the National Academy of Sciences of the United States of America.* **111**(22), pp.8013-8018.

Qian, N., Stanley, G.A., Bunte, A. and Radstrom, P. (1997). Product formation and phosphoglucomutase activities in Lactococcus lactis: Cloning and characterization of a novel phosphoglucomutase gene. *Microbiology-Sgm.* **143**, pp.855-865.

Qian, N., Stanley, G.A., Hahnhagerdal, B. and Radstrom, P. (1994). Purification and characterization of 2 phosphoglucomutases from lactococcus-lactis subsp lactis and their regulation in maltose-utilizing and glucose-utilizing cells. *Journal of Bacteriology.* **176**(17), pp.5304-5311.

Read, R.J. and Schierbeek, A.J. (1988). A phased translation function. *Journal of Applied Crystallography.* **21**, pp.490-495.

Redfern, O.C., Dessailly, B. and Orengo, C.A. (2008). Exploring the structure and function paradigm. *Current Opinion in Structural Biology.* **18**(3), pp.394-402.

Reed, M.A.C., Hounslow, A.M., Sze, K.H., Barsukov, I.G., Hosszu, L.L.P., Clarke, A.R., Craven, C.J. and Waltho, J.P. (2003). Effects of Domain Dissection on the Folding and

Stability of the 43 kDa Protein PGK Probed by NMR. *Journal of Molecular Biology.* **330**(5), pp.1189-1201.

Reyes, A.C., Amyes, T.L. and Richard, J.P. (2018). Primary Deuterium Kinetic Isotope Effects: A Probe for the Origin of the Rate Acceleration for Hydride Transfer Catalyzed by Glycerol-3-Phosphate Dehydrogenase. *Biochemistry.* **57**(29), pp.4338-4348.

Reyes, A.C., Koudelka, A.P., Amyes, T.L. and Richard, J.P. (2015a). Enzyme Architecture: Optimization of Transition State Stabilization from a Cation-Phosphodianion Pair. *Journal of the American Chemical Society.* **137**(16), pp.5312-5315.

Reyes, A.C., Zhai, X., Morgan, K.T., Reinhardt, C.J., Amyes, T.L. and Richard, J.P. (2015b). The Activating Oxydianion Binding Domain for Enzyme-Catalyzed Proton Transfer, Hydride Transfer, and Decarboxylation: Specificity and Enzyme Architecture. *Journal of the American Chemical Society.* **137**(3), pp.1372-1382.

Ricard, J., Meunier, J.C. and Buc, J. (1974). Regulatory behavior of monomeric enzymes .1. Mnemonical enzyme concept. *European Journal of Biochemistry.* **49**(1), pp.195-208.

Richard, J.P., Amyes, T.L., Goryanova, B. and Zhai, X. (2014). Enzyme architecture: on the importance of being in a protein cage. *Current Opinion in Chemical Biology.* **21**, pp.1-10.

Richard, J.P., Amyes, T.L. and Reyes, A.C. (2018). Orotidine 5 '-Monophosphate Decarboxylase: Probing the Limits of the & IT;Possible & IT; for Enzyme Catalysis. *Accounts of Chemical Research.* **51**(4), pp.960-969.

Richard, J.P. (2019). Protein flexibility and stiffness enable efficient enzymatic catalysis. *J. Am. Chem. Soc.* **141**(8), pp.3320–3331.

Ruschak, A.M. and Kay, L.E. (2010). Methyl groups as probes of supra-molecular structure, dynamics and function. *Journal of Biomolecular Nmr.* **46**(1), pp.75-87.

Ruscio, J.Z., Kohn, J.E., Ball, K.A. and Head-Gordon, T. (2009). The Influence of Protein Dynamics on the Success of Computational Enzyme Design. *Journal of the American Chemical Society.* **131**(39), pp.14111-14115.

Scherer, M., Fleishman, S.J., Jones, P.R., Dandekar, T. and Bencurova, E. (2021). Computational Enzyme Engineering Pipelines for Optimized Production of Renewable Chemicals. *Frontiers in Bioengineering and Biotechnology.* **9**.

Schimke, R.T. and Doyle, D. (1970). Control of enzyme levels in animal tissues. *Annual Review of Biochemistry.* **39**, pp.929-+.

Schnell, J.R., Dyson, H.J. and Wright, P.E. (2004). Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annual Review of Biophysics and Biomolecular Structure.* **33**, pp.119-140.

Sheldon, R.A. and Brady, D. (2021). Streamlining Design, Engineering, and Applications of Enzymes for Sustainable Biocatalysis. *Acs Sustainable Chemistry & Engineering.* **9**(24), pp.8032-8052.

Shen, Y. and Bax, A. (2010). Prediction of Xaa-Pro peptide bond conformation from sequence and chemical shifts. *Journal of Biomolecular Nmr.* **46**(3), pp.199-204.

Shen, Y. and Bax, A. (2013). Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *Journal of Biomolecular Nmr.* **56**(3), pp.227-241.

Sjoberg, A. and Hahnhagerdal, B. (1989). beta-Glucose-1-phosphate, a possible mediator for polysaccharide formation in maltose-assimilating lactococcus-lactis. *Applied and Environmental Microbiology.* **55**(6), pp.1549-1554.

Strajbl, M., Shurki, A., Kato, M. and Warshel, A. (2003). Apparent NAC effect in chorismate mutase reflects electrostatic transition state stabilization. *Journal of the American Chemical Society.* **125**(34), pp.10228-10237.

Sun, Z.Y.J., Frueh, D.P., Selenko, P., Hoch, J.C. and Wagner, G. (2005). Fast assignment of N-15-HSQC peaks using high-resolution 3D HNcocaNH experiments with non-uniform sampling. *Journal of Biomolecular Nmr.* **33**(1), pp.43-50.

Thornton, J.M., Todd, A.E., Milburn, D., Borkakoti, N. and Orengo, C.A. (2000). From structure to function: Approaches and limitations. *Nat Struct Biol.* **7**(11s), pp.991-994.

Torbeev, V.Y., Raghuraman, H., Hamelberg, D., Tonelli, M., Westler, W.M., Perozo, E. and Kent, S.B.H. (2011). Protein conformational dynamics in the mechanism of HIV-1 protease catalysis. *Proceedings of the National Academy of Sciences of the United States of America.* **108**(52), pp.20982-20987.

Tremblay, L.W., Zhang, G.F., Dai, J.Y., Dunaway-Mariano, D. and Allen, K.N. (2005). Chemical confirmation of a pentavalent phosphorane in complex with beta-phosphoglucomutase. *Journal of the American Chemical Society.* **127**(15), pp.5298-5299.

Trudeau, D.L. and Tawfik, D.S. (2019). Protein engineers turned evolutionists-the quest for the optimal starting point. *Current Opinion in Biotechnology.* **60**, pp.46-52.

Truhlar, D.G., Garrett, B.C. and Klippenstein, S.J. (1996). Current Status of Transition-State Theory. *J. Phys. Chem.* **100**(31), pp.12771-12800.

Tsang, W.Y., Amyes, T.L. and Richard, J.P. (2008). A substrate in pieces: Allosteric activation of glycerol 3-phosphate dehydrogenase (NAD(+)) by phosphite dianion. *Biochemistry.* **47**(16), pp.4575-4582.

Tsang, W.Y., Wood, B.M., Wong, F.M., Wu, W., Gerlt, J.A., Amyes, T.L. and Richard, J.P. (2012). Proton transfer from C-6 of uridine 5'-monophosphate catalyzed by orotidine 5'-monophosphate decarboxylase: Formation and stability of a vinyl carbanion intermediate and the effect of a 5-fluoro substituent. *Journal of the American Chemical Society*. **134**(35), pp.14580–14594.

Tugarinov, V., Hwang, P.M. and Kay, L.E. (2004). Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annual Review of Biochemistry.* **73**, pp.107-146.

Vagin, A. and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *Journal of Applied Crystallography.* **30**, pp.1022-1025.

Vila, J.A. (2021). Thoughts on the Protein's Native State. *Journal of Physical Chemistry Letters.* **12**(25), pp.5963-5966.

Wang, W.R., Cho, H.S., Kim, R., Jancarik, J., Yokota, H., Nguyen, H.H., Grigoriev, I.V., Wemmer, D.E. and Kim, S.H. (2002). Structural characterization of the reaction pathway in phosphoserine phosphatase: Crystallographic "snapshots" of intermediate states. *Journal of Molecular Biology.* **319**(2), pp.421-431.

Warshel, A. (1998). Electrostatic Origin of the Catalytic Power of Enzymes and the Role of Preorganized Active Sites. *J Biol Chem.* **273**(42), pp.27035-27038.

Warshel, A., Sharma, P.K., Kato, M., Xiang, Y., Liu, H. and Olsson, M.H.M. (2006). Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **106**(8), pp.3210-3235.

Warshel, A., Sharma, P.K., Kato, M., Xiang, Y., Liu, H.B. and Olsson, M.H.M. (2006). Electrostatic basis for enzyme catalysis. *Chemical Reviews.* **106**(8), pp.3210-3235.

Weber, G. (1975). Energetics of Ligand Binding to Proteins. **29**(C), pp.1-83.

Webster, C.E. (2004). High-energy intermediate or stable transition state analogue: Theoretical perspective of the active site and mechanism of beta-phosphoglucomutase. *Journal of the American Chemical Society.* **126**(22), pp.6840-6841.

Wegner, A., Meiser, J., Weindl, D. and Hiller, K. (2015). How metabolites modulate metabolic flux. *Current Opinion in Biotechnology.* **34**, pp.16-22.

Wei, G.H., Xi, W.H., Nussinov, R. and Ma, B.Y. (2016). Protein Ensembles: How Does Nature Harness Thermodynamic Fluctuations for Life? The Diverse Functional Roles of Conformational Ensembles in the Cell. *Chemical Reviews.* **116**(11), pp.6516-6551.

Welch, G.R. (1977). Role of organized multienzyme systems in cellular-metabolism - general synthesis. *Progress in Biophysics & Molecular Biology.* **32**(2), pp.103-191.

Wen, F., Nair, N.U. and Zhao, H.M. (2009). Protein engineering in designing tailored enzymes and microorganisms for biofuels production. *Current Opinion in Biotechnology.* **20**(4), pp.412-419.

Whittington, A.C., Larion, M., Bowler, J.M., Ramsey, K.M., Brueschweiler, R. and Miller, B.G. (2015). Dual allosteric activation mechanisms in monomeric human glucokinase. *Proceedings of the National Academy of Sciences of the United States of America.* **112**(37), pp.11553-11558.

Williamson, M.P. (2013). Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc.* **73**, pp.1-16.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A. and Wilson, K.S. (2011). Overview of the CCP4 suite and current developments. *Acta Cryst. D.* **67**(4), pp.235-242.

Winter, G. (2010). xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography.* **43**, pp.186-190.

Wishart, D.S., Sykes, B.D. and Richards, F.M. (1991). Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol.* **222**(2), pp.311-333.

Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E.Z. and Kern, D. (2004). Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nature Structural & Molecular Biology.* **11**(10), pp.945-949.

Wolfenden, R. and Snider, M.J. (2001). The depth of chemical time and the power of enzymes as catalysts. *Accounts of Chemical Research.* **34**(12), pp.938-945.

Wong, M., Khirich, G. and Loria, J.P. (2013). What's in Your Buffer? Solute Altered Millisecond Motions Detected by Solution NMR. *Biochemistry.* **52**(37), pp.6548-6558.

Wood, H.P., Baxter, N.J., Cruz-Navarrete, F.A., Trevitt, C.R., Hounslow, A.M. and Waltho, J.P. (2021). Enzymatic production of β-glucose 1,6-bisphosphate through manipulation of catalytic magnesium coordination. *Green Chemistry.* **23**(2), pp.752-762.

Wood, H.P., Cruz-Navarrete, F.A., Baxter, N.J., Trevitt, C.R., Robertson, A.J., Dix, S.R., Hounslow, A.M., Cliff, M.J. and Waltho, J.P. (2020). Allomorphy as a mechanism of post-translational control of enzyme activity. *Nature communications.* **11**(1), pp.5538-5538.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. and Zhang, Y. (2014). The I-TASSER suite: Protein structure and function prediction. *Nat Methods.* **12**(1), pp.7-8.

Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., Toyohara, K., Miyamoto, K., Kimura, Y. and Oda, K. (2016). A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science.* **351**(6278), pp.1196-1199.

Zamberlin, S., Antunac, N., Havranek, J. and Samarzija, D. (2012). Mineral elements in milk and dairy products. *Mljekarstvo.* **62**(2), pp.111-125.

Zeeb, M. and Balbach, J. (2004). Protein folding studied by real-time NMR spectroscopy. *Methods.* **34**(1), pp.65-74.

Zhang, C., Allen, K.N. and Dunaway-Mariano, D. (2018). Mechanism of Substrate Recognition and Catalysis of the Haloalkanoic Acid Dehalogenase Family Member alpha-Phosphoglucomutase. *Biochemistry.* **57**(30), pp.4504-4517.

Zhang, G.F., Dai, J., Wang, L.B., Dunaway-Mariano, D., Tremblay, L.W. and Allen, K.N. (2005). Catalytic cycling in beta-phosphoglucomutase: A kinetic and structural analysis. *Biochemistry.* **44**(27), pp.9404-9416.

Zhao, S., Xu, W., Jiang, W., Yu, W., Lin, Y., Zhang, T., Yao, J., Zhou, L., Zeng, Y., Li, H., Li, Y., Shi, J., An, W., Hancock, S.M., He, F., Qin, L., Chin, J., Yang, P., Chen, X., Lei, Q., Xiong, Y. and Guan, K.-L. (2010). Regulation of Cellular Metabolism by Protein Lysine Acetylation. *Science.* **327**(5968), pp.1000-1004.